manuscripta mathematica manuscript No.
(will be inserted by the editor)

Michael Drmota · Gerhard Larcher · Friedrich Pillichshammer

Precise distribution properties of the van der Corput sequence and related sequences

the date of receipt and acceptance should be inserted later

Abstract. The discrepancy is a quantitative measure for the irregularity of distribution of sequences in the unit interval. This article is devoted to the precise study of L_p–discrepancies of a special class of digital $(0,1)$–sequences containing especially the van der Corput sequence. We show that within this special class of digital $(0,1)$–sequences over $\mathbb{Z}_2$ the van der Corput sequence is the worst distributed sequence with respect to L_2–discrepancy. Further we prove that the L_p–discrepancies of the van der Corput sequence satisfy a central limit theorem and we study the discrepancy function of $(0,1)$–sequences pointwise.

Key words. Discrepancy – van der Corput sequence – digital sequence.

1. Introduction

For a sequence $x_0, x_1, \ldots$ of points in the 1-dimensional unit interval $[0,1)$ the discrepancy function Δ_N, $N \in \mathbb{N}$, is defined as

$$\Delta_N(\alpha) := A_N([0,\alpha)) - N\alpha$$

for $0 \le \alpha \le 1$, where $A_N([0,\alpha))$ denotes the number of indices i satisfying $0 \le i \le N-1$ and $x_i \in [0,\alpha)$. Now the L_p–discrepancy $L_{p,N}$, $p \ge 1$, of the sequence is defined as the L_p–norm of the discrepancy function Δ_N divided by N and is a measure for the irregularity of distribution of the first N points of the sequence in $[0,1)$ (see for example [2] or [8]), i.e., for $1 \le p < \infty$ we set

$$L_{p,N} = L_{p,N}(x_0, x_1, \ldots) := \frac{1}{N}\left(\int_0^1 |\Delta_N(\alpha)|^p \, d\alpha\right)^{\frac{1}{p}}.$$

M. Drmota: Institut für Diskrete Mathematik und Geometrie, TU Wien, Wiedner Hauptstrasse 8–10, A-1040 Wien, Austria. e-mail: michael.drmota@tuwien.ac.at

G. Larcher: Institut für Analysis, Universität Linz, Altenbergstraße 69, A-4040 Linz, Austria. e-mail: gerhard.larcher@jku.at

F. Pillichshammer: Institut für Analysis, Universität Linz, Altenbergstraße 69, A-4040 Linz, Austria. e-mail: friedrich.pillichshammer@jku.at

For $p = \infty$ we get the usual star discrepancy

$$D_N^* = D_N^*(x_0, x_1, \ldots) := \frac{1}{N} \sup_{0 \le \alpha \le 1} |\Delta_N(\alpha)|$$

of the sequence.

We consider the discrepancy of a special class of sequences in $[0,1)$, namely of so-called digital $(0,1)$–sequences. Digital $(0,1)$–sequences or more generally digital (t,s)–sequences were introduced by Niederreiter [10, 11] and they provide at the moment the most efficient method to generate sequences with small discrepancy.

We consider the discrepancy of digital $(0,1)$–sequences over $\mathbb{Z}_2$.

Definition 1. Choose a $\mathbb{N} \times \mathbb{N}$ matrix C over $\mathbb{Z}_2$ such that every left upper $m \times m$ matrix $C(m)$ has full rank over $\mathbb{Z}_2$. For $n \ge 0$ let $n = n_0 + n_1 2 + n_2 2^2 + \cdots$ be the base 2 representation of n. Then multiply the vector $\mathbf{n} = (n_0, n_1, \ldots)^T$ with the matrix C,

$$C\mathbf{n} =: (y_1(n), y_2(n), \ldots)^T \in \mathbb{Z}_2^\infty$$

and set

$$x_n := \frac{y_1(n)}{2} + \frac{y_2(n)}{2^2} + \cdots .$$

Every sequence constructed in this way is called digital $(0,1)$–sequence over $\mathbb{Z}_2$.

The most famous digital $(0,1)$–sequence over $\mathbb{Z}_2$ is the well known van der Corput sequence which is generated by the $\mathbb{N} \times \mathbb{N}$ identity matrix.

Niederreiter [10, 11] proved that for any digital $(0,1)$–sequence over $\mathbb{Z}_2$ we have

$$ND_N^* \le \frac{\log N}{2 \log 2} + O(1)$$

for any $N \in \mathbb{N}$. This was improved in [13]: for the star discrepancy D_N^* of any digital $(0,1)$–sequence over $\mathbb{Z}_2$ for every $N \in \mathbb{N}$ we have

$$ND_N^* \le N\widetilde{D}_N^* \le \frac{\log N}{3 \log 2} + 1,$$

where $\widetilde{D}_N^*$ denotes the star discrepancy of the van der Corput sequence (for the second inequality see [1]). Hence the van der Corput sequence is the worst distributed digital $(0,1)$–sequence over $\mathbb{Z}_2$ with respect to star discrepancy.

There is also a well known lower bound due to Schmidt [17] which tells us that for any sequence in $[0,1)$ for the star discrepancy D_N^* we have

$$ND_N^* \ge \frac{\log N}{66 \log 4}$$

for infinitely many values of $N \in \mathbb{N}$. Hence the star discrepancy of any digital $(0,1)$–sequence over $\mathbb{Z}_2$ is of best possible order in N.

A similar result is known for the L_2–discrepancy due to Roth [15]. There exists a constant $c > 0$ such that for the L_2–discrepancy of any sequence in $[0,1)$ we have

$$NL_{2,N} \geq c\sqrt{\log N}$$

for infinitely many values of $N \in \mathbb{N}$. It is not the case that digital $(0,1)$–sequences give in general the best order of L_2–discrepancy.

It is the aim of this paper to study precise distribution properties of a certain class of digital $(0,1)$-sequences over $\mathbb{Z}_2$. This class contains two of the most important and well known sequences namely the van der Corput sequence and the sequence generated by the matrix

$$C_1 = \begin{pmatrix} 1\,1\,1\,\dots \\ 0\,1\,1\,\dots \\ 0\,0\,1\,\dots \\ \vdots\;\vdots\;\vdots\;\ddots \end{pmatrix}. \tag{1}$$

We will call this sequence in the following the upper-**1**-sequence. The importance of this sequence is worked out in [9] or in [13].

The class of sequences we are studying is the class generated by left upper triangular matrices

$$C = \begin{pmatrix} \mathbf{a}_1 & & \\ 0 & \mathbf{a}_2 & \\ 0 & 0 & \mathbf{a}_3 \\ \dots & \dots & \dots \end{pmatrix} \tag{2}$$

with

$$\mathbf{a}_i = (1,0,0,\dots) \quad \text{or} \quad \mathbf{a}_i = (1,1,1,\dots) \quad \text{for } i \in \mathbb{N}.$$

The reason why we study these matrices is, that they seem to represent all distribution properties that can occur within digital $(0,1)$–sequences over $\mathbb{Z}_2$. These sequences are a subclass of sequences generated by NUT matrices studied by Faure in [5]. A NUT matrix over $\mathbb{Z}_2$ is a nonsingular upper triangular matrix over $\mathbb{Z}_2$ (hence the diagonal entries are all 1).

In the following section we present the main results of this paper. The proofs of these results are then given in Sections 3–7.

2. The results

First we show that our type of sequences show a very regular behavior of L_2–discrepancy. We start with the van der Corput sequence. This sequence has the highest L_2–discrepancy within all NUT sequences.

Theorem 1. *Let ω be a digital $(0,1)$–sequence over $\mathbb{Z}_2$ generated by a NUT matrix and let ω_{vdC} be the van der Corput sequence. Then we have*

$$(NL_{2,N}(\omega))^2 \leq (NL_{2,N}(\omega_{\text{vdC}}))^2 \leq \left(\frac{\log N}{6\log 2}\right)^2 + O(\log N) \tag{3}$$

and

$$\limsup_{N\to\infty} \sup_{\omega} \frac{NL_{2,N}(\omega)}{\log N} = \frac{1}{6\log 2} \tag{4}$$

where the sup *is extended over all digital* $(0,1)$*–sequences* ω *generated by a NUT matrix.*

In the average the L_2–discrepancy of the van der Corput sequence develops at a high level but very regular. For the van der Corput sequence this fact not only holds for the L_2–discrepancy but for all L_p–discrepancies and also for the star discrepancy D_N^*.

Theorem 2. *Let* D_N^* *denote the star discrepancy of the first* N *points of the van der Corput sequence. Then we have for every real* y

$$\frac{1}{M}\#\left\{N < M : ND_N^* \le \frac{1}{4}\log_2 N + y\frac{1}{4\sqrt{3}}\sqrt{\log_2 N}\right\} = \Phi(y) + o(1),$$

where

$$\Phi(y) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{t^2}{2}}\,dt$$

denotes the normal distribution function, that is, the star discrepancy satisfies a central limit theorem.

Remark 1. Faure [5, Theorem 2] proved that the extreme discrepancy $D_N(\omega)$ has the same value for any digital $(0,1)$–sequence ω over $\mathbb{Z}_2$ which is generated by a NUT matrix. Further it is well known (see [1, Théorème 1]) that for the van der Corput sequence ω_{vdC} we have $D_N(\omega_{\mathrm{vdC}}) = D_N^*(\omega_{\mathrm{vdC}})$. Hence the result from Theorem 2 remains valid if we replace the star discrepancy D_N^* of the van the Corput sequence by the extreme discrepancy of a digital $(0,1)$–sequence over $\mathbb{Z}_2$ which is generated by a NUT matrix. (For the definition of extreme discrepancy see for example [2,8,11].)

Theorem 3. *Let* $L_{p,N}$ *denote the* L_p*–discrepancy of the first* N *points of the van der Corput sequence. Then for every* $p \ge 1$ *and for every real* y *we have*

$$\frac{1}{M}\#\left\{N < M : NL_{p,N} \le \frac{1}{8}\log_2 N + y\frac{1}{8\sqrt{3}}\sqrt{\log_2 N}\right\} = \Phi(y) + o(1),$$

that is, the L_p*–discrepancy satisfies a central limit theorem.*

This regularity (for L_2–discrepancy) remains stable in a slightly weaker form for all sequences from our class as long as the density of $(1,1,1,\ldots)$ rows is not equal to 1. The average behavior of L_2–discrepancy becomes better with growing density of $(1,1,1,\ldots)$ rows.

Before we state the result we introduce the following notation. For a matrix C of the form (2), the number of $(1,0,0,\ldots)$ rows among the first m rows is denoted by $h(m)$.

Theorem 4. *Let ω be a digital $(0,1)$–sequence over $\mathbb{Z}_2$ generated by a matrix C of the form (2), where the number $h(m)$ of $(1,0,0,\ldots)$ rows among the first m rows satisfies $h(m)/\sqrt{m} \to \infty$ as $m \to \infty$. Then for any $\varepsilon > 0$ we have*

$$\lim_{m\to\infty} \frac{1}{2^m} \# \left\{ N < 2^m : (1-\varepsilon)\frac{h(m)}{8N} < L_{2,N}(\omega) < (1+\varepsilon)\frac{h(m)}{8N} \right\} = 1.$$

If $h(m) = m + o(m)$, then one can be even much more precise. It turns out that the asymptotic behaviour is almost the same as for the van der Corput sequence.

Theorem 5. *Let ω be a digital $(0,1)$–sequence over $\mathbb{Z}_2$ generated by a matrix C of the form (2), where the number $h(m)$ of $(1,0,0,\ldots)$ rows among the first m rows satisfies $h(m) = m + o(m)$. Then for every real y we have*

$$\frac{1}{2^m} \# \left\{ N < 2^m : N L_{2,N}(\omega) \le \frac{1}{8} h(m) + y\frac{\sqrt{m}}{8\sqrt{3}} \right\} = \Phi(y) + o(1),$$

that is, the L_2–discrepancy satisfies a central limit theorem.

It should be noted that there is a central limit theorem if $h(m)/\sqrt{m} \to \infty$ and not only in the case $h(m) = m + o(m)$. However, it seems to be difficult to get the precise normalization (compare with Section 5).

If $h(m)/\sqrt{m} \to 0$ then the behaviour of $L_{2,N}$ is completely different.

Theorem 6. *Let ω be a digital $(0,1)$–sequence over $\mathbb{Z}_2$ generated by a matrix C of the form (2), where the number $h(m)$ of $(1,0,0,\ldots)$ rows among the first m rows satisfies $h(m)/\sqrt{m} \to 0$ as $m \to \infty$. Then there exists a constant $d > 0$ such that for every real $y \ge 1/\sqrt{48}$ we have*

$$\frac{1}{2^m} \# \left\{ N < 2^m : N L_{2,N}(\omega) \le y\sqrt{m} \right\} = 2\Phi\left(2\sqrt{\frac{y^2 - 1/48}{d}} \right) - 1 + o(1).$$

In particular for every $\varepsilon > 0$ we have

$$\lim_{m\to\infty} \frac{1}{2^m} \# \left\{ N < 2^m : N L_{2,N}(\omega) < (\log N)^{\frac{1}{2}+\varepsilon} \right\} = 1.$$

We would hope from the result of Theorem 6 that the L_2–discrepancy of the upper-**1**-sequence is small for all N. This however is not the case.

Theorem 7. *For the upper-**1**-sequence ω we have*

$$N L_{2,N}(\omega) > c \log N$$

for infinitely many $N \in \mathbb{N}$, where $c > 0$ is an absolute constant.

At a next step we start an even more detailed investigation. We do not study averages or extreme values of the discrepancy function, but we consider it pointwise.

From [16] or [18] it is known that for the van der Corput sequence the discrepancy function

$$\Delta_N^{\mathrm{vdC}}(\alpha) = A_N([0,\alpha)) - N\alpha$$

is bounded in N if and only if α has a finite base 2 representation $(*)$. Here the "if-part" is almost trivial, whereas the "only if-part" needs some investigation. In the following we will give a very strict quantitative version of this fact.

For $\alpha \in [0,1)$ with base 2 representation $\alpha = a_1/2 + a_2/2^2 + \cdots$ and $m \in \mathbb{N}$ define

$$f_\alpha(m) := \#\{j < m \,:\, a_j \neq a_{j+1}\}$$

and

$$g_\alpha(m) := \frac{1}{2}\sum_{u=0}^{m-1} \|2^u\alpha\|,$$

where $\|\cdot\|$ is the distance to the nearest integer function. Then we have

Theorem 8. *Let $\Delta_N^{\mathrm{vdC}}(\alpha)$ be the discrepancy function of the first N points of the van der Corput sequence. Let $\alpha \in [0,1)$ with an infinite base 2 representation. Then for $\varepsilon > 0$ we have*

$$\lim_{m\to\infty} \frac{1}{2^m}\#\left\{N \le 2^m : (1-\varepsilon)g_\alpha(m) < \Delta_N^{\mathrm{vdC}}(\alpha) < (1+\varepsilon)g_\alpha(m)\right\} = 1.$$

The functions f_α and g_α are intimately connected (see Lemma 5) so that from this theorem we obtain the following corollary.

Corollary 1 *Let $\Delta_N^{\mathrm{vdC}}(\alpha)$ be the discrepancy function of the first N points of the van der Corput sequence. Let $\alpha \in [0,1)$ with an infinite base 2 representation. Then for $\varepsilon > 0$ we have*

$$\lim_{m\to\infty} \frac{1}{2^m}\#\left\{N \le 2^m : (1-\varepsilon)f_\alpha(m) < \Delta_N^{\mathrm{vdC}}(\alpha)\right\} = 1.$$

Note that Corollary 1 is a strong quantitative version of $(*)$.

The quantity f_α plays also an important role for the discrepancy of arbitrary digital $(0,1)$–sequences over $\mathbb{Z}_2$. This is shown in the following intermediate result.

Theorem 9. *Let $\Delta_N(\alpha)$ be the discrepancy function of the first N points of a digital $(0,1)$–sequence over $\mathbb{Z}_2$. For $\alpha \in [0,1)$ and $N \in \mathbb{N}$ we have*

$$|\Delta_N(\alpha)| \le f_\alpha(\lfloor \log_2 N \rfloor) + 4.$$

Remark 2. From Corollary 1 and Theorem 9 we obtain a result from Faure [4, Corollaire 2].

Following the proofs of $(*)$ in [16] or [18] it is an easy example to extend this proof to the case of a certain type of digital $(0,1)$–sequences over $\mathbb{Z}_2$.

Definition 2. A digital $(0,1)$–sequence over $\mathbb{Z}_2$ will be called "of finite row type" if every row of the generating matrix contains only finitely many ones.

We then have

Theorem 10. *For the discrepancy function $\Delta_N^{\mathrm{fr}}(\alpha)$ of a finite row type $(0,1)$–sequence over $\mathbb{Z}_2$ we have: $\Delta_N^{\mathrm{fr}}(\alpha)$ is bounded in N if and only if α has a finite base 2 representation.*

This result follows for example by an obvious adaption of the proof of the Theorem in [7].

The problem seems to become more involved for general digital $(0,1)$–sequences over $\mathbb{Z}_2$. For example the upper-**1**-sequence is not of finite row type. We will investigate the boundedness of the discrepancy function $\Delta_N^{\mathbf{1}}(\alpha)$ of this sequence and give a corresponding quantitative result, thereby suggesting that the result of Theorem 10 should hold also for general digital $(0,1)$–sequences over $\mathbb{Z}_2$.

Theorem 11. *The discrepancy function $\Delta_N^{\mathbf{1}}(\beta)$ of the upper-**1**-sequence is bounded if and only if β has a finite base 2 representation. If the base 2 representation of β is infinite, then for every m there exists an $N \leq 2^m$ such that*

$$\Delta_N^{\mathbf{1}}(\beta) \geq \frac{1}{8}\sqrt{f_\beta(m) - 1}.$$

The following sections are devoted to the proofs of the results.

3. The L_2–discrepancy of digital sequences

This section is devoted to give some general results concerning the L_2–discrepancy and to give the proof of Theorem 1. First we recall the following general result.

Lemma 1 (Formula of Koksma). *For any sequence $\omega = (x_n)_{n\geq 0}$ in $[0,1)$ we have*

$$(NL_{2,N}(\omega))^2 = \left(\sum_{n=0}^{N-1}\left(\frac{1}{2} - x_n\right)\right)^2 + \frac{1}{4\pi^2}(NF_N(\omega))^2,$$

where

$$F_N(\omega) := \frac{1}{N}\left(2\sum_{m=1}^{\infty}\frac{1}{m^2}\left|\sum_{n=0}^{N-1} \mathrm{e}^{2\pi i m x_n}\right|^2\right)^{1/2}$$

is the diaphony of the sequence.

Proof. The formula follows by considering the Fourier coefficients of the discrepancy function and an application of Parsevals identity. See [8].□

In the next lemma we give the order of diaphony of a NUT sequence.

Lemma 2. *For the diaphony of a digital* $(0,1)$*–sequence* ω *over* $\mathbb{Z}_2$ *generated by a NUT matrix we have*

$$NF_N(\omega) = O(\sqrt{\log N}).$$

Proof. From [5, Theorem 4] it follows that for the diaphony F_N of a digital $(0,1)$–sequence over $\mathbb{Z}_2$ generated by a NUT matrix we have

$$(NF_N(\omega))^2 = \pi^2 \sum_{j=1}^{\infty} \left\| \frac{N}{2^j} \right\|^2.$$

(We remark that Faure proved this result in a much more general setting.) Choose $m \in \mathbb{N}$ such that $2^{m-1} \leq N < 2^m$. Then we have

$$\begin{aligned}(NF_N(\omega))^2 &= \pi^2 \sum_{j=1}^{m} \left\| \frac{N}{2^j} \right\|^2 + \pi^2 \sum_{j=m+1}^{\infty} \left(\frac{N}{2^j} \right)^2 \\ &\leq \pi^2 \sum_{j=0}^{m-1} \left\| 2^j \frac{N}{2^m} \right\|^2 + \frac{\pi^2}{3} = O(m),\end{aligned}$$

by [9, Theorem 2]. Since $2^{m-1} \leq N < 2^m$, the result follows.□

Recall that Walsh-functions in base 2 can be defined as follows: for a non-negative integer k with base 2 representation $k = k_m 2^m + \cdots + k_1 2 + k_0$ and a real x with (canonical) base 2 representation $x = \frac{x_1}{2} + \frac{x_2}{2^2} + \cdots$ the k-th Walsh function in base 2 is defined as

$$\mathrm{wal}_k(x) := (-1)^{x_1 k_0 + x_2 k_1 + \cdots + x_{m+1} k_m}.$$

Now we consider $\sum_{n=0}^{N-1}(1/2 - x_n)$. From the Walsh series expansion of the function $x \mapsto 1/2 - x$ we obtain

$$\sum_{n=0}^{N-1} \left(\frac{1}{2} - x_n \right) = \sum_{r=0}^{\infty} \frac{1}{2^{r+2}} \sum_{n=0}^{N-1} \mathrm{wal}_{2^r}(x_n).$$

If $(x_n)_{n \geq 0}$ is a digital sequence in base 2, then for any $m \in \mathbb{N}$ we have

$$\sum_{n=0}^{2^m-1} \mathrm{wal}_{2^r}(x_n) = \begin{cases} 0 & \text{if } r < m, \\ 2^m & \text{if } r \geq m. \end{cases}$$

Now choose m such that $2^m \leq N < 2^{m+1}$. Then we obtain

$$\sum_{n=0}^{N-1} \left(\frac{1}{2} - x_n \right) = \frac{1}{2} + \sum_{r=0}^{\infty} \frac{1}{2^{r+2}} \sum_{n=2^m}^{N-1} \mathrm{wal}_{2^r}(x_n).$$

If $r \geq m+1$, we have $\mathrm{wal}_{2^r}(x_n) = 1$ such that

$$\sum_{r=m+1}^{\infty} \frac{1}{2^{r+2}} \sum_{n=2^m}^{N-1} \mathrm{wal}_{2^r}(x_n) = \frac{1}{4}\frac{N-2^m}{2^m}$$

and therefore we obtain

$$\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right) = \frac{1}{2} + \frac{1}{4}\frac{N-2^m}{2^m} + \sum_{r=0}^{m} \frac{1}{2^{r+2}} \sum_{n=2^m}^{N-1} \mathrm{wal}_{2^r}(x_n).$$

Note that $\mathrm{wal}_{2^r}(x_n) = (-1)^{x_n(r+1)}$, where $x_n(r)$ denotes the r-th digit in the binary digit expansion of x_n.

Assume now that the digital sequence $(x_n)_{n\geq 0}$ is generated by a NUT matrix $C = (c_{i,j})_{i,j=1,2,\ldots}$. Especially we have $c_{i,1} = \cdots = c_{i,i-1} = 0$ and $c_{i,i} = 1$ for all $i \geq 1$. Hence the r-th coordinate of x_n is given by

$$x_n(r) = c_{r,1}n_0 \oplus c_{r,2}n_1 \oplus c_{r,3}n_2 \oplus \ldots,$$

where $n = n_0 + n_1 2 + n_2 2^2 + \cdots$. Now we have

(a) if $r = m$:

$$\frac{1}{2^{m+2}} \sum_{n=2^m}^{N-1} \mathrm{wal}_{2^m}(x_n) = \frac{1}{2^{m+2}} \sum_{n=2^m}^{N-1} (-1)^{c_{m+1,m+1}} = \frac{1}{4}\left(1-\frac{N}{2^m}\right).$$

(b) if $r < m$:

$$\sum_{n=2^m}^{N-1} \mathrm{wal}_{2^r}(x_n) = \sum_{n=2^m}^{N-1} (-1)^{c_{r+1,r+1}n_r + \cdots + c_{r+1,m}n_{m-1} + c_{r+1,m+1}}.$$

Define now $U := N - 2^m = N_0 + N_1 2 + \cdots + N_{m-1}2^{m-1}$. Then we obtain

$$\begin{aligned}
&\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right) \\
&= \frac{1}{2} + \sum_{r=0}^{m-1} \frac{1}{2^{r+2}} (-1)^{c_{r+1,m+1}} \sum_{n=2^m}^{N-1} (-1)^{c_{r+1,r+1}n_r + \cdots + c_{r+1,m}n_{m-1}} \\
&= \frac{1}{2} + \sum_{r=0}^{m-1} \frac{1}{2^{r+2}} (-1)^{c_{r+1,m+1}} \sum_{n=0}^{U-1} (-1)^{c_{r+1,r+1}n_r + \cdots + c_{r+1,m}n_{m-1}}.
\end{aligned}$$

Now we need the following lemma.

Lemma 3. *Let the non-negative integer U have binary expansion $U = U_0 + U_1\,2 + \cdots + U_{m-1}\,2^{m-1}$. For any non-negative integer $n \leq U-1$ let $n = n_0 + n_1\,2 + \cdots + n_{m-1}\,2^{m-1}$ be the binary representation of n. For $0 \leq p \leq m-1$*

let $U(p) := U_0 + \cdots + U_p\, 2^p$. Let $b_0, b_1, \ldots, b_{m-1}$ be arbitrary elements of $\mathbb{Z}_2$, not all zero. Then

$$\sum_{n=0}^{U-1} (-1)^{b_0 n_0 + \cdots + b_{m-1} n_{m-1}} = (-1)^{b_{w+1} U_{w+1} + \cdots + b_{m-1} U_{m-1}} (2^w + (-1)^{U_w} (U(w) - 2^w)),$$

where w is minimal such that $b_w = 1$.

Proof. The result easily follows from splitting up the sum.

$$\begin{aligned}
&\sum_{n=0}^{U-1} (-1)^{b_0 n_0 + \cdots + b_{m-1} n_{m-1}} \\
&= \sum_{n=0}^{2^{w+1}(U_{w+1} + \cdots + U_{m-1} 2^{m-w-2}) - 1} (-1)^{n_w} (-1)^{b_{w+1} n_{w+1} + \cdots + b_{m-1} n_{m-1}} \\
&\quad + \sum_{n=0}^{U(w)-1} (-1)^{n_w} (-1)^{b_{w+1} U_{w+1} + \cdots + b_{m-1} U_{m-1}} \\
&= 0 + (-1)^{b_{w+1} U_{w+1} + \cdots + b_{m-1} U_{m-1}} \sum_{n=0}^{U(w)-1} (-1)^{n_w} \\
&= (-1)^{b_{w+1} U_{w+1} + \cdots + b_{m-1} U_{m-1}} \times \begin{cases} U(w) & \text{if } U(w) < 2^w, \\ 2^{w+1} - U(w) & \text{if } U(w) \geq 2^w. \end{cases}
\end{aligned}$$

The result follows.□

Since $c_{r+1,r+1} = 1$ with Lemma 3 we obtain

$$\begin{aligned}
&\sum_{n=0}^{N-1} \left(\frac{1}{2} - x_n \right) \\
&= \frac{1}{2} + \sum_{r=0}^{m-1} \frac{1}{2^{r+2}} (-1)^{c_{r+1,m+1}} (-1)^{c_{r+1,r+2} N_{r+1} + \cdots + c_{r+1,m} N_{m-1}} \times \\
&\qquad \times (2^r + (-1)^{N_r} (N(r) - 2^r)),
\end{aligned} \tag{5}$$

where $2^m \leq N < 2^{m+1}$. Note that

$$\frac{2^r + (-1)^{N_r} (N(r) - 2^r)}{2^{r+1}} = \left\| N(r)/2^{r+1} \right\| = \left\| N/2^{r+1} \right\|$$

and hence we have the following result.

Proposition 1. *For any digital* $(0,1)$*–sequence* $(x_n)_{n\geq 0}$ *over* $\mathbb{Z}_2$ *which is generated by a NUT matrix* $C = (c_{i,j})_{i,j=1,2,\ldots}$ *we have*

$$\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right)$$
$$= \frac{1}{2}\left(1+\sum_{r=0}^{m-1}\|2^{-r-1}N\|(-1)^{c_{r+1,m+1}}(-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,m}N_{m-1}}\right),$$

where $2^m \leq N < 2^{m+1}$.

From Proposition 1 we obtain

Corollary 2 *Let* $(x_n)_{n\geq 0}$ *be a digital* $(0,1)$*–sequence over* $\mathbb{Z}_2$ *which is generated by a NUT matrix and let* $(y_n)_{n\geq 0}$ *be the van der Corput sequence. Then for any* $N \in \mathbb{N}$ *we have*

$$\left|\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right)\right| \leq \sum_{n=0}^{N-1}\left(\frac{1}{2}-y_n\right).$$

Now we can give the proof of Theorem 1.

Proof of Theorem 1. Inequality (3) follows from Lemma 1, Lemma 2, Corollary 2, [9, Theorem 2] and the fact that the diaphony has the same value for any digital $(0,1)$–sequence over $\mathbb{Z}_2$ generated by a NUT matrix (this follows from Faure's formula [5, Theorem 4]). Equality (4) follows from (3) together with [13, Corollary 1].□

4. The van der Corput sequence

From [3, 13] we know that for the discrepancy function of the van der Corput sequence ω_{vdC} we have

$$\Delta_N^{\mathrm{vdC}}(\alpha) \geq 0$$

for any $\alpha \in [0,1]$. Further in [14] Proinov and Atanassov showed that $L_{1,N}(\omega_{\mathrm{vdC}}) = \frac{1}{2}D_N^*(\omega_{\mathrm{vdC}})$. Now we get

$$D_N^*(\omega_{\mathrm{vdC}}) = 2L_{1,N}(\omega_{\mathrm{vdC}}) = \frac{2}{N}\int_0^1 \Delta_N^{\mathrm{vdC}}(\alpha)d\alpha = \frac{2}{N}\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right).$$

Therefore together with Proposition 1 we obtain the following formula for the star discrepancy of the van der Corput sequence; see also [1].

Proposition 2. *For the star discrepancy* D_N^* *of the first* N *points of the van der Corput sequence* ω_{vdC} *we have*

$$ND_N^*(\omega_{\mathrm{vdC}}) = 1+\sum_{r=1}^{m}\left\|2^{-r}N\right\| = \sum_{r=1}^{\infty}\left\|2^{-r}N\right\|,$$

where $2^m \leq N < 2^{m+1}$.

Now we give the proof of Theorem 2.

Proof of Theorem 2. We will not prove the limit relation of Theorem 2 but

$$\frac{1}{M}\#\left\{N<M : ND_N^* \le \frac{1}{4}\log_2 M + y\frac{1}{4\sqrt{3}}\sqrt{\log_2 M}\right\} = \Phi(y)+o(1). \quad (6)$$

It is easy to see that both limit relations are equivalent. This follows directly by restricting on N with $M/(\log M) \le N \le M$. Of course, N with $N < M/(\log M)$ do not matter in the limit. For the *remaining ones* we have $|\log M - \log N| \ll \log\log M$ and this difference does not matter in the limit, either.

The idea of the proof of (6) is to use Proposition 2 and approximate D_N^* as sum of weakly dependent random variables, in particular they are stronly mixing with $\alpha(k) = O(2^{-k})$.

First, let us consider the (easier) case when $M = 2^m$, that is, the digits N_j of the binary expansions $N = N_0 + 2N_1 + \cdots + 2^{m-1}N_{m-1}$ can be considered as independent random variables that are uniformly distributed on $\{0,1\}$ (by assuming that all numbers $N < 2^m$ have equal probability 2^{-m}). Then

$$U_r = \left\{\frac{N}{2^r}\right\} = \frac{N_{r-1}}{2} + \frac{N_{r-2}}{2^2} + \cdots + \frac{N_0}{2^r}$$

is very close to a random variable that is uniformly distributed on $[0,1)$. In fact, we can add *missing digits* $N_{-1}, N_{-2}, \ldots$ that are independent and uniformly distributed on $\{0,1\}$ and get

$$U_r' = U_r + \sum_{j=1}^{\infty} \frac{N_{-j}}{2^{r+j}}.$$

Then U_r' is exactly uniformly distributed on $[0,1)$ and $|U_r - U_r'| \le 2^{-r}$.

Furthermore, U_r and U_{r+k} get more and more independent as k gets large. More precisely, $U_{r+k} - 2^{-k}U_r$ and U_r are independent. This means, that a slight modification of order 2^{-k} makes U_r and U_{r+k} independent.

Note that

$$ND_N^* = \sum_{r=1}^{m} \|U_r\| + O(1)$$

for all $N < 2^m$. Thus, ND_N^* can be represented (up to a small error term) by a sum of weakly dependent random variables.

We further note that the above *smoothing* by introducing missing digits can be also obtained by considering the following slight variation of the above probability model. Let $\tilde{N}$ be a random variable that is uniformly distributed on $[0,2^m)$ then $\tilde{U}_r = \{\tilde{N}/2^r\}$ is uniformly distributed on $[0,1)$ and the common distribution of $(\tilde{U}_r)_{1\le r\le m}$ is the same as that of $(U_r')_{1\le r\le m}$. Furthermore, if we set $\omega = \tilde{N}/2^m$ and $\tilde{V}_r = U_{m-r+1}$ then ω is uniformly distributed on $[0,1)$ and

$$V_r = \{\omega 2^r\}.$$

Since $|V_r - U_{m-r+1}| \leq 2^{-(k-r+1)}$ we also have

$$ND_N^* = \sum_{r=1}^{m} \|\omega 2^r\| + O(1).$$

Now, by using well know limit theorems for lacunary sequence (compare with [6] and [12], alternatively we can use [19]) it follows that

$$S_m = \sum_{r=1}^{m} \|\omega 2^r\|$$

satisfies a central limit theorem. Since $\mathbf{E}\, S_m = m/4$ and $\mathbf{Var}\, S_m = m/48$ we thus have

$$\frac{S_m - m/4}{\sqrt{m/48}} \to N(0,1)$$

and also, all moments converge. Since $ND_N^* = S_m + O(1)$ we get the same limit relation for ND_N^* if N is uniformly distributed on $\{0,1,\ldots,2^m-1\}$. This proves the theorem for $M = 2^m$.

The general case, where M is not a power of 2 can be reconstructed from the case of $M = 2^m$ by considering moments. For example, if $M = 2^{m_1} + 2^{m_2} + \cdots + 2^{m_s}$ (with $m_1 > m_2 > \cdots > m_s \geq 0$) then

$$\begin{aligned}\sum_{N<M} ND_N^* = \sum_{N<2^{m_1}} D_N^* + \sum_{N<2^{m_2}} (N+2^{m_1})D^*_{N+2^{m_1}} \\ \cdots + \sum_{N<2^{m_s}} (N+2^{m_1}+\cdots+2^{m_{s-1}})D^*_{N+2^{m_1}+\cdots+2^{m_{s-1}}}\end{aligned}$$

By using the representation $ND_N^* = \sum_{r=1}^{\infty} \|N2^{-r}\|$ it directly follows that

$$(N+2^{m_1}+\cdots+2^{m_j})D^*_{N+2^{m_1}+\cdots+2^{m_j}} = ND_N^* + O(j).$$

Hence we can approximate $\sum_{N<M} ND_N^*$ by sums of the form $\sum_{N<2^m} ND_N^*$. After some algebra this shows that all centralized moments of ND_N^* with $N<M$ converge to the moments of the Gaussian distribution. This completes the proof of the general case. □

The central limit theorem for all L_p–discrepancies of the van der Corput sequence (Theorem 3) is a consequence from Theorem 2.

Proof of Theorem 3. As in the proof of Theorem 2 we just work out the case $M = 2^m$. Further, it is sufficient to consider integral p since the mapping $p \mapsto L_{p,N}$ is monotone.

For integral p set

$$d_p := \sum_{r_1,\ldots,r_p=1}^{m} \prod_{i=1}^{p} \left\| \frac{N}{2^{r_i}} \right\| 2^{-k(r_1,\ldots,r_p)},$$

where $k(r_1,\ldots,r_p)$ denotes the number of different r_i's. Then

$$d_p \le (NL_{p,N})^p \le \sum_{k=0}^{p} \binom{p}{k} d_k$$

(see [13]) and consequently

$$(NL_{p,N})^p = \left(\frac{1}{2}\sum_{r=1}^{m} \left\| \frac{N}{2^r} \right\| \right)^p + O(m^{p-1}).$$

Thus, as long as we can assure that

$$\sum_{r=1}^{m} \left\| \frac{N}{2^r} \right\| \gg m \tag{7}$$

then it follows that

$$NL_{p,N} = \frac{1}{2}\sum_{r=1}^{m} \left\| \frac{N}{2^r} \right\| + O(1) = \frac{1}{2} ND_N^* + O(1).$$

Consequently, Theorem 3 follows from Theorem 2 provided (7) holds.

Of course, (7) is not satisfied for all $N < 2^m$. However, by Chebyshev's inequality we have for all $\eta > 0$

$$\frac{1}{2^m}\# \left\{ N < 2^m : \left| \sum_{r=0}^{m-1} \left\| \frac{N}{2^r} \right\| - m/4 \right| \ge \eta \right\} \ll \frac{m}{\eta^2}.$$

Hence, if we use $\eta = m/8$ it follows that (7) holds with at most $O(2^m/m)$ exceptions. But these exceptions do not count in the limit. This completes the proof of Theorem 3. □

5. Sequences generated by special matrices

As announced in the introduction of this paper here we consider digital $(0,1)$–sequences over $\mathbb{Z}_2$ with generator matrices C of the form given in (2). If $\mathbf{a}_i = (1,0,0,\ldots)$, then we say "row i is of type zero" and if $\mathbf{a}_i = (1,1,1,\ldots)$, then we say "row i is of type one". For $m \in \mathbb{N}$ we denote by $h(m)$ the number of zero type rows among the first m rows of the matrix C. (Hence $h(m) \le m$ for any $m \in \mathbb{N}$.)

We start with a lemma that will be useful in the sequel.

Lemma 4. *Let $(x_n)_{n\ge 0}$ be a digital $(0,1)$–sequence over $\mathbb{Z}_2$ generated by a matrix C of the form (2). Then there exist two positive constants d_1, d_2 such that*

$$d_1 2^m m \le \sum_{N=2^m}^{2^{m+1}-1} \left(\sum_{n=0}^{N-1} \left(\frac{1}{2} - x_n \right) \right)^2 - \frac{2^m}{2^6} h(m)^2 \le d_2 2^m m.$$

Proof. With Proposition 1 we get

$$\sum_{N=2^m}^{2^{m+1}-1}\left(\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right)\right)^2$$
$$= \frac{2^m}{4} + \sum_{N=2^m}^{2^{m+1}-1}\frac{1}{2}\sum_{r=0}^{m-1}\left\|\frac{N}{2^{r+1}}\right\|(-1)^{c_{r+1,m+1}}(-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,m}N_{m-1}}$$
$$+\sum_{N=2^m}^{2^{m+1}-1}\frac{1}{4}\sum_{r,s=0}^{m-1}\left\|\frac{N}{2^{r+1}}\right\|\cdot\left\|\frac{N}{2^{s+1}}\right\|$$
$$\times(-1)^{c_{r+1,m+1}}(-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,m}N_{m-1}}$$
$$\times(-1)^{c_{s+1,m+1}}(-1)^{c_{s+1,s+2}N_{s+1}+\cdots+c_{s+1,m}N_{m-1}}.$$

First we consider

$$\Sigma_1 := \sum_{N=2^m}^{2^{m+1}-1}\frac{1}{2}\sum_{r=0}^{m-1}\left\|\frac{N}{2^{r+1}}\right\|(-1)^{c_{r+1,m+1}}(-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,m}N_{m-1}}$$
$$= \frac{1}{2}\sum_{r=0}^{m-1}(-1)^{c_{r+1,m+1}}\sum_{N_0,\ldots,N_{m-1}=0}^{1}\left\|\frac{N}{2^{r+1}}\right\|$$
$$\times(-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,m}N_{m-1}}$$
$$= \frac{1}{2}\sum_{\substack{r=0\\ \text{row } r+1 \text{ is of type zero}}}^{m-1}\sum_{N_0,\ldots,N_{m-1}=0}^{1}\left\|\frac{N}{2^{r+1}}\right\|$$
$$-\frac{1}{2}\sum_{\substack{r=0\\ \text{row } r+1 \text{ is of type one}}}^{m-1}\sum_{N_0,\ldots,N_{m-1}=0}^{1}\left\|\frac{N}{2^{r+1}}\right\|(-1)^{N_{r+1}+\cdots+N_{m-1}}.$$

Since

$$\sum_{N_0,\ldots,N_{m-1}=0}^{1}\left\|\frac{N}{2^{r+1}}\right\| = 2^{m-2}$$

and

$$\sum_{N_0,\ldots,N_{m-1}=0}^{1}\left\|\frac{N}{2^{r+1}}\right\|(-1)^{N_{r+1}+\cdots+N_{m-1}} = \begin{cases}0 & \text{if } r \le m-2,\\ 2^{m-2} & \text{if } r = m-1,\end{cases}$$

we obtain

$$\Sigma_1 = 2^{m-3}(h(m)-y_m),$$

where $y_m = 1$ if row m is of type 1 and $y_m = 0$ otherwise. Hence

$$\sum_{N=2^m}^{2^{m+1}-1}\left(\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right)\right)^2 = \frac{2^m}{4}+\frac{2^m}{2^3}(h(m)-y_m)+\frac{1}{4}\sum_{r,s=0}^{m-1}\Sigma(r,s), \quad (8)$$

where

$$\begin{aligned}
&\Sigma(r,s) \\
&:= \sum_{N=2^m}^{2^{m+1}-1} \left\| \frac{N}{2^{r+1}} \right\| \cdot \left\| \frac{N}{2^{s+1}} \right\| (-1)^{c_{r+1,m+1}} (-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,m}N_{m-1}} \\
&\qquad\qquad \times (-1)^{c_{s+1,m+1}} (-1)^{c_{s+1,s+2}N_{s+1}+\cdots+c_{s+1,m}N_{m-1}}.
\end{aligned}$$

Assume first that $r = s$. Then we have

$$\begin{aligned}
\Sigma(r,r) &= \sum_{N=2^m}^{2^{m+1}-1} \left\| \frac{N}{2^{r+1}} \right\|^2 = \sum_{N_0,\ldots,N_r=0}^{1} \left\| \frac{N_r}{2} + \frac{N_{r-1}}{2^2} + \cdots + \frac{N_0}{2^{r+1}} \right\|^2 \frac{2^m}{2^{r+1}} \\
&= \frac{2^m}{2^{r+1}} \sum_{N_0,\ldots,N_{r-1}=0}^{1} \left(\left(\frac{N_{r-1}}{2^2} + \cdots + \frac{N_0}{2^{r+1}} \right)^2 \right. \\
&\qquad\qquad \left. + \left(\frac{1}{2} - \frac{N_{r-1}}{2^2} - \cdots - \frac{N_0}{2^{r+1}} \right)^2 \right) \\
&= \frac{2^m}{2^{r+1}} \sum_{N=0}^{2^r-1} \left(\left(\frac{N}{2^{r+1}} \right)^2 + \left(\frac{1}{2} - \frac{N}{2^{r+1}} \right)^2 \right) = \frac{2^m}{3 \cdot 2^{2r+3}} (2^{2r+1} + 1).
\end{aligned}$$

Assume now that $r < s$. Then we have

$$\begin{aligned}
&\Sigma(r,s) = (-1)^{c_{r+1,m+1}+c_{s+1,m+1}} \\
&\quad \times \sum_{N_0,\ldots,N_s=0}^{1} \left\| \frac{N}{2^{r+1}} \right\| \cdot \left\| \frac{N}{2^{s+1}} \right\| (-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,s+1}N_s} \\
&\quad \times \sum_{N_{s+1},\ldots,N_{m-1}=0}^{1} (-1)^{N_{s+1}(c_{r+1,s+2}+c_{s+1,s+2})+\cdots+N_{m-1}(c_{r+1,m}+c_{s+1,m})}.
\end{aligned}$$

If $s < m-1$ then we have

$$\begin{aligned}
\Sigma(r,s) &= (-1)^{c_{r+1,m+1}+c_{s+1,m+1}} \frac{2^m}{2^{s+1}} \\
&\quad \times \sum_{N_0,\ldots,N_s=0}^{1} \left\| \frac{N}{2^{r+1}} \right\| \cdot \left\| \frac{N}{2^{s+1}} \right\| (-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,s+1}N_s},
\end{aligned}$$

if $c_{r+1,j} \oplus c_{s+1,j} = 0$ for all $s+2 \le j \le m$. But in this case we also have $c_{r+1,m+1} \oplus c_{s+1,m+1} = 0$ and hence we have

$$\begin{aligned}
\Sigma(r,s) &= \frac{2^m}{2^{s+1}} \sum_{N_0,\ldots,N_s=0}^{1} \left\| \frac{N}{2^{r+1}} \right\| \cdot \left\| \frac{N}{2^{s+1}} \right\| (-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,s+1}N_s} \\
&= \frac{2^m}{2^{s+1}} \sum_{N_0,\ldots,N_r=0}^{1} \left\| \frac{N}{2^{r+1}} \right\| \\
&\quad \times \sum_{N_{r+1},\ldots,N_s=0}^{1} \left\| \frac{N}{2^{s+1}} \right\| (-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,s+1}N_s}.
\end{aligned}$$

Otherwise, i.e., if $c_{r+1,j} \oplus c_{s+1,j} = 1$ for all $s+2 \le j \le m$, we have

$$\Sigma(r,s) = 0.$$

Let $c_{r+1,j} \oplus c_{s+1,j} = 0$ for all $s+2 \le j \le m$. We consider two cases:

1. If row $r+1$ is of type one, then we have

$$\begin{aligned}
&\sum_{N_{r+1},\ldots,N_s=0}^{1} \left\| \frac{N}{2^{s+1}} \right\| (-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,s+1}N_s} \\
&= \sum_{N_{r+1},\ldots,N_s=0}^{1} \left\| \frac{N}{2^{s+1}} \right\| (-1)^{N_{r+1}+\cdots+N_s}.
\end{aligned}$$

We compute the last sum:

$$\begin{aligned}
&\sum_{N_{r+1},\ldots,N_s=0}^{1} \left\| \frac{N}{2^{s+1}} \right\| (-1)^{N_{r+1}+\cdots+N_s} \\
&= \sum_{N_{r+1},\ldots,N_{s-1}=0}^{1} (-1)^{N_{r+1}+\cdots+N_{s-1}} \left(\frac{N_{s-1}}{2} + \cdots + \frac{N_0}{2^s} - \frac{1}{2} \right) \\
&= \sum_{N_{r+1},\ldots,N_{s-1}=0}^{1} (-1)^{N_{r+1}+\cdots+N_{s-1}} \left(\frac{N_{s-1}}{2} + \cdots + \frac{N_0}{2^s} \right) \\
&= -\frac{1}{2} \sum_{N_{r+1},\ldots,N_{s-2}=0}^{1} (-1)^{N_{r+1}+\cdots+N_{s-2}}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&\sum_{N_{r+1},\ldots,N_s=0}^{1} \left\| \frac{N}{2^{s+1}} \right\| (-1)^{N_{r+1}+\cdots+N_s} \\
&= \begin{cases} 0 & \text{if } r \le s-3, \\ -\frac{1}{2} & \text{if } r = s-2, \\ \frac{N_{s-1}}{2} + \cdots + \frac{N_0}{2^s} - \frac{1}{2} & \text{if } r = s-1, \end{cases}
\end{aligned}$$

and hence

$$\Sigma(r,s)=\begin{cases}0 & \text{if } r\le s-3,\\ -2^{m-5} & \text{if } r=s-2,\\ 0 & \text{if } r=s-1.\end{cases}$$

2. If row $r+1$ is of type zero, then we have

$$\Sigma(r,s)=2^{m-1-s}\sum_{N_0,\ldots,N_s=0}^{1}\left\|\frac{N}{2^{r+1}}\right\|\cdot\left\|\frac{N}{2^{s+1}}\right\|.$$

We compute the last sum:

$$\begin{aligned}
&\sum_{N_0,\ldots,N_s=0}^{1}\left\|\frac{N}{2^{r+1}}\right\|\cdot\left\|\frac{N}{2^{s+1}}\right\|\\
&=\sum_{N_0,\ldots,N_s=0}^{1}\left\|\frac{N_r}{2}+\cdots+\frac{N_0}{2^{r+1}}\right\|\cdot\left\|\frac{N_s}{2}+\cdots+\frac{N_0}{2^{s+1}}\right\|\\
&=\sum_{N_0,\ldots,N_{s-1}=0}^{1}\left\|\frac{N_r}{2}+\cdots+\frac{N_0}{2^{r+1}}\right\|\\
&\qquad\times\left(\frac{N_{s-1}}{2^2}+\cdots+\frac{N_0}{2^{s+1}}+\frac{1}{2}-\frac{N_{s-1}}{2^2}-\ldots-\frac{N_0}{2^{s+1}}\right)\\
&=\frac{1}{2}\sum_{N_0,\ldots,N_r=0}^{1}\left\|\frac{N_r}{2}+\cdots+\frac{N_0}{2^{r+1}}\right\|2^{s-1-r}\\
&=2^{s-2-r}\sum_{N_0,\ldots,N_{r-1}=0}^{1}\frac{1}{2}=2^{s-3-r}2^r=2^{s-3}.
\end{aligned}$$

Therefore we have $\Sigma(r,s)=2^{m-4}$.

Assume now that $s=m-1$. Then we have

$$\begin{aligned}
&\Sigma(r,m-1)=(-1)^{c_{r+1,m+1}+c_{m,m+1}}\\
&\times\sum_{N_0,\ldots,N_{m-1}=0}^{1}\left\|\frac{N}{2^{r+1}}\right\|\cdot\left\|\frac{N}{2^m}\right\|(-1)^{c_{r+1,r+2}N_{r+1}+\cdots+c_{r+1,m}N_{m-1}}.
\end{aligned}$$

If row $r+1$ is of type zero, then we have

$$\begin{aligned}
&\Sigma(r,m-1)\\
&=(-1)^{c_{m,m+1}}\sum_{N_0,\ldots,N_{m-1}=0}^{1}\left\|\frac{N_r}{2}+\cdots+\frac{N_0}{2^{r+1}}\right\|\cdot\left\|\frac{N_{m-1}}{2}+\cdots+\frac{N_0}{2^m}\right\|\\
&=\frac{1}{2}(-1)^{c_{m,m+1}}\sum_{N_0,\ldots,N_{m-2}=0}^{1}\left\|\frac{N_r}{2}+\cdots+\frac{N_0}{2^{r+1}}\right\|=(-1)^{c_{m,m+1}}2^{m-4}.
\end{aligned}$$

If row $r+1$ is of type one, then we obtain after some elementary calculations that

$$\Sigma(r, m-1) = \begin{cases} 0 & \text{if } r \leq m-4, \\ (-1)^{c_{m,m+1}} 2^{m-5} & \text{if } r = m-3, \\ 0 & \text{if } r = m-2. \end{cases}$$

Altogether for $r < s$ we have

$$\Sigma(r,s) = \begin{cases} 2^{m-4} & \text{if } c_{r+1,j} \oplus c_{s+1,j} = 0 \quad \forall j \geq s+2,\ s < m-1, \\ & \text{and row } r+1 \text{ is of type zero,} \\ -2^{m-5} & \text{if } c_{r+1,j} \oplus c_{s+1,j} = 0 \quad \forall j \geq s+2,\ s < m-1, \\ & \text{and row } r+1 \text{ is of type one,} \\ & \text{and } r = s-2, \\ (-1)^{c_{m,m+1}} 2^{m-4} & \text{if } s = m-1, \\ & \text{and row } r+1 \text{ is of type zero,} \\ (-1)^{c_{m,m+1}} 2^{m-5} & \text{if } s = m-1, r = m-3 \\ & \text{and row } r+1 \text{ is of type one,} \\ 0 & \text{otherwise.} \end{cases}$$

Therefore we obtain

$$\begin{aligned} \sum_{r,s=0}^{m-1} \Sigma(r,s) &= 2 \sum_{\substack{r,s=0 \\ r<s}}^{m-1} \Sigma(r,s) + \sum_{r=0}^{m-1} \Sigma(r,r) \\ &= 2 \sum_{\substack{r,s=0 \\ r<s \\ c_{r+1,j} \oplus c_{s+1,j} = 0 \ \forall j \geq s+2 \\ \text{row } r+1 \text{ is of type } 0}}^{m-1} \frac{2^m}{2^4} - 2 \sum_{\substack{r=0 \\ c_{r+1,j} \oplus c_{m,j} = 0 \ \forall j \geq 2 \\ \text{row } r+1 \text{ is of type } 0}}^{m-2} \frac{2^m}{2^4} \\ &\quad -2 \sum_{\substack{s=2 \\ c_{s-1,j} \oplus c_{s+1,j} = 0 \ \forall j \geq 2 \\ \text{row } s-1 \text{ is of type } 1}}^{m-1} \frac{2^m}{2^5} + 2 \sum_{\substack{r=0 \\ \text{row } r+1 \text{ is of type } 0}}^{m-2} (-1)^{c_{m,m+1}} \frac{2^m}{2^4} \\ &\quad +2(-1)^{c_{m,m+1}} \frac{2^m}{2^5} q_m + \sum_{r=0}^{m-1} \frac{2^m}{3 \cdot 2^{2r+3}} (2^{2r+1}+1), \end{aligned}$$

where $q_m = 1$ if row $m-2$ is of type one and $q_m = 0$ otherwise. From this we obtain

$$\begin{aligned} \sum_{r,s=0}^{m-1} \Sigma(r,s) &= \frac{2^m}{2^3} \sum_{\substack{r,s=0 \\ r<s \\ \text{row } r+1 \text{ is of type } 0 \\ \text{row } s+1 \text{ is of type } 0}} 1 + O(2^m m) \\ &= \frac{2^m}{2^4} h(m)^2 + O(2^m m). \end{aligned}$$

Inserting this in Eq. (8) we get

$$\sum_{N=2^m}^{2^{m+1}-1} \left(\sum_{n=0}^{N-1} \left(\frac{1}{2} - x_n \right) \right)^2 = \frac{2^m}{2^6} h(m)^2 + O(2^m m)$$

and we have proved the right inequality. On the other hand we obtain

$$\begin{aligned}\sum_{r,s=0}^{m-1} \Sigma(r,s) &\geq \frac{2^m}{2^4}h(m)^2 - 2^{m-3}h(m-1) - 2^{m-4}(m-2-(h(m)-1)) \\ &\quad -2^{m-3}h(m-1) - 2^{m-4} + \frac{2^m}{12}m + \frac{2^{m-1}}{9} - \frac{1}{9\cdot 2^{m+1}} \\ &\geq \frac{2^m}{2^4}h(m)^2 + 2^m m\frac{1}{48} - 2^m h(m)\frac{3}{16} + \frac{2^m}{18} - \frac{1}{9\cdot 2^{m+1}}.\end{aligned}$$

Inserting this in Eq. (8) we get

$$\begin{aligned}&\sum_{N=2^m}^{2^{m+1}-1}\left(\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right)\right)^2 \\ &\geq \frac{2^m}{4} + \frac{2^m}{8}h(m) - \frac{2^m}{8} \\ &\quad + \frac{1}{4}\left(\frac{2^m}{2^4}h(m)^2 + 2^m m\frac{1}{48} - 2^m h(m)\frac{3}{16} + \frac{2^m}{18} - \frac{1}{9\cdot 2^{m+1}}\right) \\ &\geq \frac{2^m}{2^6}h(m)^2 + \frac{2^m}{192}m,\end{aligned}$$

and we are done. □

Remark 3. Note that Lemma 4 also implies that there are two positive constants d_1', d_2' with

$$d_1' 2^m m \leq \sum_{N<2^m}\left(\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right)\right)^2 - \frac{2^m}{2^6}h(m)^2 \leq d_2' 2^m m.$$

We just have to sum up and to note that $0 \leq h(m+1) - h(m) \leq 1$.

By Proposition 1 we have

$$\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right) = \frac{1}{2}\left(1 + \sum_{r=0}^{m-1} s_r \|2^{-r-1}N\|\right),$$

where $s_r = 1$ if row $r+1$ is of type zero and $s_r = (-1)^{1+N_{r+1}+\cdots+N_m}$ if row $r+1$ is of type one. This means that we are in a situation that is quite similar to that of the proof of Theorem 2. We have

$$2\sum_{n=0}^{N-1}\left(\frac{1}{2}-x_n\right) = \sum_{r=1}^{m} s_{m-r}\|\omega 2^r\| + O(1).$$

The *only but important difference* is the appearance of the signs s_r. If row $r+1$ is of type zero the sign s_r is deterministically equal to 1. On the other hand if row $r+1$ is of type one then s_r is a random sign that obtains

± 1 with equal probability $\frac{1}{2}$. Furthermore, s_r is independent of $\|2^{-r-1}N\|$. Finally, as in the proof of Theorem 2 it follows that $s_r\|2^{-r-1}N\|$ constitute weakly dependent random variables. More precisely, $s_r\|2^{-r-1}N\|$ and $s_{r+k}\|2^{-r-k-1}N\|$ can be made independent by making a slight modification of $s_{r+k}\|2^{-r-k-1}N\|$ of order 2^{-k}. This implies that the sequence $s_r\|2^{-r-1}N\|$, $r < m$, is strongly mixing with $\alpha(k) = O(2^{-k})$. Set

$$T_m := \sum_{r=0}^{m-1} s_r \|2^{-r-1}N\|$$

Then we have $\mathbf{E}\, T_m = \frac{1}{4}h(m) + O(1)$[1] and by Lemma 4 it follows that there are two positive constants $\tilde{d}_1, \tilde{d}_2$ with $\tilde{d}_1 m \le \mathbf{Var}\, T_m \le \tilde{d}_2 m$. Hence, by [19] the sum T_m satisfies a central limit theorem.

The exact behaviour of $\mathbf{Var}\, T_m$ depends on the local structure of $h(m)$ and it seems that there is no reasonable expression for $\mathbf{Var}\, T_m$, in particular if $h(m)$ and m are of the same order of magnitude (compare with the proof of Lemma 4). Only if $h(m) = o(m)$ and if $h(m) = m + o(m)$ we get proper asymptotic expansions without any further work.

If $h(m) = m + o(m)$ then we can *neglect* (more or less) the rows of type one and we get $\mathbf{Var}\, T_m = m/48 + o(m)$ similarly to the case of the van der Corput sequence.

If $h(m) = o(m)$ then almost all signs are present. For a moment, let us assume that all rows are of type one, that is $h(m) = 0$. By definition it follows that the joint distributions $(s_r\|2^{-r-1}N\|, s_{r+k}\|2^{-r-k-1}N\|)$ are (asymptotically) the same for all r. Hence we obtain that the covariance $\mathbf{Cov}(s_r\|2^{-r-1}N\|, s_{r+k}\|2^{-r-k-1}N\|)$ just depends (asymptotically) on the difference k and consequently $\mathbf{Var}\, T_m = dm + O(1)$ for some constant d. Since $d \ge \tilde{d}_2$ it follows that $d > 0$. From that we also get that $\mathbf{Var}\, T_m = dm + o(m)$ if $h(m) = o(m)$.

The *proof of Theorem 4* is now a direct consequence of Chebyshev' inequality combined with Lemma 1 and Lemma 2. We have

$$\frac{1}{2^m} \#\{N < 2^m : |T_m - h(m)/4| \ge \eta\} \ll \frac{m}{\eta^2}.$$

Setting $\eta = \varepsilon h(m)$, using the assumption that $m/h(m)^2 \to \infty$, and observing (from Lemma 2) that $NF_N = O(\sqrt{m})$ the result follows. □

Proof of Theorem 5. By Lemma 1 we have

$$NL_{2,N} = \frac{1}{2}\sqrt{T_m^2 + O(m)}.$$

Since $h(m) \sim m$ the result follows from the general central limit theorem of T_m. Note that we also have a central limit theorem for $NL_{2,N}$ if

[1] By definition $\mathbf{E}(s_r\|2^{-r-1}N\|) = 0$ if row $r+1$ is of type one and $\mathbf{E}(s_r\|2^{-r-1}N\|) = \frac{1}{4} + O(2^{-r})$ if row $r+1$ is of type zero.

$h(m)/\sqrt{m} \to \infty$, however, we do not know the exact behaviour of the variance (just lower and upper bounds of order m) so that we cannot formulate an explicit limit theorem. □

Proof of Theorem 6. Again by Lemma 1 we have $NL_{2,N} \le y\sqrt{m}$ if and only if

$$\frac{1}{4}T_m^2 + \frac{1}{4}\sum_{j=0}^{m-1} \|2^{-j}N\|^2 \le y^2 m.$$

By Chebyshev's inequality it follows that for all but $O(2^m m^{-1/3})$ numbers $N < 2^m$ we have

$$\frac{1}{4}\sum_{j=0}^{m-1} \|2^{-j}N\|^2 = \frac{m}{48} + O(m^{2/3}).$$

This means that this sum behaves almost deterministically. The *random part* is almost concentrated in T_m. We get (for most N)

$$|T_m| \le 2\sqrt{y^2 - \frac{1}{48}}\,\sqrt{m}\,(1+o(1)).$$

Finally (since $h(m)/\sqrt{m} \to 0$) we have

$$\frac{1}{2^m}\#\{N < 2^m : |T_m| \le z\sqrt{dm}\} = \Phi(z) + \Phi(-z) + o(1) = 2\Phi(z) - 1 + o(1).$$

This proves the result. □

6. The upper-1-sequence

In this section we give the proof of Theorem 7.

Proposition 3. *Let $(x_n)_{n\ge 0}$ be the upper-**1**-sequence. Then for any $N \in \mathbb{N}$ of the form $N = 2^m + \frac{2^{m+2}-1}{5}$ with $m \equiv 2 \pmod 4$ we have*

$$\left|\sum_{n=0}^{N-1}\left(\frac{1}{2} - x_n\right)\right| = \left|\frac{\lfloor \mathrm{ld} N\rfloor}{20} - \frac{41}{75} + \frac{13}{150\cdot 2^m}\right|.$$

This result together with Lemma 1 proves Theorem 7.

Proof of Proposition 3. After some straightforward calculations from formula (5) we obtain

$$\begin{aligned}
&\sum_{n=0}^{N-1}\left(\frac{1}{2} - x_n\right)\\
&= \frac{1 + (-1)^{N_0+\cdots+N_{m-1}}}{4} - (-1)^{N_0+\cdots+N_{m-1}}\sum_{r=0}^{m-1}\frac{N(r)}{2^{r+2}}(-1)^{N_0+\cdots+N_{r-1}}.
\end{aligned}$$

Let now $N = 2^m + \frac{2^{m+2}-1}{5}$, $m \equiv 2 \pmod 4$. With $k := (m-2)/4$ we have

$$N = 2^0 + 2^1 + 2^4 + 2^5 + \cdots + 2^{4k} + 2^{4k+1} + 2^m.$$

Hence $N_0 + \cdots + N_{m-1} \equiv 0 \pmod 2$ such that

$$\sum_{n=0}^{N-1} \left(\frac{1}{2} - x_n \right) = \frac{1}{2} - \sum_{r=0}^{m-1} \frac{N(r)}{2^{r+2}} (-1)^{N_0 + \cdots + N_{r-1}}. \tag{9}$$

Now for $0 \leq r \leq m-1$ we have:

1. If $r = 4k$, then $N_0 + \cdots + N_{4k-1} \equiv 0 \pmod 2$ and

$$N(r) = \sum_{i=0}^{k} 2^{4i} + \sum_{i=0}^{k-1} 2^{4i+1} = 2^{4k}\frac{6}{5} - \frac{1}{5}.$$

2. If $r = 4k+1$, then $N_0 + \cdots + N_{4k} \equiv 1 \pmod 2$ and

$$N(r) = \sum_{i=0}^{k} 2^{4i} + \sum_{i=0}^{k} 2^{4i+1} = 2^{4k}\frac{16}{5} - \frac{1}{5}.$$

3. If $r = 4k+2$, then $N_0 + \cdots + N_{4k+1} \equiv 0 \pmod 2$ and

$$N(r) = 2^{4k}\frac{16}{5} - \frac{1}{5}.$$

4. If $r = 4k+3$, then $N_0 + \cdots + N_{4k+2} \equiv 0 \pmod 2$ and

$$N(r) = 2^{4k}\frac{16}{5} - \frac{1}{5}.$$

Now the result follows by splitting up the sum in (9).□

7. Intervals with unbounded remainder

We start this section with the proof of the subsequent lemma which shows that the functions $g_\alpha(m)$ and $f_\alpha(m)$ defined in Section 2 are of the same order of magnitude.

Lemma 5. *For any $\alpha \in [0,1)$ we have*

$$\frac{1}{4}(f_\alpha(m) + 1) \leq \sum_{u=0}^{m-1} \|2^u \alpha\| \leq f_\alpha(m) + 1.$$

Proof. Let $\alpha = a_1/2 + a_2/2^2 + \cdots$. Let $1 \le l_1 < l_2 < \ldots < l_k < m$ such that

$$a_1 = \ldots = a_{l_1} \neq a_{l_1+1} = \ldots = a_{l_2} \neq a_{l_2+1} = \ldots\ldots \neq a_{l_k+1} = \ldots = a_m.$$

Of course it is $k = f_\alpha(m)$. Further define $l_0 := 0$ and $l_{k+1} := m$. Then we have

$$\sum_{u=0}^{m-1} \|2^u\alpha\| = \sum_{i=0}^{k} \sum_{u=l_i}^{l_{i+1}-1} \|2^u\alpha\|. \tag{10}$$

To estimate the innermost sum in the above equation we have to consider two cases. For short we write $p = l_i$ and $q = l_{i+1}$ (note that $q \ge 1$).

1. If $a_{p+1} = \ldots = a_q = 0$ and $a_{q+1} = 1$, then

$$\begin{aligned}\sum_{u=p}^{q-1} \|2^u\alpha\| &= \frac{a_{p+1}}{2} + \cdots + \frac{a_q}{2^q} + \frac{a_{q+1}}{2^{q+1}} + \cdots \\ &+ \frac{a_{p+2}}{2} + \cdots + \frac{a_q}{2^{q-1}} + \frac{a_{q+1}}{2^q} + \cdots \\ &\vdots \\ &+ \frac{a_q}{2} + \frac{a_{q+1}}{2^2} + \cdots \\ &\ge \frac{1}{2^{q+1}} + \cdots + \frac{1}{2^2} = \frac{1}{2}\left(1 - \frac{1}{2^q}\right) \ge \frac{1}{4},\end{aligned}$$

since $q \ge 1$. In the same way we obtain

$$\sum_{u=p}^{q-1} \|2^u\alpha\| \le \frac{1}{2^q} + \frac{1}{2^{q-1}} + \cdots + \frac{1}{2} \le 1.$$

2. If $a_{p+1} = \ldots = a_q = 1$ and $a_{q+1} = 0$, then

$$\begin{aligned}\sum_{u=p}^{q-1} \|2^u\alpha\| = 1 &- \left(\frac{a_{p+1}}{2} + \cdots + \frac{a_q}{2^q} + \frac{a_{q+1}}{2^{q+1}} + \cdots\right) \\ &+ 1 - \left(\frac{a_{p+2}}{2} + \cdots + \frac{a_q}{2^{q-1}} + \frac{a_{q+1}}{2^q} + \cdots\right) \\ &\vdots \\ &+ 1 - \left(\frac{a_q}{2} + \frac{a_{q+1}}{2^2} + \cdots\right) \\ &\geq 1 - \left(-\frac{1}{2^{q+1}} + \frac{1}{2} + \frac{1}{2^2} + \cdots\right) \\ &+ 1 - \left(-\frac{1}{2^q} + \frac{1}{2} + \frac{1}{2^2} + \cdots\right) \\ &\vdots \\ &+ 1 - \left(-\frac{1}{2^2} + \frac{1}{2} + \frac{1}{2^2} + \cdots\right) \\ &= \frac{1}{2^2} + \cdots + \frac{1}{2^{q+1}} + (q-p)\left(1 - \frac{1}{2} - \frac{1}{2^2} - \ldots\right) \geq \frac{1}{4},\end{aligned}$$

since $q \geq 1$. In the same way we obtain

$$\sum_{u=p}^{q-1} \|2^u\alpha\| \leq \frac{1}{2^q} + \frac{1}{2^{q-1}} + \cdots + \frac{1}{2} \leq 1.$$

Inserting these estimates in (10) we obtain

$$\sum_{u=0}^{m-1} \|2^u\alpha\| \geq \sum_{i=0}^{k} \frac{1}{4} \geq \frac{k+1}{4}$$

and

$$\sum_{u=0}^{m-1} \|2^u\alpha\| \leq \sum_{i=0}^{k} 1 \leq k+1.$$

This is the desired result.□

Now we can give the proof of Theorem 8.

Proof of Theorem 8. Define

$$\tilde{g}_\alpha(m) := \frac{1}{2}\sum_{u=0}^{m-1} \|2^u\alpha(m)\|,$$

where $\alpha(m)$ denotes the smallest m-bit number larger or equal to α. If α is greater than $1 - 2^{-m}$, then $\alpha(m) := 1$.

From [13] we know that for $N \le 2^m$ we have

$$\Delta_N^{\mathrm{vdC}}(\alpha) = \delta_{2^m}^H(\alpha(m), N/2^m) + N(\alpha(m) - \alpha),$$

where $\delta_{2^m}^H$ denotes the discrepancy function of the 2-dimensional Hammersley point set in base 2 with 2^m points. From the definition of the Hammersley point set it follows that $\delta_{2^m}^H(\alpha,\beta) = \delta_{2^m}^H(\beta,\alpha)$ for all $0 \le \alpha, \beta \le 1$. Hence we have

$$\Delta_N^{\mathrm{vdC}}(\alpha) = \delta_{2^m}^H(N/2^m, \alpha(m)) + N(\alpha(m) - \alpha). \tag{11}$$

From the proof of [13, Lemma 4] we find that

$$\frac{1}{2^m}\sum_{N=0}^{2^m-1} \delta_{2^m}^H(N/2^m, \alpha(m)) = \frac{1}{2}\sum_{u=0}^{m-1} \|2^u\alpha(m)\| \quad \text{and} \tag{12}$$

$$\frac{1}{2^m}\sum_{N=0}^{2^m-1} \delta_{2^m}^H(N/2^m, \alpha(m))^2 = \left(\frac{1}{2}\sum_{u=0}^{m-1} \|2^u\alpha(m)\|\right)^2 + \frac{1}{4}\sum_{u=0}^{m-1} \|2^u\alpha(m)\|^2. \tag{13}$$

Now we consider

$$\Sigma := \frac{1}{2^m}\sum_{N=1}^{2^m} \left(\delta_{2^m}^H(N/2^m, \alpha(m)) - \widetilde{g}_\alpha(m)\right)^2.$$

With Eq. (12) and (13) we obtain

$$\Sigma = \frac{1}{4}\sum_{u=0}^{m-1} \|2^u\alpha(m)\|^2.$$

Now we have

$$\begin{aligned}
&\frac{1}{4}\sum_{u=0}^{m-1} \|2^u\alpha(m)\|^2 = \frac{1}{2^m}\sum_{N=1}^{2^m} \left(\delta_{2^m}^H(N/2^m, \alpha(m)) - \widetilde{g}_\alpha(m)\right)^2 \\
&\ge \frac{1}{2^m}\left(\varepsilon\widetilde{g}_\alpha(m)\right)^2 \#\left\{N \le 2^m : \left|\delta_{2^m}^H(N/2^m, \alpha(m)) - \widetilde{g}_\alpha(m)\right| \ge \varepsilon\widetilde{g}_\alpha(m)\right\} \\
&= \frac{1}{2^m}\left(\varepsilon\widetilde{g}_\alpha(m)\right)^2 \left(2^m - \#\left\{N \le 2^m : \left|\delta_{2^m}^H(N/2^m, \alpha(m)) - \widetilde{g}_\alpha(m)\right| < \varepsilon\widetilde{g}_\alpha(m)\right\}\right).
\end{aligned}$$

From this we obtain

$$\begin{aligned}
&\frac{1}{2^m}\#\left\{N \le 2^m : \left|\delta_{2^m}^H(N/2^m, \alpha(m)) - \widetilde{g}_\alpha(m)\right| < \varepsilon\widetilde{g}_\alpha(m)\right\} \\
&\qquad\qquad \ge 1 - \frac{\frac{1}{4}\sum_{u=0}^{m-1} \|2^u\alpha(m)\|^2}{\varepsilon^2\widetilde{g}_\alpha(m)^2}.
\end{aligned}$$

Of course it is $\|x\|^2 \leq \|x\|$ and hence we have

$$0 \leq \frac{\frac{1}{2}\sum_{u=0}^{m-1} \|2^u \alpha(m)\|^2}{\widetilde{g}_\alpha(m)^2} \leq \frac{1}{\widetilde{g}_\alpha(m)}.$$

From Lemma 5 we obtain

$$\widetilde{g}_\alpha(m) \geq \frac{1}{8}(f_{\alpha(m)}(m)+1) \geq \frac{1}{8} f_\alpha(m),$$

since $f_\alpha(m)$ and $f_{\alpha(m)}(m)$ differ at most by 1. If α has an infinite base 2 representation, then we have $f_\alpha(m) \longrightarrow \infty$ as $m \longrightarrow \infty$ and therefore we obtain

$$\lim_{m\to\infty} \frac{1}{2^m} \# \left\{ N \leq 2^m : \left|\delta^H_{2^m}(N/2^m, \alpha(m)) - \widetilde{g}_\alpha(m)\right| < \varepsilon \widetilde{g}_\alpha(m) \right\} = 1.$$

With Eq. (11) we obtain

$$\lim_{m\to\infty} \frac{\# \left\{ N \leq 2^m : |\Delta_N^{\mathrm{vdC}}(\alpha) - N(\alpha(m)-\alpha) - \widetilde{g}_\alpha(m)| < \varepsilon \widetilde{g}_\alpha(m) \right\}}{2^m} = 1.$$

For $N \leq 2^m$ we always have $0 \leq N(\alpha(m)-\alpha) \leq 1$. Since α has an infinite base 2 representation we have already seen that $\lim_{m\to\infty} \widetilde{g}_\alpha(m) = \infty$. Hence for $\varepsilon > 0$ we have

$$\lim_{m\to\infty} \frac{1}{2^m} \# \{ N \leq 2^m : -\varepsilon \widetilde{g}_\alpha(m) < N(\alpha(m)-\alpha) < \varepsilon \widetilde{g}_\alpha(m) \} = 1$$

and therefore

$$\lim_{m\to\infty} \frac{1}{2^m} \# \left\{ N \leq 2^m : (1-\varepsilon)\widetilde{g}_\alpha(m) < \Delta_N^{\mathrm{vdC}}(\alpha) < (1+\varepsilon)\widetilde{g}_\alpha(m) \right\} = 1.$$

Since α and $\alpha(m)$ differ at most by $1/2^m$ we find that $g_\alpha(m)$ and $\widetilde{g}_\alpha(m)$ differ at most by 1 and hence the result follows.□

Proof of Corollary 1. The result follows from Theorem 8 together with Lemma 5.□

Proof of Theorem 9. Let $x_0, x_1, \ldots$ be the digital $(0,1)$–sequence in over $\mathbb{Z}_2$ and let $2^{m-1} < N \leq 2^m$. As in the proof of [13, Lemma 2] we find that

$$\Delta_N(\alpha) = \delta_{2^m}(N/2^m, \alpha(m)) + N(\alpha(m)-\alpha), \tag{14}$$

where $\alpha(m)$ is as in the proof of Theorem 8. Further δ_{2^m} denotes the discrepancy function of the digital net $(k/2^m, x_k)$, $0 \leq k \leq 2^m - 1$. From [9, Theorem 1] we know that

$$\delta_{2^m}(N/2^m, \alpha(m)) = \sum_{u=0}^{m-1} \|2^u \alpha(m)\| \varepsilon_u,$$

where $\varepsilon_u \in \{-1,0,1\}$ depending on $u, N, \alpha(m)$. Inserting in Eq. (14) and applying the triangle inequality we obtain

$$|\Delta_N(\alpha)| \le \sum_{u=0}^{m-1} \|2^u \alpha(m)\| + 1 \le f_{\alpha(m)}(m) + 2.$$

Since $f_\alpha(m)$ and $f_{\alpha(m)}(m)$ differ at most by 1 we obtain

$$|\Delta_N(\alpha)| \le f_\alpha(m) + 3.$$

Since $m-1 \le \log_2 N \le m$ and since $f_\alpha(m) \le f_\alpha(m-1)+1$ we obtain

$$|\Delta_N(\alpha)| \le f_\alpha(\lfloor \log_2 N \rfloor) + 4.$$

This is the desired result.□

Finally we give the proof of Theorem 11.

Proof of Theorem 11. We again start with (for $N \le 2^m$)

$$\Delta_N^{\mathbf{1}}(\beta) = \delta_{2^m}^{\mathbf{1}}(N/2^m, \beta(m)) + N(\beta(m) - \beta).$$

Here $\delta_{2^m}^{\mathbf{1}}(x,y)$ denotes the discrepancy function of the two-dimensional digital net $(k/2^m, y_k)$, $0 \le k \le 2^m - 1$ ($y_0, y_1, \ldots$ denotes the upper-$\mathbf{1}$-sequence). Then we use the formula for $\delta_{2^m}^{\mathbf{1}}(x,y)$ given in [9, page 406] ($\delta_{2^m}^{\mathbf{1}}$ is denoted by Δ there):

$$\delta_{2^m}^{\mathbf{1}}(\alpha, \beta(m)) = \sum_{u=0}^{m-1} \|2^u \beta(m)\| \rho(u),$$

$$\alpha = a_1/2 + \cdots + a_m/2^m,\ \beta(m) = b_1/2 + \cdots + b_m/2^m,\ \widetilde{a}_i := a_1 \oplus \ldots \oplus a_{m+1-i},$$

$$\rho(u) := (-1)^{\widetilde{a}_{u+2}} (\widetilde{a}_{u+1} \oplus \widetilde{a}_{u+2} \oplus \widetilde{a}_{r(u)} \oplus \widetilde{a}_{r(u)+1}),$$

and

$$r(u) := \begin{cases} 0 & \text{if } u = 0, \\ 0 & \text{if } b_j = \widetilde{a}_j \text{ for } j = 1, \ldots, u, \\ \max\{j \le u : b_j \ne \widetilde{a}_j\} & \text{otherwise,} \end{cases}$$

and where we have to set $\widetilde{a}_{r(u)} \oplus \widetilde{a}_{r(u)+1} := 0$ if $r(u) = 0$.

We consider in the following

$$\begin{aligned} S_m(\beta(m)) &:= \sum_{N=0}^{2^m-1} \left(\delta_{2^m}^{\mathbf{1}}(N/2^m, \beta(m))\right)^2 \\ &= \sum_{u=0}^{m-1} \sum_{v=0}^{m-1} \|2^u \beta(m)\| \cdot \|2^v \beta(m)\| \sum_{N=0}^{2^m-1} \rho(u)\rho(v), \end{aligned}$$

i.e., we have to consider

$$\Sigma_{u,v} := \sum_{N=0}^{2^m-1} \rho(u)\rho(v).$$

We distinguish two cases:

1. If $u = v$, then we have

$$\sum_{N=0}^{2^m-1} \rho(u)^2 = \sum_{N=0}^{2^m-1} \widetilde{a}_{u+1} \oplus \widetilde{a}_{u+2} \oplus \widetilde{a}_{r(u)} \oplus \widetilde{a}_{r(u)+1}.$$

Now the summation over $N = 0, \ldots, 2^m - 1$ means summation over all $a_1, \ldots, a_m \in \{0,1\}$ and this means summation over all $\widetilde{a}_1, \ldots, \widetilde{a}_m \in \{0,1\}$. Hence (we use shorthand a_i for $\widetilde{a}_i$ and note that replacing in the definition of $r(u)$ the maximum by the minimum does not change the result but simplifies the notation) we have

$$\begin{aligned}
\Sigma_{u,u} &= \sum_{a_1,\ldots,a_m=0}^{1} a_{u+1} \oplus a_{u+2} \oplus a_{r(u)} \oplus a_{r(u)+1} \\
&= \sum_{j=1}^{u-1} \sum_{\substack{a_k=b_k \ \forall 1\le k\le j-1 \\ a_j \ne b_j \\ a_{j+1},\ldots,a_m \in \{0,1\}}} a_{u+1} \oplus a_{u+2} \oplus a_j \oplus a_{j+1} \\
&\quad + \sum_{\substack{a_k=b_k \ \forall 1\le k\le u-1 \\ a_u \ne b_u \\ a_{u+1},\ldots,a_m \in \{0,1\}}} a_u \oplus a_{u+2} + \sum_{\substack{a_k=b_k \ \forall 1\le k\le u \\ a_{u+1},\ldots,a_m \in \{0,1\}}} a_{u+1} \oplus a_{u+2} \\
&= \sum_{j=1}^{u-1} 4 \cdot 2^{m-j-3} + 2^{m-u-1} + 2 \cdot 2^{m-u-2} = 2^{m-1}.
\end{aligned}$$

2. If $u < v$. Here we have to consider two sub-cases, namely $u < v-1$ and $u = v-1$. Then

$$\begin{aligned}
\Sigma_{u,v} \; = & \sum_{j=1}^{u-1} \sum_{\substack{a_k=b_k \;\; \forall 1\le k\le j-1 \\ a_j\neq b_j \\ a_{j+1},\ldots,a_m\in\{0,1\}}} (-1)^{a_{u+1}+a_{v+1}} (a_{u+1}\oplus a_{u+2}\oplus (a_j\oplus a_{j+1})) \times \\
& \qquad\qquad\qquad \times (a_{v+1}\oplus a_{v+2}\oplus (a_j\oplus a_{j+1})) \\
& + \sum_{\substack{a_k=b_k \;\; \forall 1\le k\le u-1 \\ a_u\neq b_u \\ a_{u+1},\ldots,a_m\in\{0,1\}}} (-1)^{a_{u+1}+a_{v+1}} (a_u\oplus a_{u+2}) \\
& \qquad\qquad\qquad \times (a_{v+1}\oplus a_{v+2}\oplus (a_u\oplus a_{u+1})) \\
& + \sum_{j=u+1}^{v-1} \sum_{\substack{a_k=b_k \;\; \forall 1\le k\le j-1 \\ a_j\neq b_j \\ a_{j+1},\ldots,a_m\in\{0,1\}}} (-1)^{a_{u+1}+a_{v+1}} (a_{u+1}\oplus a_{u+2}) \\
& \qquad\qquad\qquad \times (a_{v+1}\oplus a_{v+2}\oplus (a_j\oplus a_{j+1})) \\
& + \sum_{\substack{a_k=b_k \;\; \forall 1\le k\le v-1 \\ a_v\neq b_v \\ a_{v+1},\ldots,a_m\in\{0,1\}}} (-1)^{a_{u+1}+a_{v+1}} (a_{u+1}\oplus a_{u+2})(a_v\oplus a_{v+2}) \\
& + \sum_{\substack{a_k=b_k \;\; \forall 1\le k\le v \\ a_{v+1},\ldots,a_m\in\{0,1\}}} (-1)^{a_{u+1}+a_{v+1}} (a_{u+1}\oplus a_{u+2})(a_{v+1}\oplus a_{v+2}) \\
=: & \; \Sigma_1+\Sigma_2+\Sigma_3+\Sigma_4+\Sigma_4.
\end{aligned}$$

The innermost "senseless" brackets around $a_j\oplus a_{j+1}$ serve to remind that this term is zero if $r(u)=0$.
For all these sums it is easily shown that they equal zero. As demonstration we show this for the "most complicated" case, namely for Σ_1. To that end we have to consider

$$\sum_{a_{u+1},a_{v+1},e,f\in\{0,1\}} (-1)^{a_{u+1}+a_{v+1}} (a_{u+1}\oplus e)(a_{v+1}\oplus f) = \sum_{a_{u+1},a_{v+1}\in\{0,1\}} (-1)^{a_{u+1}+a_{v+1}} = 0$$

(here e (resp. f) plays the role of $a_{u+2}\oplus(a_j\oplus a_{j+1})$ (resp. $a_{v+2}\oplus(a_j\oplus a_{j+1})$) which move independently between 0 and 1). Hence $\Sigma_1=0$, and analogously $\Sigma_{u,v}=0$ for $|u-v|>1$.
If $v=u+1$ then in the above setting Σ_1 and Σ_5 have to be rewritten:

$$\begin{aligned}
\Sigma_1 = & \sum_{j=1}^{u-1} \sum_{\substack{a_k=b_k \;\; \forall 1\le k\le j-1 \\ a_j\neq b_j \\ a_{j+1},\ldots,a_m\in\{0,1\}}} (-1)^{a_v+a_{v+1}} (a_v\oplus a_{v+1}\oplus (a_j\oplus a_{j+1})) \times \\
& \qquad\qquad\qquad \times (a_{v+1}\oplus a_{v+2}\oplus (a_j\oplus a_{j+1})).
\end{aligned}$$

Now

$$\sum_{a_j,a_{j+1}} \sum_{a_v,a_{v+1}} (-1)^{a_v+a_{v+1}}(a_v \oplus a_{v+1} \oplus (a_j \oplus a_{j+1}))$$
$$\times \sum_{a_{v+2}} (a_{v+1} \oplus a_{v+2} \oplus (a_j \oplus a_{j+1}))$$
$$= 2 \sum_{a_v,a_{v+1}} (-1)^{a_v+a_{v+1}}[(a_v \oplus a_{v+1}) + (a_v \oplus a_{v+1} \oplus 1)]$$
$$= 2 \sum_{a_v,a_{v+1}} (-1)^{a_v+a_{v+1}} = 0,$$

hence again $\Sigma_1 = 0$. Similarly Σ_5 and therefore again $\Sigma_{u,v} = 0$.

We conclude

$$\frac{1}{2^m} S_m(\beta(m)) = \frac{1}{2} \sum_{u=0}^{m-1} \|2^u \beta(m)\|^2.$$

So for every m there exists an $N < 2^m$ such that

$$\delta^{\mathbf{1}}_{2^m}(N/2^m, \beta(m)) \geq \sqrt{\frac{1}{2} \sum_{u=0}^{m-1} \|2^u \beta(m)\|^2}$$

and therefore for every m there exists an $N < 2^m$ such that

$$\Delta^{\mathbf{1}}_N(\beta) \geq \sqrt{\frac{1}{2} \sum_{u=0}^{m-1} \|2^u \beta(m)\|^2 - 1} \geq \frac{1}{8} \sqrt{f_\beta(m) - 1}.$$

□

It should be possible by investigating the Walsh series for $\Delta(\alpha, \beta)$ given in [9, Theorem 1] in a similar way to handle also the general case. However this is a quite technical task and is referred to later work.

Acknowledgements. The authors are supported by the Austrian Research Foundation (FWF), Project S 8305. The third author is also supported by the FWF project P17022-N12.

References

[1] R. Béjian, H. Faure, Discrépance de la suite de van der Corput. C. R. Acad. Sci., Paris, Sér. A 285 (1977), 313–316.

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

[3] H. Faure, Discrépances de suites associées a un système de numération (en dimension un). Bull. Soc. math. France 109 (1981), 143–182.

[4] H. Faure, Étude des restes pour les suites de Van Der Corput généralisées. J. Number Theory 16 (1983), 376–394.

[5] H. Faure, Discrepancy and Diaphony of digital $(0,1)$-sequences in prime base. Submitted, 2004.

[6] V. F. Gaposhkin, Lacunary sequences and independent functions, Russian Math. Surveys 21 (1966), 3–82.

[7] P. Hellekalek, On regularities of the distribution of special sequences, Monatsh. Math. 90 (1980), 291–295.

[8] L. Kuipers, H. Niederreiter, Uniform Distribution of Sequences. John Wiley, New York, 1974.

[9] G. Larcher, F. Pillichshammer, Sums of Distances to the Nearest Integer and the Discrepancy of Digital Nets. Acta Arith. 106 (2003), 379–408.

[10] H. Niederreiter, Point sets and sequences with small discrepancy. Monatsh. Math. 104 (1987), 273–337.

[11] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods. No. 63 in CBMS-NSF Series in Applied Mathematics. SIAM, Philadelphia, 1992.

[12] W. Philipp, W. Stout, Almost sure invariance principles for partial sums of weakly dependent random variables. Mem. Amer. Math. Soc. 2, no. 161 (1975).

[13] F. Pillichshammer, On the discrepancy of $(0,1)$–sequences, J. Number Theory 104 (2004), 301–314.

[14] P.D. Proinov, E.Y. Atanassov, On the distribution of the van der Corput generalized sequences. C. R. Acad. Sci. Paris Sér. I Math. 307 (1988), 895–900.

[15] K.F. Roth, On irregularities of distribution. Mathematika 1 (1959), 73–79.

[16] W.M. Schmidt, Irregularities of distribution VI. Comput. Math. 24 (1972), 63–74.

[17] W.M. Schmidt, Irregularities of distribution VII. Acta Arith. 21 (1972), 45–50.

[18] L. Shapiro, Regularities of distribution, in “Studies in Probability and Ergodic Theory” (G.C. Rota, Ed.), pp. 135–154, Academic Press New York/San Francisco/London, 1978.

[19] J. Sunklodas, The rate of convergence in the central limit theorem for strongly mixing random variables. Litovsk. Mat. Sb. 24 (1984), no. 2, 174–185.