

# Recent Progress in Improvement of Extreme Discrepancy and Star Discrepancy of One-dimensional Sequences.

*Victor Ostromoukhov*

Université de Montreal, LIRIS: CNRS, Université Lyon 1

In this communication, we report on recent progress in improvement of extreme discrepancy and star discrepancy of one-dimensional sequences. Namely, we present a permutation of “Babylonian” sequences in base 60, which improves the best known results for star discrepancy obtained by Henri Faure in 1981 [5], and a permutation of sequences in base 84, which improves the best known results for extreme discrepancy obtained by Henri Faure in 1992 [6]. Our best result for star discrepancy in base 60 is  $32209/(35400 \log 60) \approx 0.222223$  (Faure’s best result in base 12 is  $1919/(3454 \log 12) \approx 0.223585$ ); our best result for extreme discrepancy in base 84 is  $130/(83 \log 84) \approx 0.353494$  (Faure’s best result in base 36 is  $23/(35 \log 6) \approx 0.366758$ ).

## 1 Introduction

Variance reduction in quasi-Monte Carlo integration is tightly related to uniformity of distributions of the point sets, which sample the integrand. Among different metrics for evaluation of the uniformity of distributions, *star discrepancy* and *extreme discrepancy* play a special role. In fact, it has been shown [8] that the variance of an integral estimation is bounded by an expression which depends on star discrepancy and extreme discrepancies. Schmidt [9] estimated the lower bounds of star and extreme discrepancies for an arbitrary sequence of points. This theoretical estimation has been later improved by BÉjjan [1]. A thorough description of the problem, the main results and the relevant bibliography can be found in Niederreiter’s book [8].

The first low-discrepancy sequences are due to van der Corput [11]. BÉjjan and Faure [2] estimated the asymptotic behavior of star and extreme discrepancies of the van der Corput sequences. Different constructions for building low-discrepancy sequences have been proposed and evaluated by Borel [3], Braaten and Weller [4], Lapeyre and Pagès [7] and Thomas [10]. In 1981, Faure [5] proposed different generalized (permuted) van der Corput sequences in base 12, having the smallest asymptotic star and extreme discrepancies. In 1989, Thomas [10] improved Faure’s result for extreme discrepancy by a small amount. In 1992, Faure [6] further improved extreme discrepancy, using generalized van der Corput sequences in base 36. Faure’s constructions for star discrepancy (1981, base 12) and for extreme dis-

crepancy (1992, base 36) remain the best known to date results for one-dimensional sequences.

In this paper, we improve Faure's results for extreme discrepancy and star discrepancy of one-dimensional sequences. Our best result for star discrepancy in base 60 is  $32209/(35400 \log 60) \approx 0.222223$  (Faure's best result in base 12 is  $1919/(3454 \log 12) \approx 0.223585$ ); our best result for extreme discrepancy in base 84 is  $130/(83 \log 84) \approx 0.353494$  (Faure's best result in base 36 is  $23/(35 \log 6) \approx 0.366758$ ).

First, let us recall some definitions commonly used in the specialized literature [5, 6, 8].

Let  $X = (x_n)_{n \geq 1}$  be a sequence defined on one-dimensional interval  $[0, 1]$ , and  $A(\alpha, N, X)$  the number of  $n \leq N$  such that  $0 \leq x_n < \alpha$ . The *remainder*  $E$  is defined as  $E(\alpha, N, X) = A(\alpha, N, X) - \alpha N$ ;  $E([\alpha, \beta]; N, X) = E(\beta, N, X) - E(\alpha, N, X)$ , where  $0 \leq x_n < \alpha < \beta \leq 1$ .

The *extreme discrepancy* is defined as  $D(N, X) = \sup_{\alpha, \beta} |E([\alpha, \beta]; N, X)|$ , and the *star discrepancy* is defined as  $D^*(N, X) = \sup_{\alpha, \beta} |E(\alpha, N, X)|$ .

The superior limits of extreme and star discrepancy are defined as

$$s(X) = \overline{\lim}_N (D(N) / \log(N))$$

and

$$s^*(X) = \overline{\lim}_N (D^*(N) / \log(N)).$$

Given an integer  $n \geq 1$  in  $b$ -adic representation  $\sum_{j=0}^{\infty} a_j(n) n^j$  and the sequences of permutations  $(\sigma_j)_{j \geq 0}$  of the set  $\{0, 1, \dots, b-1\}$ , the generalized van der Corput sequence  $S_{b, \sigma}$  in fixed base  $b$  is defined by

$$S_{b, \sigma} = \sum_{j=0}^{\infty} \sigma_j(a_j(n)) n^{-j-1}. \quad (1)$$

In this article, we consider only position-independent permutations, that is permutations  $(\sigma_j)_{j \geq 0}$  which are identical for any position  $j$  in the generalized van der Corput sequence in Equation (1);  $j$  can be omitted.

Let  $Z_b^\sigma = (\sigma(0)/b, \dots, \sigma(b-1)/b)$ . For any integer  $h$  such that  $0 \leq h < b-1$ , the functions  $\Psi_{b, \sigma}^-, \Psi_{b, \sigma}^+$  and  $\Psi_{b, \sigma}$  are defined as follows:

$$\Psi_{b, \sigma}^+(x) = \begin{cases} \max_h (A([0, h/b]; k; Z_b^\sigma) - hx) & \text{if } 0 \leq h \leq \sigma(h-1), \\ \max_h ((b-h)x - A([h/b, 1]; k; Z_b^\sigma)) & \text{if } \sigma(h-1) < h < b, \end{cases} \quad (2)$$

$$\Psi_{b, \sigma}^-(x) = \begin{cases} \max_h (hx - A([0, h/b]; k; Z_b^\sigma)) & \text{if } 0 \leq h \leq \sigma(h-1), \\ \max_h (A([h/b, 1]; k; Z_b^\sigma) - (b-h)x) & \text{if } \sigma(h-1) < h < b, \end{cases} \quad (3)$$

and

$$\Psi_{b, \sigma}(x) = \Psi_{b, \sigma}^+(x) + \Psi_{b, \sigma}^-(x). \quad (4)$$

The terms  $\alpha_{b, \sigma}^+, \alpha_{b, \sigma}^-$  and  $\alpha_{b, \sigma}$  are defined as follows:

$$\begin{aligned}
\alpha_{b,\sigma}^+ &= \inf_{n \geq 1} \sup_{x \in R} \left( \frac{1}{n} \sum_{j=1}^n \Psi_{b,\sigma}^+ \left( \frac{x}{b^j} \right) \right), \\
\alpha_{b,\sigma}^- &= \inf_{n \geq 1} \sup_{x \in R} \left( \frac{1}{n} \sum_{j=1}^n \Psi_{b,\sigma}^- \left( \frac{x}{b^j} \right) \right), \quad \text{and} \\
\alpha_{b,\sigma} &= \inf_{n \geq 1} \sup_{x \in R} \left( \frac{1}{n} \sum_{j=1}^n \Psi_{b,\sigma} \left( \frac{x}{b^j} \right) \right). \tag{5}
\end{aligned}$$

Three theorems by Faure [5] relate the terms of extreme discrepancy  $D(S_{b,\sigma}, N)$  and star discrepancy  $D^*(S_{b,\sigma}, N)$ , as well as the partial terms  $D^+(S_{b,\sigma}, N)$  and  $D^-(S_{b,\sigma}, N)$ , with functions  $\Psi_{b,\sigma}^+$ ,  $\Psi_{b,\sigma}^-$  and  $\Psi_{b,\sigma}$  defined in Equations (2) to (4). Also, they allow to express the superior limits of extreme discrepancy  $s(S_{b,\sigma})$  and star discrepancy  $s^*(S_{b,\sigma})$  in terms of  $\alpha_{b,\sigma}^+$ ,  $\alpha_{b,\sigma}^-$  and  $\alpha_{b,\sigma}$ :

**Theorem 1 (Faure 1981)** *The terms of extreme and star discrepancy of  $S_{b,\sigma}$  can be expressed, for any  $N \geq 1$ , as follows*

$$\begin{aligned}
D^+(S_{b,\sigma}, N) &= \sum_{j=1}^{\infty} \Psi_{b,\sigma}^+ \left( \frac{N}{b^j} \right), \\
D^-(S_{b,\sigma}, N) &= \sum_{j=1}^{\infty} \Psi_{b,\sigma}^- \left( \frac{N}{b^j} \right), \\
D(S_{b,\sigma}, N) &= \sum_{j=1}^{\infty} \Psi_{b,\sigma} \left( \frac{N}{b^j} \right), \quad \text{and} \\
D^*(S_{b,\sigma}, N) &= \max(D^+(S_{b,\sigma}, N), D^-(S_{b,\sigma}, N)).
\end{aligned}$$

**Theorem 2 (Faure 1981)** *The asymptotic term of the extreme discrepancy of  $S_{b,\sigma}$  can be expressed in terms of the constant  $\alpha_{b,\sigma}$ , defined in Equation (5):*

$$s(S_{b,\sigma}) = \overline{\lim}_{N \rightarrow \infty} \frac{D(S_{b,\sigma}, N)}{\log N} = \frac{\alpha_{b,\sigma}}{\log b}.$$

**Theorem 3 (Faure 1981)** *Let  $A \subset N$  defined as  $A = \bigcup_{H=1}^{\infty} A_H$  and  $A_H = \{H(H-1) + 1, \dots, H^2\}$ . Let  $\sigma$  be any permutation of  $\{0, \dots, b-1\}$ , and  $\tau$  be a permutation defined as  $\tau(k) = b-1-k$ , where  $0 \leq k \leq b-1$ . Then, the permutation  $\Sigma_A = (\sigma_j)_{j \geq 1}$  is defined as  $\sigma_j = \sigma$  if  $j \in A$  and  $\sigma_j = \tau \circ \sigma$  if  $j \notin A$ . The asymptotic behavior of the star discrepancy of  $S_{b,\Sigma_A}$  can be expressed in terms of  $\alpha_{b,\sigma}^+$  and  $\alpha_{b,\sigma}^-$  as follows:*

$$s^*(S_{b,\Sigma_A}) = \overline{\lim}_{N \rightarrow \infty} \frac{D^*(S_{b,\sigma}, N)}{\log N} = \frac{\alpha_{b,\sigma}^+ + \alpha_{b,\sigma}^-}{2 \log b}.$$

## 2 Main results

Let  $\sigma_{84}$  be a permutation in base 84:

$$\begin{aligned} \sigma_{84} = & (0, 22, 64, 32, 50, 76, 10, 38, 56, 18, 72, 45, 6, 28, 59, 79, 41, 13, 67, 25, 54, \\ & 2, 36, 70, 16, 48, 81, 30, 61, 8, 43, 74, 20, 52, 4, 34, 66, 15, 46, 77, 26, 11, 62, \\ & 39, 82, 57, 23, 69, 33, 3, 51, 19, 73, 42, 7, 60, 29, 80, 47, 14, 65, 35, 1, 53, 24, \\ & 68, 12, 40, 78, 58, 27, 5, 44, 71, 17, 55, 37, 83, 21, 49, 75, 9, 31, 63). \end{aligned} \quad (6)$$

**Theorem 4** *Let base  $b = 84$  and the permutation  $\sigma_{84}$ , defined in Equation (6). The superior limit of the extreme discrepancy of the sequence  $S_{84, \sigma_{84}}$  is*

$$s(S_{84, \sigma_{84}}) = 130 / (83 \log 84) \approx 0.353494.$$

Let  $\sigma_{60}$  be a permutation in base 60:

$$\begin{aligned} \sigma_{60} = & (0, 15, 30, 40, 2, 48, 20, 35, 8, 52, 23, 43, 12, 26, 55, 4, 32, 45, 17, 37, \\ & 6, 50, 28, 10, 57, 21, 41, 13, 33, 54, 1, 25, 46, 18, 38, 5, 49, 29, 9, 58, \\ & 22, 42, 14, 34, 53, 3, 27, 47, 16, 36, 7, 51, 19, 44, 31, 11, 56, 24, 39, 59). \end{aligned} \quad (7)$$

**Theorem 5** *Let base  $b = 60$  and the permutation  $\sigma_{60}$ , defined in Equation (7). The superior limit of the star discrepancy of the sequence  $S_{60, \Sigma_A}$  is*

$$s^*(S_{60, \Sigma_A}) = 32209 / (35400 \log 60) \approx 0.222223.$$

### 3 Upper and Lower Bounds of $s(S_{84, \sigma_{84}})$ and $s^*(S_{60, \Sigma_A})$

It may be interesting to evaluate numerically the upper and lower bounds of the extreme discrepancy  $s(S_{84, \sigma_{84}})$ . Here, we follow Faure's method presented in [5], Section 5.2.1.

To obtain a lower bound of  $s(S_{84, \sigma_{84}})$ , we compute  $(1/v)F_v(a/(b^v - 1))$  for given integers  $a$  and  $v$  so that  $1 \leq a \leq b^v$ . For  $v = 1$  and  $a = 16$ , we get  $s(S_{84, \sigma_{84}}) \geq 0.353494 \dots = 130 / (83 \log 84)$ . Note that we get the same value of  $s(S_{84, \sigma_{84}})$  by exact calculation, presented in Section 4.2. To obtain an upper bound, we need to compute  $F_n(x)$  up to a sufficiently big  $n$ , then evaluate the expression  $\alpha$ . Namely, for  $n = 6$ ,  $F_6(x)$  reaches its maximum at  $x = 120475271600 / 84^6$ . For this  $x$ ,  $s(S_{84, \sigma_{84}})$  can be calculated:  $s(S_{84, \sigma_{84}}) = 207668158967 / (131736761856 \log 84) \approx 0.355778$ . Therefore, our numerical evaluation of lower and upper bounds of  $s(S_{84, \sigma_{84}})$  can be formulated as follows:

$$0.353494 \leq s(S_{84, \sigma_{84}}) \leq 0.355778.$$

Note that this numerical estimation already surpasses the best Faure's result of  $s(S_{36, \sigma_{36}})$  in base 36.

Similarly, to obtain a lower bound of  $s^*(S_{60, \Sigma_A})$ , we compute  $(1/v)F_v(a/(b^v - 1))$  for given integers  $a$  and  $v$  so that  $1 \leq a \leq b^v$ . For  $v = 2$  and  $a = 1239$ , we get  $s^*(S_{60, \Sigma_A}) \geq 0.222218 \dots = 111 / (122 \log 60)$ . To obtain an upper bound, we need to compute  $F_n(x)$  up to a sufficiently big  $n$ , then evaluate the expression  $\alpha$ . Namely, for  $n = 8$ ,  $F_8(x)$  reaches its maximum at  $x = 57822845901639 / 60^8$ . For this  $x$ ,  $s^*(S_{60, \Sigma_A})$

can be calculated:  $s^*(S_{60,\Sigma_A}) \approx 0.223424$ . Therefore, our numerical evaluation of lower and upper bounds of  $s^*(S_{60,\Sigma_A})$  can be formulated as follows:

$$0.222218 \leq s^*(S_{60,\Sigma_A}) \leq 0.223424.$$

Note that this numerical estimation already improves the best Faure's result of  $s^*(S_{12,\Sigma_A})$  in base 12.

## 4 Proofs

The proofs of Theorems 4 and 5 follow the main line of the proofs provided by Henri Faure in [5, 6].

First, we build the functions  $\Psi_{b,\sigma}^+(x)$ ,  $\Psi_{b,\sigma}^-(x)$  and  $\Psi_{b,\sigma}(x)$ . Then, based on Theorems 1 and 2, we express  $s(S_{b,\sigma})$  in terms of  $\Psi_{b,\sigma}$ . We perform numerical investigation of this function, make an induction hypothesis and prove it.

Similarly, we express  $s^*(S_{b,\sigma})$  in terms of  $\Psi_{b,\sigma}^-$  and  $\Psi_{b,\sigma}^+$ , based on theorems Theorems 1 and 3. We make an induction hypothesis and prove it.

As in [5, 6], we introduce the function

$$F_n(x) = \sum_{k=0}^{n-1} \Psi(xb^k), \quad (8)$$

where  $\Psi(xb^k)$  is the piecewise affine function defined in Equation (4), and express  $\alpha = \inf_{n \geq 1} (\max_{x \in [0,1]} F_n(x)/n)$ .

### 4.1 Function $\Psi_{84,\sigma_{84}}(x)$

Finding  $\Psi_{b,\sigma}^+(x)$ ,  $\Psi_{b,\sigma}^-(x)$  and  $\Psi_{b,\sigma}(x)$  is a tedious work. These functions should be presented as piecewise affine functions on well-defined intervals. As, for definition of  $s(S_{84,\sigma_{84}})$ , we need the function  $\Psi_{84,\sigma_{84}}(x)$  only, we omit here, for the reasons of compactness, the intermediate expressions for  $\Psi_{84,\sigma_{84}}^+(x)$  and  $\Psi_{84,\sigma_{84}}^-(x)$ .

The exact definition of the function  $\Psi_{84,\sigma_{84}}(x)$  defined on intervals  $I_h^1 = [h/84, (h+1)/84]$  is presented in Table 1. Each interval  $I_h^1$  can also be expressed as a set of affine subintervals. Thus, the interval  $[0, 1]$  is expressed as a set of 216 affine subintervals. Figure 1 (left) shows the function  $\Psi_{84,\sigma_{84}}(x)$  visually.

### 4.2 Proof of Theorem 4

Following [5, 6], we define  $\Psi_{b,\sigma_{84}}^-, \Psi_{b,\sigma_{84}}^+$  and  $\Psi_{b,\sigma_{84}}$  on intervals  $I_h^n = [h/b^n, (h+1)/b^n]$ . The interval  $I_h^n$  is called *dominated* if there exists a set  $J$  of integers with  $h \notin J$  such that  $F_n(x) \leq \max_{j \in J} F_n(x + (j-h)/b^n)$  for all  $x \in I_h^n$ . Otherwise, the interval is *dominant*.

Numerical investigations shows that there are three dominant intervals when  $n = 1$ :  $J_{28}^1, J_{52}^1$  and  $J_{55}^1$ . But, for higher  $n$ , there are exactly two dominant intervals. For

**Table 1.**  $\Psi_{84, \sigma_{84}}(x)$  defined on intervals  $I_h^1 = [h/84, (h+1)/84]$ 

h	$\Psi_{84, \sigma_{84}}(x)$ , maximum of linear functions	h	$\Psi_{84, \sigma_{84}}(x)$ , maximum of linear functions
0	{83x}	42	{5 - 7x, 44x - 21}
1	{1 - x, 61x}	43	{22 - 40x, 26x - 12}
2	{2 - 23x, 41x}	44	{21 - 37x, 14x - 6, 59x - 30}
3	{3 - 43x, 31x}	45	{15 - 25x, 21x - 10, 54x - 28}
4	{2 - 11x, 41x - 1}	46	{18 - 30x, 8x - 3, 26x - 13}
5	{4 - 43x, 21x}	47	{34 - 58x, 11 - 17x, 29x - 15}
6	{3 - 21x, 31x - 1}	48	{17 - 27x, 11x - 5, 54x - 30}
7	{5 - 41x, 37x - 2}	49	{12 - 18x, 16x - 8}
8	{6 - 47x, 2 - 7x, 51x - 4}	50	{42 - 68x, 27x - 15}
9	{5 - 33x, 3 - 15x, 29x - 2, 55x - 5}	51	{36 - 57x, 51x - 30, 64x - 38}
10	{5 - 29x, 27x - 2}	52	{14 - 20x, 9 - 12x, 39x - 23}
11	{9 - 57x, 5 - 27x, 25x - 2}	53	{30 - 45x, 21x - 12, 46x - 28}
12	{8 - 45x, 9x, 49x - 6}	54	{26 - 38x, 13x - 7, 33x - 20, 59x - 37}
13	{7 - 35x, 15x - 1}	55	{18 - 25x}
14	{4 - 15x, 37x - 5}	56	{23x - 14}
15	{10 - 47x, 3 - 9x, 45x - 7, 61x - 10}	57	{24 - 33x, 24x - 15}
16	{6 - 23x, 62x - 11}	58	{43 - 60x, 16 - 21x, 32x - 21, 72x - 49}
17	{6 - 22x, 12x - 1}	59	{10 - 12x, 19x - 12}
18	{14 - 58x, 9 - 35x, 15x - 2, 60x - 12}	60	{38 - 51x, 10 - 12x, 20x - 13, 38x - 26}
19	{7 - 24x, 23x - 4, 57x - 12}	61	{35 - 46x, 19 - 24x, 6x - 3, 59x - 42}
20	{8 - 27x, 10x - 1}	62	{20 - 25x, 6x - 3}
21	{3 - 6x, 25x - 5}	63	{9 - 10x, 27x - 19}
22	{17 - 59x, 3 - 6x, 24x - 5, 46x - 11}	64	{29 - 36x, 11x - 7, 24x - 17}
23	{12 - 38x, 7 - 20x, 12x - 2, 51x - 13}	65	{48 - 60x, 13 - 15x, 35x - 26, 58x - 44}
24	{7 - 19x, 12x - 2}	66	{11 - 12x, 22x - 16}
25	{23 - 72x, 11 - 32x, 21x - 5, 60x - 17}	67	{51 - 62x, 23x - 17}
26	{9 - 24x, 33x - 9}	68	{51 - 61x, 38 - 45x, 9x - 6, 47x - 37}
27	{9 - 23x}	69	{32 - 37x, 15x - 11}
28	{25x - 7}	70	{14 - 15x, 35x - 28}
29	{22 - 59x, 13 - 33x, 21x - 6}	71	{43 - 49x, 9 - 9x, 45x - 37}
30	{19 - 49x, 10 - 24x, 9x - 2, 61x - 21}	72	{23 - 25x, 27x - 22, 57x - 48}
31	{10 - 23x, 33x - 11}	73	{25 - 27x, 29x - 24}
32	{21 - 51x, 8 - 17x, 37x - 13}	74	{50 - 55x, 27 - 29x, 15x - 12, 33x - 28}
33	{20 - 47x, 68x - 26}	75	{47 - 51x, 7x - 5, 47x - 41}
34	{8 - 16x, 18x - 6, 47x - 18}	76	{35 - 37x, 41x - 36}
35	{12 - 25x, 13x - 4}	77	{30 - 31x, 23x - 20}
36	{14 - 29x, 17x - 6}	78	{58 - 61x, 27x - 24}
37	{31 - 67x, 41x - 17, 61x - 26}	79	{55 - 57x, 13x - 11, 31x - 28}
38	{12 - 23x, 27x - 11, 40x - 17}	80	{52 - 53x, 31 - 31x, 23x - 21}
39	{9 - 16x, 18x - 7}	81	{60 - 61x, 31 - 31x, 33x - 31}
40	{33 - 66x, 22 - 43x, 7x - 2, 40x - 18}	82	{51 - 51x, x}
41	{23 - 44x, 7x - 2}	83	{83 - 83x}

**Table 2.**  $\Psi_{60,\sigma_{60}}^+(x)$  defined on intervals  $I_h^1 = [h/60, (h+1)/60]$ . Note that  $\Psi_{60,\sigma_{60}}^-(x) = 0$  on  $[0, 1]$ .

h	$\Psi_{60,\sigma_{60}}^+(x)$ , maximum of linear functions	h	$\Psi_{60,\sigma_{60}}^+(x)$ , maximum of linear functions
0	{59x}	30	{46x - 22}
1	{1 - x, 44x}	31	{9 - 14x, 26x - 12}
2	{2 - 16x, 29x}	32	{20 - 34x, 3 - 3x}
3	{3 - 31x, 19x}	33	{3 - 3x, 26x - 13}
4	{19x, 57x - 3}	34	{21 - 34x, 17 - 27x, 13x - 6}
5	{2 - 3x}	35	{13x - 6}
6	{2 - 3x}	36	{6 - 7x}
7	{2 - 3x}	37	{6 - 7x, 9x - 4}
8	{2 - 3x, 19x - 1}	38	{9x - 4, 26x - 15}
9	{5 - 21x}	39	{11 - 14x}
10	{2 - 3x, 36x - 5}	40	{9 - 11x, 29x - 18}
11	{6 - 24x, 2 - 3x}	41	{23 - 31x}
12	{7x, 36x - 6}	42	{29x - 19, 36x - 24}
13	{7 - 24x, 16x - 2, 29x - 5}	43	{19 - 24x, 16x - 10}
14	{9 - 31x}	44	{34 - 44x}
15	{33x - 7}	45	{44x - 32}
16	{9 - 27x, 3 - 5x, 24x - 5}	46	{14 - 16x, 24x - 17}
17	{12 - 36x, 3 - 5x}	47	{30 - 36x, 7 - 7x}
18	{3 - 5x, 24x - 6}	48	{7 - 7x, 29x - 22}
19	{13 - 36x, 11x - 2}	49	{27 - 31x, 9x - 6}
20	{14x - 3}	50	{21x - 16}
21	{18 - 46x, 5 - 9x}	51	{18 - 19x}
22	{5 - 9x, 7x - 1}	52	{18 - 19x, 4x - 2, 36x - 30}
23	{7x - 1}	53	{23 - 24x, 4x - 2}
24	{7 - 13x}	54	{4x - 2}
25	{7 - 13x, 27x - 10}	55	{4x - 2}
26	{16 - 33x, 12 - 24x, 14x - 5}	56	{54 - 56x, 24 - 24x}
27	{14x - 5, 36x - 15}	57	{24 - 24x, 21x - 19}
28	{13 - 24x, 14x - 5}	58	{39 - 39x, x}
29	{24 - 46x}	59	{59 - 59x}

example, when  $n = 2$ , the dominant intervals are  $J_{2420}^2$  and  $J_{4636}^2$ . Further numerical investigations allow us to make the following induction hypothesis: for any  $n > 1$ , the index  $h_n$  of dominant intervals  $J_{h_n}^n$  is either

$$h_n = -\frac{16}{83} + \frac{509}{83} 3^{n+1} 28^{n-1}$$

or

$$h_n = \frac{16}{83} + \frac{797}{83} 3^n 28^{n-1}.$$

In these intervals,  $F_n$  is the affine function in form  $p_n(x - h_n/84^n) + q_n$ , where the coefficients  $p_n$  and  $q_n$  are either

$$p_n = -\frac{61}{83} + \frac{703384^n}{83}; \quad q_n = \frac{130n}{83} + \frac{1434^{2-n}21^{-n}}{6889} + \frac{11715}{192892}$$

or

$$p_n = -\frac{23}{83} + \frac{699584^n}{83}; \quad q_n = \frac{130n}{83} + \frac{1434^{2-n}21^{-n}}{6889} + \frac{11715}{192892}.$$

In both cases,  $\max\{F_n(x) \mid x \in J_{h_n}^n\} = q_n$ .

Our induction hypothesis can be easily checked for  $n = 1$ . Let us suppose that it holds for an arbitrary  $n \geq 1$ . To check that it holds for  $n + 1$ , we need to add  $\Psi(xb^n)$  to  $F_n(x)$  on  $J_{h_n}^n$  and check that  $F_{n+1}(x)$  is still dominant on  $J_{h_{n+1}}^{n+1}$ . We performed this checking for each affine subinterval of definition of the function  $\Psi_{84, \sigma_{84}}(x)$ , and verified that our induction hypothesis holds: the intervals  $J_{h_{n+1}}^{n+1}$  are dominant.

There, we have proved that

$$d_n = \max_{x \in [0,1]} F_n(x)/n = q_n$$

and

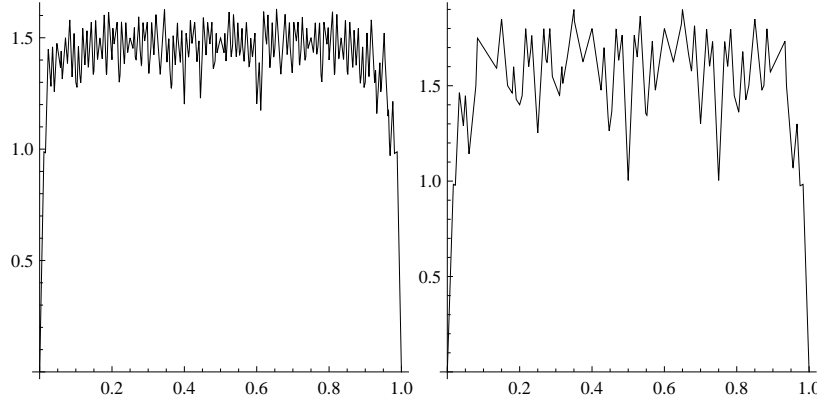
$$\alpha_{84, \sigma_{84}} = \inf_{n \geq 1} d_n/n = \lim_{n \rightarrow \infty} d_n/n = 130/83.$$

Consequently,

$$s(S_{84, \sigma_{84}}) = 130/(83 \log 84) \approx 0.353494.$$

### 4.3 Functions $\Psi_{60, \sigma_{60}}^+(x)$ and $\Psi_{60, \sigma_{60}}^-(x)$

For the definition of  $s^*(S_{60, \Sigma_A})$ , we need the function  $\Psi_{60, \sigma_{60}}^+(x)$  and  $\Psi_{60, \sigma_{60}}^-(x)$ .



**Fig. 1.** Graphical representation of the functions  $\Psi_{84, \sigma_{84}}(x)$  and  $\Psi_{60, \sigma_{60}}^+(x)$ , as defined in Equations (2) to (4), for two particular cases explored in this paper. Left: the function  $\Psi_{84, \sigma_{84}}(x)$ . Right: the function  $\Psi_{60, \sigma_{60}}^+(x)$ . Both are defined on the interval  $[0, 1]$ . Note that  $\Psi_{60, \sigma_{60}}^-(x) = 0$  on  $[0, 1]$ .



The exact definition of the function  $\Psi_{60,\sigma_{60}}^+(x)$  defined on intervals  $I_h^1 = [h/60, (h+1)/60]$  is presented in Table 2. Each interval  $I_h^1$  is also expressed as a set of affine subintervals. Thus, the interval  $[0, 1]$  is expressed as a set of 102 affine subintervals.  $\Psi_{60,\sigma_{60}}^-(x) = 0$  on  $[0, 1]$ . Figure 1 (right) shows the function  $\Psi_{60,\sigma_{60}}^+(x)$  visually.

#### 4.4 Proof of Theorem 5

In this case,  $\Psi_{60,\sigma_{60}}^-(x) = 0$  on  $[0, 1]$ . Consequently,  $\Psi_{60,\sigma_{60}}(x) = \Psi_{60,\sigma_{60}}^+(x)$ .

Numerical investigations shows that there are exactly two dominant intervals, for any  $n \geq 1$ . When  $n = 1$ , the dominant intervals are  $J_{21}^1$  and  $J_{39}^1$ . When  $n = 2$ , the dominant intervals are  $J_{1239}^2$  and  $J_{2361}^2$ . Further numerical investigations allow us to make the following induction hypothesis: for any  $n \geq 1$ , the index  $h_n$  of dominant intervals  $J_{h_n}^n$  is either

$$h_n = \frac{21}{61} ((-1)^n - 60^n)$$

or

$$h_n = \frac{1}{61} (21(-1)^n + 2^{2n+3}3^n5^{n+1}).$$

In these intervals,  $F_n$  is the affine function in form  $p_n(x - h_n/60^n) + q_n$ , where the coefficients  $p_n$  and  $q_n$  are either

$$p_n = \frac{1793}{59} 2^{2n+1} 15^n - \frac{46}{59}; \quad q_n = \frac{32209n}{17700} + \frac{492^{1-2n} 3^{1-n} 5^{-n}}{3481} + \frac{82369}{104430}$$

or

$$p_n = \frac{2}{59} (118760^n - 7); \quad q_n = \frac{32209n}{17700} + \frac{492^{1-2n} 3^{1-n} 5^{-n}}{3481} + \frac{82369}{104430},$$

and  $\max\{F_n(x) \mid x \in J_{h_n}^n\} = q_n$ .

Our induction hypothesis can be easily checked for  $n = 1$ . Let us suppose that it holds for an arbitrary  $n \geq 1$ . To check that it holds for  $n + 1$ , we need to add  $\Psi^+(xb^n)$  to  $F_n(x)$  on  $J_{h_n}^n$  and check that  $F_{n+1}(x)$  is still dominant on  $J_{h_{n+1}}^{n+1}$ . We performed this checking for each affine subinterval of definition of the function  $\Psi_{60,\sigma_{60}}^+(x)$ , and verified that our induction hypothesis holds: the intervals  $J_{h_{n+1}}^{n+1}$  are dominant.

There, we have proved that

$$d_n = \max_{x \in [0,1]} F_n(x)/n = q_n$$

and

$$\alpha_{60,\sigma_{60}}^+ = \inf_{n \geq 1} d_n/n = \lim_{n \rightarrow \infty} d_n/n = 32209/17700; \quad \alpha_{60,\sigma_{60}}^- = 0.$$

Consequently,

$$s^*(S_{60,\Sigma_A}) = (\alpha_{60,\sigma_{60}}^+ + \alpha_{60,\sigma_{60}}^-)/(2 \log 60) = 32209/(35400 \log 60) \approx 0.222223.$$

## 5 Search Method

Looking for good permutations for large bases  $b$  is a difficult task. In fact, an exhaustive search would require explicit evaluation of  $b!$  cases. For example, in base 84, this would require studying  $84! > 10^{127}$  sequences, which is obviously not tractable with modern computers. In this section, we shortly sketch our search method, which makes this difficult task manageable. We illustrate our method using base  $b = 84$ .

Our search method consists in three separate steps. First, we build a (pruned) tree of all possible permutations. The root node is, by convention, 0. The first level of the tree contains 83 branches: it can be any number in the range  $[1, 83]$ . Every branch of the first level contains 82 branches of the second level, etc. We build the tree, starting from the root. When a new branch is added, this corresponds to building a partial permutation. For example, adding the branch '3' to the root '0', is equivalent to building the beginning of the permutation sequence  $\sigma = (0, 3, \dots)$ . At this point, the whole permutation sequence  $\sigma$  is unknown; therefore, we can not build the functions  $\Psi_{b,\sigma}^-$ ,  $\Psi_{b,\sigma}^+$  and  $\Psi_{b,\sigma}$ , as defined in Equations (2) to (4). Nevertheless, we can evaluate the discrepancy for this partial subset of  $k$  elements  $(\sigma(0)/b, \dots, \sigma(k-1)/b)$ ,  $k < b$ . If the discrepancy value of this particular sequence is bigger than a certain pruning threshold value  $T$ , the branch is pruned away. The choice for the pruning threshold is a delicate task: if it is too large, the tree after all pruning operations may contain a huge number of branches. If the pruning threshold  $T$  is too small, the final tree may contain no branches at all. Choosing the right threshold value  $T$  requires many trial-and-errors and some intuition. At the end of the first step, the entire tree of all possible permutations is built. Thanks to pruning, it contains a reasonably small number of branches. For each possible permutation sequence  $\sigma^{(i)}$ , which corresponds to one leaf of the tree, the discrepancy of the first 84 elements of the sequence is below the threshold  $T$ .

At the beginning of the second step, we have a list of permutation sequences  $\sigma^{(i)}$ , one sequence per leaf of the tree. For each sequence  $\sigma^{(i)}$ , the functions  $\Psi_{b,\sigma}^-$ ,  $\Psi_{b,\sigma}^+$  and  $\Psi_{b,\sigma}$  are built according to Equations (2) to (4). Consequently, the terms  $F_n(x)$  can be evaluated according to Equation (8). We sort the sequences  $\sigma^{(i)}$  according to the value of  $F_2(x)$ , calculated for the first  $84^2$  terms of each permutation sequence  $\sigma^{(i)}$ .

During the third step, we study more carefully the permutation sequences  $\sigma^{(i)}$  with the smallest values of  $F_2(x)$ . The behavior of  $F_n(x)$  is studied for  $n > 2$ ; for each  $n$ , the maxima of  $F_n(x)$  are determined. Finally, an induction hypothesis is emitted, and the value of  $p_n, q_n, \alpha_{84}, \sigma_{84}$ , etc. are calculated. A special program checks the induction hypothesis (validation of the dominant intervals), as described in Section 4.2.

## 6 Conclusions

In this contribution, we have shown two permutations in bases 60 and 84, which improve the best known values of asymptotic star and extreme discrepancies of one-dimensional sequences. Our numerical exploration, based on the methodology described in Section 5, has shown that, in general, asymptotic terms of star and extreme

discrepancies decrease as the values of the base  $b$  become bigger. The decrease is not linear, and some particular bases, namely  $b = 60$  and  $b = 84$ , allow particularly low asymptotic terms of star and extreme discrepancies. Our current methodology allows the exploration of integer bases  $b < 100$ . A challenging future step would be developing a more powerful method of search for “good permutations” in larger bases, which could approach the theoretical lower bounds of star and extreme discrepancies, predicted by Schmidt and B ejian.

## Bibliography

- [1] R. B ejian. Minoration de la discr epance d’une suite quelconque sur  $T$ . *Acta Arith.*, 41:185–202, 1982.
- [2] R. B ejian and H. Faure. Discr epance de la suite de van der Corput. *C. R. Acad. Sci. Paris, S erie A*, 285:313–316, 1977.
- [3] J.-P. Borel. Self-similar measures and sequences. *J. of Numb. Theory*, 31(2):208–241, 1989.
- [4] E. Braaten and W. Weller. An improved low-discrepancy sequence for multidimensional quasi-Monte Carlo integration. *J. Comput. Phys.*, 33:249–258, 1987.
- [5] H. Faure. Discr epance des suites associ ees   un syst eme de num eration (en dimension un). *Bull. Soc. Math. France*, 109:143–182, 1981.
- [6] H. Faure. Good permutations for extreme discrepancy. *J. Numb. Theory*, 42:47–56, 1992.
- [7] B. Lapeyre and G. Pag es. Familles de suites discr epance faible obtenues par it eration de transformations de  $[0,1]$ . *Note aux C.R.A.S., S erie I*, 17:507–509, 1989.
- [8] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, volume 63 of *SIAM CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, PA, 1992.
- [9] W. M. Schmidt. Irregularities of distribution. VII. *Acta Arith.*, 21:45–50, 1972.
- [10] A. Thomas. Discr epance en dimension un. *Ann. Fac. Sci. Toulouse, S erie Math.*, 10(3):369–399, 1989.
- [11] J. G. van der Corput. Verteilungsfunktionen I. *Akademie van Wetenschappen*, 38:813–821, 1935.