Tractability using periodized generalized Faure sequences

Christiane Lemieux

Department of Statistics and Actuarial Science, University of Waterloo
E-mail: clemieux@uwaterloo.ca

Abstract

In [1], a construction closely related to generalized Faure sequences was introduced. Unlike generalized Faure sequences, this construction is extensible in the dimension and lends itself well to high-dimensional problems. In this paper, we consider a weighted space of functions for which this construction is particularly well suited and we use it to prove tractability in that space. This space of functions is a special case of one in which we have "finite-order weights", which were studied recently by several authors. We compare our findings with these previously obtained results and provide numerical results to illustrate the practical potential of our approach.

Keywords: tractability, generalized Faure sequence, effective dimension, finite-order weights.

1. Introduction

Low-discrepancy sequences have proven to be useful as deterministic counterpart to Monte Carlo methods for the approximation of integrals of the form

$$I_s(f) = \int_{[0,1)^s} f(\boldsymbol{x}) d\boldsymbol{x},$$

where f is a real-valued function defined over the unit hypercube $I^s := [0,1)^s$. The idea is to construct a sequence of points $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ in $[0,1)^s$ that is evenly distributed and then form the approximation

$$Q_{N,s}(f; P_N) = \frac{1}{n} \sum_{i=1}^{n} f(\boldsymbol{x}_i)$$

for $I_s(f)$, where $P_N = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$ contains the first N points of the sequence. For a sequence that is sufficiently well distributed over $[0,1)^s$ — as measured by the concept of discrepancy to be defined in the next section — and for smooth enough functions, the error $|I_s(f) - Q_{N,s}(f)|$ can then be shown to be of order $O((\log N)^s/N)$, which is asymptotically better than the (probabilistic) $O(1/\sqrt{N})$ associated with random sampling [2].

The Faure sequences and their generalizations are among the most widely used low-discrepancy sequences. These sequences are of interest because they are constructed so that the quality parameter t measuring their uniformity (or equidistribution) is equal to 0, which is the best possible value. However, the price to pay to achieve this is that the base b used to construct the points (i.e., the coordinates of $\boldsymbol{x}_i$ are defined through their base b expansion) must be at least as large as the dimension s. The problem is then that as the base increases, larger values of N are required for the low discrepancy regime to settle in [26, Sec. 5.4].

In this paper, we investigate a construction introduced in [1] that is closely related to Faure sequences, but is such that the base b can be fixed as s increases. Here we call this construction a *periodized generalized Faure sequence* and it is presented in Section 2. In Section 3, we determine a class of functions for which the error achieved by this construction behaves well enough so that integration in that class is *QMC-tractable* or even *QMC-strongly tractable*. We then illustrate through a few numerical examples in Section 4 how this construction performs on different problems, and conclude in Section 5 with a few final remarks.

2. Periodized Generalized Faure Sequences

We start by describing the *van der Corput sequence in base* $b \geq 2$, whose nth term is defined as

$$S_b(n) = \sum_{r=0}^{\infty} a_r(n) b^{-r-1},$$

where $a_r(n)$ is the rth digit of the b-adic expansion of $n-1 = \sum_{r=0}^{\infty} a_r(n)\, b^r$.

This construction forms the basis of several *low-discrepancy sequences*, which are sequences of points that fill the space I^s in such a way that the fraction of the first N points of the sequence included in a given subinterval becomes arbitrarily close, as N goes to infinity, to the volume of that

subinterval. More precisely, denote by $\mathcal{P}_N = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\} \subseteq I^s$ the first N points of the sequence and by $\mathcal{J}^*$ the set of intervals of I^s of the form $J = \prod_{j=1}^{s}[0, z_j)$, where $0 < z_j \leq 1$. Then the *discrepancy function* of $\mathcal{P}_N$ on such an interval J is the difference

$$E(J; \mathcal{P}_N) = \frac{A(J; \mathcal{P}_N)}{N} - V(J),$$

where $A(J; \mathcal{P}_N) = \#\{n; 1 \leq n \leq N, X_n \in J\}$ is the number of points in $\mathcal{P}_N$ that fall in the subinterval J, and $V(J) = \prod_{j=1}^{s} z_j$ is the volume of J.

Then, the *star discrepancy* D^* of $\mathcal{P}_N$ is defined by

$$D^*(\mathcal{P}_N) = \sup_{J \in \mathcal{J}^*} |E(J; \mathcal{P}_N)|.$$

For an infinite sequence X, we denote by $D^*(N, X)$ the discrepancy of its first N points. As sequence satisfying $D^*(N, X) \in O(N^{-1}(\log N)^s)$ is said to be a *low-discrepancy sequence*.

When designing an s-dimensional low-discrepancy sequence from van der Corput sequences in base b, one must first apply a linear transformation in each dimension to the digits $a_0(n), a_1(n), \ldots$ before outputting a number between 0 and 1, an idea originally proposed by Sobol' in [3]. As in [1], we call this a *linearly scrambled* van der Corput sequence. For a prime base b, it is obtained by choosing a matrix C with elements in $\mathbb{Z}_b$ and an infinite number of rows and columns, and then defining the nth term of this sequence as

$$S_b^C(n) := \sum_{r=0}^{\infty} \frac{y_{n,r}}{b^{r+1}} \quad \text{in which} \quad y_{n,r} = \sum_{k=0}^{\infty} c_{r+1,k+1} a_k(n), \tag{1}$$

where $c_{r,k}$ is the kth element on the rth row of C. Note that the second summation is finite and performed in $\mathbb{Z}_b$, but the first one can be infinite and is performed in $\mathbb{R}$, with the possibility that $y_{n,r} = b - 1$ for all but finitely many r. This may result in having points with coordinates equal to 1 and thus outside I^s, an issue that can be handled by the use of a truncation operator as in, e.g., [4, 5]. In this paper, we focus on cases where C is a nonsingular upper triangular matrix, in which case the first summation is finite and therefore the issue of having points outside I^s is avoided.

To assess the quality of point sets obtained by such linear transformations, the quality parameter t is often used. It was introduced by Niederreiter in [6]

to define the concept of (t, s)-sequences, and uses the concept of elementary intervals in base b, which are subsets of I^s of the form

$$\prod_{j=1}^{s} \left[\frac{l_j}{b^{r_j}}, \frac{l_j + 1}{b^{r_j}} \right),$$

where l_j and r_j are integers with $0 \leq l_j < b^{r_j}$ and $r_j \geq 0$, for $j = 1, \ldots, s$. A (t, m, s)-net is defined as a point set P_N with $N = b^m$ points such that any elementary interval of volume b^{-M} with $M := r_1 + \ldots + r_s \leq m - t$ contains b^{m-M} points from P_N. A (t, s)-sequence in base b is a sequence of points $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ such that $\{\boldsymbol{x}_{kb^m+1}, \ldots, \boldsymbol{x}_{(k+1)b^m}\}$ is a (t, m, s)-net for all $m \geq t$ and all $k \geq 0$. A digital (t, s)-sequence over $\mathbb{Z}_b$ is a (t, s)-sequence obtained by juxtaposing linearly scrambled van der Corput sequences in base b.

For digital sequences over $\mathbb{Z}_b$ for a given base b, the first construction that was defined so that $t = 0$ was given by Faure in [7]. It is obtained by choosing a prime base $b \geq s$ and using generating matrices C_j given by the $(j-1)$th power of the (upper triangular) Pascal matrix P_b in $\mathbb{Z}_b$.

It was then shown by Tezuka in [8] that a more general construction for $(0, s)$-sequences is obtained by taking

$$C_j = A_j P_b^{j-1}, \qquad j = 1, \ldots, s, \tag{2}$$

where each A_j is a nonsingular lower triangular (NLT) matrix. This family of constructions is called *generalized Faure sequences* in [8].

Because $(0, s)$-sequences need to be defined in a base $b \geq s$, as mentioned in [9], they do not have the property of being *extensible in the dimension*. A family of constructions with this property is such that if the dimension of a problem is not known ahead of time — this is typical for certain types of simulations that end with a random stopping time — then coordinates can be added on the fly. The problem with $(0, s)$-sequences is that if we choose a base $b \geq s$, then for any $r > b$ we need to choose a new base in order to define a $(0, r)$-sequence, and thus we cannot simply extend each s-dimensional point of the sequence into an r-dimensional one.

Now, if we weaken the property of $(0, s)$-sequences of having $t = 0$ by introducing another quality parameter to replace t, then we can create a construction closely connected to $(0, s)$-sequences, but that has the advantage of being extensible in the dimension. Before describing this construction, we first define this new quality parameter, which has similarities with other criteria discussed in, e.g., [10].

Definition 1. *For a positive integer $k \leq s$, the quality parameter t_k of an s-dimensional digital sequence $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots\}$ over $\mathbb{Z}_b$ is defined as the smallest integer so that for each $u = \{j_1, \ldots, j_r\}$ satisfying $1 \leq j_1 < \cdots < j_r \leq s$, $1 \leq |u| = r \leq k$, and $r(u) := j_r - j_1 + 1 \leq k$, the corresponding projection*

$$\{(x_{n,j_1}, x_{n,j_2}, \ldots, x_{n,j_r}), n \geq 1\}$$

is a (t_k, r)-sequence.

The quality parameter t_k thus focuses on projections over subsets of indices $u = \{j_1, \ldots, j_r\}$ that span a *range* $r(u)$ no larger than k. Note that for a (t, s)-sequence, we have $t_s = t$. Also, if $t_k = 0$, then $t_j = 0$ for all $1 \leq j \leq k$.

The idea proposed in [1] is to fix the base b, and for any dimension $s \geq 1$, to construct an s-dimensional digital sequence over $\mathbb{Z}_b$ using generating matrices of the form

$$C_j = A_j P_b^{j-1}, \qquad j = 1, \ldots, s. \tag{3}$$

Moreover, to make this construction truly extensible in the dimension, it is important to have a predetermined way to choose the scrambling matrices A_j. The idea proposed in [1] is to choose a period $p \approx b/2$ and then let

$$A_j = f_{(j \bmod p)} I, \tag{4}$$

where f_l is the integer in $\mathbb{Z}_b$ with the lth smallest value of θ_b^f, defined as

$$\theta_b^f = \max_{1 \leq N \leq b} \left(T^2(N, S_b^{fI}) - \frac{N^2}{12b^2} \right),$$

and where $T(N, S)$ measures the L_2-discrepancy of the first N points of S. Multipliers f for which θ_b^f is small thus give rise to good one-dimensional sequences. This approach for selecting scrambling matrices was used to define different types of generalized Faure sequences in [1]. The L_2-discrepancy is used to select these factors (as opposed to the star discrepancy, for example) because it can be computed easily and is therefore better suited to this type of computer search. Finally, the reason why we suggest to take the period p to be around $b/2$ is that from experience, we find that once we have enumerated the first half of the possible multipliers f according to θ_b^f, their quality starts deteriorating and we eventually hit "bad" multipliers that do not sufficiently "scramble" the sequence, with 1 and $b - 1$ being the worst cases.

Summing up, we have the following definition:

Definition 2. *Let b be a prime number. A* periodized generalized Faure sequence (PGFS) in base b *is based on generating matrices of the form (3).*

Note that the matrices A_j as defined in (4) with $p > 1$ make the period of the sequence $C_1, C_2, \ldots$ longer than that of the sequence $P_b, P_b^1, P_b^2, \ldots$, which equals b since $P_b^j = P_b^{j+lb}$ for any $l \geq 1$. However, the corresponding PGFS still has bad projections over indices of the form $\{j, j + lb\}$ for any $j, l \geq 1$, as the points accumulate around the b points of a lattice determined by the generating vector $\frac{1}{b}[f_{j \bmod p}, f_{j+lb \bmod p}]$. Thus, while this is better than having all the points on the main diagonal of $[0, 1)^2$ — which is what these two-dimensional projections would be if A_j were the identity matrix for all j — a PGFS should not be used on problems for which these projections correspond to important components of the integrand. Finally, we note that a PGFS with matrices A_j as specified in (4) can handle problems where the dimension is unbounded, because once b and p are chosen and the factors $f \in \mathbb{Z}_b$ are sorted according to θ_b^f, *no extra parameters need to be chosen.*

The following result can be proved easily using the well-known fact that any projection of a $(0, s)$-sequence has a quality parameter $t = 0$.

Proposition 1. *Let b be a prime number. For any $s \geq 1$, an s-dimensional PGFS in base b has a quality parameter $t_b = 0$.*

The purpose of the next section is to introduce classes of functions for which constructions having $t_R = 0$ for some range $R \geq 1$ are well suited. A PGFS with $b \geq R$ gives us such a construction, and the key property that will allow us to prove tractability results is summarized in the following lemma, which gives an upper bound on the star-discrepancy of the projections of such sequences.

Lemma 1. *Let b be a prime and let P_N be the first $N \geq 1$ points of a digital sequence over $\mathbb{Z}_b$ with $t_b = 0$. Let $u = \{j_1, \ldots, j_{|u|}\}$ and let P_N^u be the projection of P_N over u, i.e., $P_N^u = \{(x_{n,j_1}, \ldots, x_{n,j_{|u|}}), n = 1, \ldots, N\}$. Then for any nonempty subset $u \subseteq \{1, \ldots, s\}$ such that $r(u) \leq b$, the star discrepancy of P_N^u satisfies*

$$D^*(P_N^u) \leq \frac{1}{N} \frac{b+1}{2b} (b \log_b(bN))^{|u|}.$$

Before we give the proof of this result, note that any PGFS in base b (including the one based on $A_j = I$ for each j, i.e., where an s-dimensional

sequence is obtained by repeating the b first coordinates of a Faure sequence about s/b times) is such that $t_b = 0$, and therefore satisfies the tractability results in the next section. However, in practice it seems clear that for finite values of N and/or problems that may not fit the classes of functions described in the next section, it is important to "scramble" the sequence according to some carefully chosen matrices A_j, as we propose to do via (4). Results comparing these two options are presented in Section 4.

Proof. The proof is very similar to that of Lemma 1 in [11]. We start with the case where $b \geq 3$, and use the bound given in [2, Theorem 4.12] and the fact that P_N^u is a $(0, |u|)$-sequence if $r(u) \leq b$, which implies that

$$ND^*(P_N^u) \leq \frac{b-1}{2} \sum_{i=1}^{|u|} \binom{|u|-1}{i-1} \binom{k+1}{i} \left(\frac{b-1}{2}\right)^{i-1} + \frac{1}{2} \sum_{i=0}^{|u|-1} \binom{|u|-1}{i} \left(\binom{k+1}{i} + \binom{k}{i}\right) \left(\frac{b-1}{2}\right)^{i} =: \Sigma_1 + \Sigma_2,$$

for $N \geq 1$, where k is the largest integer such that $b^k \leq N$, which means $k \leq \log_b N$. In addition and as in [12], we use the fact that $\binom{k+1}{i} \leq \frac{(k+1)^i}{i!} \leq (k+1)^{|u|}$. Thus we have that

$$\begin{aligned}
\Sigma_1 &\leq \frac{b-1}{2} \sum_{i=1}^{|u|} \binom{|u|-1}{i-1} (k+1)^{|u|} \left(\frac{b-1}{2}\right)^{i-1} \\
&\leq \frac{b-1}{2} (\log_b N + 1)^{|u|} \sum_{i=1}^{|u|} \binom{|u|-1}{i-1} \left(\frac{b-1}{2}\right)^{i-1} \\
&= \frac{b-1}{2} (\log_b bN)^{|u|} \left(\frac{b+1}{2}\right)^{|u|-1} \leq \frac{b-1}{2b} (b \log_b bN)^{|u|}.
\end{aligned}$$

Similarly, for the second term Σ_2 we have

$$\Sigma_2 \leq (\log_b bN)^{|u|} \left(\frac{b+1}{2}\right)^{|u|-1} \leq \frac{1}{b} (b \log_b bN)^{|u|}.$$

Therefore we have

$$ND^*(P_N^u) \leq \frac{b+1}{2b} (b \log_b bN)^{|u|}. \tag{5}$$

The case where $b = 2$ is handled similarly, making use of [2, Theorem 4.13], and the bound (5) also holds in that case. □

3. Tractability results

Tractability is a term that refers to a given class of functions and that quantifies in a very specific way how much computational effort must be spent in order to achieve a given level of integration error. Our treatment of tractability is largely based on [13]. More precisely, for a given Hilbert space H_s of functions $f : I^s \to \mathbb{R}$, norm $\|\cdot\|_{H_s}$, and a given $\epsilon > 0$, we consider integration algorithms defined by a point set P_N used to construct the approximation $Q_{N,s}(f, P_N)$. We then look at the worst-case error over the class H_s, which is defined as

$$e(P_N; H_s) = \sup\{|I_s(f) - Q_{N,s}(f; P_N)| : f \in H_s, \|f\|_{H_s} \leq 1\}, \tag{6}$$

and compare it with the initial error, defined as

$$e(0; H_s) = \sup\{|I_s(f)| : f \in H_s, \|f\|_{H_s} \leq 1\}.$$

We then define $n(\epsilon, H_s)$ as the smallest n for which there exists P_n such that $e(P_n; H_s) \leq \epsilon e(0; H_s)$, where ϵ is in $(0, 1)$.

The behavior of the quantity $n(\epsilon, H_s)$ as a function of the desired error level ϵ and the dimension s is what defines the tractability of integration over H_s. More precisely, integration over H_s is said to be *QMC-tractable* if there exist non-negative numbers C, p, and q such that

$$n(\epsilon, H_s) \leq C\epsilon^{-p}s^q \qquad \text{for all } \epsilon \in (0, 1) \text{ and all } s \geq 1. \tag{7}$$

So tractability means the number of sample points required to decrease the error so that it is within ϵ of the initial error grows polynomially fast with s, and is therefore considered "manageable". As in [13], the numbers p and q are called the $\epsilon-$ *and* $s-$*exponents of QMC-tractability*, respectively. If $q = 0$ in (7), then integration over H_s is said to be *QMC-strongly tractable*, and the infimum of the numbers p satisfying (7) with $q = 0$ is called the $\epsilon-$*exponent of QMC-strong tractability*.

We first define the space of functions for which a PGFS seems to be naturally designed. We follow an approach similar to [13] and assume the space H_s we work with is a *reproducing kernel Hilbert space*. Such spaces are characterized by a reproducing kernel $K_s(\boldsymbol{x}, \boldsymbol{y})$ defined over $I^s \times I^s$ and such that $f(\cdot) = \langle f(\boldsymbol{x}), K_s(\boldsymbol{x}, \cdot)\rangle$ for any $f \in H_s$, where $\langle \cdot, \cdot \rangle$ is the scalar product associated with H_s. As in [13], we assume a reproducing kernel of the form

$$K_s(\boldsymbol{x}, \boldsymbol{y}) = \sum_{u \subseteq \{1,\ldots,s\}} \gamma_{s,u} \prod_{j \in u} \eta_j(x_j, y_j), \tag{8}$$

where

$$\eta_j(x,y) = \frac{1}{2}B_2(|x-y|) + (x-\frac{1}{2})(y-\frac{1}{2}) + \mu_j(x) + \mu_j(y) + m_j$$

and the weights $\gamma_{s,u}$ are arbitrary non-negative numbers, with the assumption that $\gamma_{s,\emptyset} = 1$, and the product over j in (8) is taken to be 1 when $u = \emptyset$. The function $B_2(\cdot)$ is the Bernoulli polynomial of degree 2, i.e., $B_2(x) = x^2 - x + \frac{1}{6}$. The function $\mu_j(\cdot)$ is assumed to have a derivative that exists (except in possibly finitely many points), is bounded in $[0,1]$, and such that $\int_0^1 \mu_j(x)dx = 0$. We also have $m_j = \int_0^1 (\mu_j'(x))^2 dx$. The reproducing kernel (8) is weighted, with the weights $\gamma_{s,u}$ moderating the importance of the different subsets u of variables. It is interesting to note that while the weights do not affect membership to H_s, they affect the value of the norm $\|\cdot\|_{H_s}$ and therefore come into play when we study the worst-case error (6). More precisely, the inner product on a weighted space H_s has the form

$$\langle f, g\rangle = \sum_{u\subseteq\{1,\ldots,s\}} \gamma_{s,u}^{-1}\langle f, g\rangle_u,$$

where $\langle f, g\rangle_u$ is the inner product on the space with reproducing kernel $\prod_{j\in u}\eta_j(x_j, y_j)$, defined e.g. in [14, p. 38]. Hence we see that, as the weights get smaller, we impose more and more conditions on functions $f \in H_s$ in order to have $\|f\|_{H_s} \leq 1$. See [14, Sec. 2.5] for more information.

As in [13], we consider two choices for μ_j, which respectively lead to an *anchored Sobolev space* denoted $H(K_{s,A})$ or an *unanchored Sobolev space* denoted $H(K_{s,B})$. The first choice is to take

$$\mu_j(x) = \mu_{j,A}(x) := \max(x, a_j) - \frac{1}{2}(x^2 + a_j^2) - \frac{1}{3} \text{ with arbitrary } a_j \in [0,1],$$

which leads to $\eta_{j,A}(x,y) = \min(|x_j - a_j|, |y - a_j|)$ if $(x - a_j)(y - a_j) > 0$ and 0 otherwise. The point $(a_1, \ldots, a_s)$ is called the *anchor*. In this case, $m_j = a_j^2 - a_j + \frac{1}{3} \in [1/12, 1/3]$. The second choice is to take $\mu_j(x) = 0$, which leads to $\eta_{j,B}(x,y) = \frac{1}{2}B_2(|x-y|) + \left(x - \frac{1}{2}\right)\left(y - \frac{1}{2}\right)$ and $m_j = 0$ in this case.

Tractability over weighted spaces usually occurs by choosing weights that prevent all subsets of variables from having the same weight. Two ways of doing this that have been explored by many authors (in addition to [13], see, for example [15, 16]) are (1) product-type weights; (2) finite-order weights. The first case refers to weights of the form $\gamma_{s,u} = \prod_{j\in u}\gamma_j$ for some non-negative numbers $\gamma_j \leq 1, j = 1, \ldots, s$. The second case refers to weights for

which there exists an integer q such that $\gamma_{s,u} = 0$ for all u with $|u| > q$. That is, the weights are 0 when the subset u contains more than q indices.

Here, we propose a special case of the latter, which makes use of the notion of range $r(u)$ defined earlier. This type of weights is defined as follows.

Definition 3. *A set of weights $\{\gamma_{s,u}\}_{u\subseteq\{1,\dots,s\}}$ is said to be of* finite-range *if there exists an integer $R \in \{0,\dots,s-1\}$ (called the range) such that $\gamma_{s,u} = 0$ if $r(u) = \max_j\{j \in u\} - \min_j\{j \in u\} + 1 > R$.*

The motivation for this definition takes roots in the connection between weights and functional ANOVA decompositions [17, 18], which amount to writing a function f in $\mathcal{L}^2$ as

$$f(\boldsymbol{x}) = \sum_{u\subseteq\{1,\dots,s\}} f_u(\boldsymbol{x}),$$

where each component $f_u(\boldsymbol{x})$ depends only on the variables x_j such that $j \in u$. These components also satisfy $\int_{I^s} f_u(\boldsymbol{x})d\boldsymbol{x} = 0$ for any non-empty u, and $\int_{I^s} f_\emptyset(\boldsymbol{x})d\boldsymbol{x} = I_s(f)$ otherwise. In addition, this decomposition is orthogonal, so that $\int_{I^s} f_u(\boldsymbol{x})f_v(\boldsymbol{x})d\boldsymbol{x} = 0$ for any distinct $u, v \subseteq \{1,\dots,s\}$. We refer the reader to [19] for more information on the explicit definition of the ANOVA components f_u.

The ANOVA decomposition provides an appropriate setup to define the concept of *effective dimension* [19], by considering the so-called *Sobol' sensitivity indices* [20] $S_u := \sigma_u^2/\sigma^2 \in [0,1]$, where

$$\sigma_u^2 = \int f_u^2(\boldsymbol{x})d\boldsymbol{x} \qquad \text{and } \sigma^2 = \int_{I^s} f^2(\boldsymbol{x})d\boldsymbol{x} - (I_s(f))^2.$$

That is, S_u is equal to the fraction of the variance of f explained by the component f_u. Based on this, the effective dimension of f in the superposition sense (in proportion p) is defined as the smallest integer d_s such that

$$\sum_{u:|u|\le d_s} S_u \ge p.$$

The connection with finite-order weights is that, as proved in [21, Thm 4], for the unanchored Sobolev space $H(K_{s,B})$ the ANOVA decomposition is equivalent to the (unique) orthogonal decomposition

$$f(\boldsymbol{x}) = \sum_u \tilde{f}_u(\boldsymbol{x}),$$

such that $\tilde{f}_u \in H(K_{u,s,B})$, where $K_{u,s,B}$ is given by the term $\gamma_{s,u} \prod_{j \in u} \eta_j(x_j, y_j)$ in the definition (8). That is, the unanchored Sobolev space $K_{s,B}$ is given by the sum $K_{s,B}(\boldsymbol{x}, \boldsymbol{y}) = \sum_u K_{u,s,B}(\boldsymbol{x}, \boldsymbol{y})$. From this result, it is clear that if a function is in $H_s(K_{s,B})$ with weights of finite-order d_s, then its effective dimension in the superposition sense is d_s in proportion 1.

The introduction of a constraint on the range is motivated by the observation that in practice, functions $f(\boldsymbol{x})$ that represent a given stochastic model (examples will be given in Section 4) are often such that the ANOVA components f_u with a significant Sobol' index S_u are those such that not only $|u|$ is not too large, but also such that the indices in u are not too far apart. This is because the index j in x_j is often related to time in a dynamic model. For example, x_j might be the source of randomness to simulate an asset price at time j. A notion of effective dimension that was introduced in [22] with that motivation in mind is to say that f has an effective dimension of d_c in the *successive dimensions sense* (in proportion p) if

$$\sum_{u:r(u) \le d_c} S_u \ge p.$$

Hence we have the following result:

Proposition 2. *A function $f(\boldsymbol{x}) \in H(K_{s,B})$ with weights of finite-range R has an ANOVA decomposition such that $\sigma_u^2 = 0$ if $r(u) > R$ and therefore the effective dimension of f in the superposition sense and the successive dimensions sense is no larger than R.*

We can now establish an upper bound on the integration error of functions in $H(K_{s,A})$ or $H(K_{s,B})$ and with weights of finite range R, when using a digital sequence over $\mathbb{Z}_b$ such that $t_R = 0$, as achieved by a PGFS with $b \ge R$. This result is the counterpart of Theorem 10 in [13].

Theorem 1. *Let $H(K_s)$ be the anchored Sobolev space $H(K_{s,A})$ with an arbitrary anchor $\boldsymbol{a}$, or the unanchored Sobolev space $H(K_{s,B})$, and assume we have finite-range weights $\{\gamma_{s,u}\}_{u \subseteq \{1,\dots,s\}}$ of range $R \ge 1$. Let P_N be the first N points of a digital sequence over $\mathbb{Z}_b$ such that $t_R = 0$, where $b \ge R$. Then*

$$e^2(P_N; H(K_s)) \le \frac{1}{N^2} \sum_{\substack{\emptyset \ne u \subseteq \{1,\dots,s\} \\ r(u) \le R}} \gamma_{s,u} \left(\frac{b+1}{2b}\right)^2 (2b \log_b bN)^{2|u|}. \qquad (9)$$

Proof. The proof is very similar to the one for Theorem 10 in [13]. We start by considering the anchored Sobolev space with the anchor $\boldsymbol{a} = (1, \dots, 1)$. The corresponding kernel is

$$K_{s,A}(\boldsymbol{x}, \boldsymbol{y}) = 1 + \sum_{\emptyset \neq u \subseteq \{1,\dots,s\}} \gamma_{s,u} \prod_{j \in u} \min(1 - x_j, 1 - y_j).$$

As explained in [13], in this case we have

$$e^2(P_N; H(K_{s,A})) = \sum_{\emptyset \neq u \subseteq \{1,\dots,s\}} \gamma_{s,u} \int_{[0,1)^{|u|}} E^2(J(\boldsymbol{x}_u, \underline{1}), P_N) d\boldsymbol{x}_u, \tag{10}$$

where $J(\boldsymbol{x}_u, \underline{1}) = \prod_{j \in u}[0, x_j)$. We also know that

$$\int_{[0,1)^{|u|}} E^2(J(\boldsymbol{x}_u, \underline{1})); P_N) d\boldsymbol{x}_u \leq (D^*(P_N^u))^2.$$

Hence from Lemma 1 and for u such that $r(u) \leq b$ we have that

$$\int_{[0,1)^{|u|}} E^2(J(\boldsymbol{x}_u, \underline{1})); P_N) d\boldsymbol{x}_u \leq \frac{1}{N^2} \left(\frac{b+1}{2b}\right)^2 (b \log_b(bN))^{2|u|}.$$

Therefore, using (10) we get

$$e^2(P_N; H(K_{s,A})) \leq \frac{1}{N^2} \sum_{\emptyset \neq u \subseteq \{1,\dots,s\}, r(u) \leq R} \gamma_{s,u} \left(\frac{b+1}{2b}\right)^2 (b \log_b(bN))^{2|u|},$$

which proves the result for $\boldsymbol{a} = (1, \dots, 1)$.

The case of an arbitrary anchor is handled similarly, and here we do not reproduce all the steps from [13]. The important step to point out (and that is different from the case where $\boldsymbol{a} = (1, \dots, 1)$) is that for a general anchor,

$$e^2(P_N; H(K_{s,A})) \leq \sum_{\emptyset \neq u \subseteq \{1,\dots,s\}} \gamma_{s,u} 4^{|u|} (D^*(P_N^u))^2.$$

Hence we have

$$e^2(P_n; H(K_{s,A})) \leq \frac{1}{N^2} \sum_{\emptyset \neq u \subseteq \{1,\dots,s\}, r(u) \leq R} \gamma_{s,u} \left(\frac{b+1}{2b}\right)^2 (2b \log_b(bN))^{2|u|}. \tag{11}$$

Next we consider the case of an unanchored Sobolev space. Again, following the approach in [13] (which contains several steps that apply to any low-discrepancy point set and not only the ones considered in that paper), we use the fact that

$$e^2(P_N; H(K_{s,B})) \leq \sum_{\emptyset \neq u \subseteq \{1,\ldots,s\}} \gamma_{s,u} 4^{|u|} (D^*(P_N^u))^2,$$

which means the bound (11) holds in that case as well. □

Remark 1. *We note that the corresponding bound in [13], obtained for Niederreiter sequences, is of a form similar to (9) but with a term $j \log_2(j+b)$ in the product over $j \in u$. Hence the terms comprised in the sum over the subsets u grow as the indices in u get larger. This problem is avoided with our construction and its pairing to the class of functions we consider.*

The above bound on the integration error allows us to get the next result, which studies the tractability of integration for classes of functions having finite-range weights. It is the analog of Theorem 11 in [13].

Theorem 2. *Let $\{\gamma_{s,u}\}_{s\geq 1, u\subseteq\{1,\ldots,s\}}$ be weights of finite-range $R \geq 1$. Let P_N be the first N points of a digital sequence over $\mathbb{Z}_b$ with $b \geq R$ and such that $t_R = 0$ for all $s \geq 1$. (a) Consider the anchored Sobolev space $H(K_{s,A})$ with an arbitrary anchor $\boldsymbol{a}$ and weights $\{\gamma_{s,u}\}_{u\subseteq\{1,\ldots,s\}}$. Then we have*

$$\frac{e(P_N; H(K_{s,A}))}{e(0; H(K_s))} \leq C(b) \frac{1}{N} (\log_b bN)^b,$$

where $C(b) = (4\sqrt{3}b)^b$. Furthermore, for any arbitrary $\delta > 0$ there exists a constant C_δ independent of s and N such that

$$\frac{e(P_N; H(K_{s,A}))}{e(0; H(K_{s,A}))} \leq C_\delta N^{-1+\delta}.$$

Hence we have QMC-strong tractability with $\epsilon-$exponent 1.

(b) Consider the unanchored Sobolev space $H(K_{s,B})$ with weights $\{\gamma_{s,u}\}_{u\subseteq\{1,\ldots,s\}}$. (i) If there exists c_B^* such that $\gamma_{s,u} \leq c_B^*$ for all u and $s \geq 1$, then

$$\frac{e(P_N; H(K_{s,B}))}{e(0; H(K_{s,B}))} \leq C_1(b) \frac{s^{1/2}}{N} (\log_b bN)^b,$$

where

$$C_1(b) = \sqrt{c_B^*}(\sqrt{8}b)^b.$$

(ii) If we further assume that

$$\mathcal{M} = \sup_{s=1,2,\ldots} \sum_{u:0\le r(u)\le R} \gamma_{s,u} < \infty \tag{12}$$

then

$$\frac{e(P_N; H(K_{s,B}))}{e(0; H(K_{s,B}))} \le C_2(b)\frac{1}{N}(\log_b bN)^b,$$

where $C_2(b) = \sqrt{\mathcal{M}}(2b)^b$. *Furthermore, for any arbitrary* $\delta > 0$ *there exists a constant* $C_{B,\delta}$ *independent of* s *and* N *such that*

$$\frac{e(P_N; H(K_{s,B}))}{e(0; H(K_{s,B}))} \le C_{B,\delta} N^{-1+\delta}.$$

Hence we have QMC-strong tractability with $\epsilon-$*exponent 1.*

Proof. Consider the anchored Sobolev space $H(K_{s,A})$. Following [13], we have that the initial error $e(0; H(K_{s,A})$ satisfies

$$e^2(0; H(K_{s,A})) = \sum_{u\subseteq\{1,\ldots,s\}} \gamma_{s,u} \prod_{j\in u} m_j.$$

If we have arbitrary weights of finite range R, then from Theorem 1 we have

$$\begin{aligned}
\frac{e^2(P_n; H(K_{s,A}))}{e^2(0; H(K_{s,A}))} &\le \frac{1}{N^2} \frac{\sum_{0\le r(u)\le R} \left(\frac{b+1}{2b}\right)^2 \gamma_{s,u} (2b\log_b bN)^{2|u|}}{\sum_{0\le r(u)\le R} \gamma_{s,u} \prod_{j\in u} m_j} \\
&\le \frac{12^R}{N^2} \frac{\sum_{0\le r(u)\le R} \gamma_{s,u} \left(\frac{b+1}{2b}\right)^2 (2b\log_b bN)^{2|u|}}{\sum_{0\le r(u)\le R} \gamma_{s,u}} \\
&\le \frac{12^R}{N^2} (2b\log_b bN)^{2R} \left(\frac{b+1}{2b}\right)^2 \le \frac{12^b}{N^2} (2b\log_b bN)^{2b},
\end{aligned}$$

since $R \le b$ and $b \ge 2$. The second inequality above follows from the fact that $m_j \in [1/12, 1/3]$. Therefore

$$\frac{e(P_N; H(K_{s,A}))}{e(0; H(K_{s,A}))} \le \frac{1}{N}(2\sqrt{3})^b (2b)^b (\log_b bN)^b, \tag{13}$$

and since the bound on the right-hand side of (13) does not depend on s, we clearly have QMC-strong tractability. Also, from (13) and for any $\delta > 0$, we can write

$$e(P_N; H(K_{s,A}) \le e(0; H(K_{s,A}))\frac{1}{N}(\log_b N + 1)^b C(b) \le e(0; H(K_{s,A}))C_\delta N^{-1+\delta}$$

for all $N \ge 1$ and for some $C_\delta > 0$. This is because for any $\delta > 0$, there exists a constant $K(b, \delta)$ such that $(\log_b N + 1)^b \le N^\delta$ for all $N \ge K(b, \delta)$. Therefore, we have

$$\frac{(\log_b N + 1)^b}{N} C(b) \le C_\delta N^{-1+\delta} \text{ for all } N \ge 1$$

if we take $C_\delta = C(b)(\log_b K(b, \delta) + 1)^b$. Hence for fixed $\epsilon > 0$ we can take $N = \lceil \epsilon^{-1/(1-\delta)}/C_\delta \rceil$ to get $e(P_N; H(K_{s,A}) \le \epsilon e(0; H(K_{s,A}))$, and thus the ϵ-exponent of QMC-strong tractability is 1.

The case of $H(K_{s,B})$ is handled similarly, but with the difference that $e^2(0; H(K_{s,B})) = 1$ and thus in case (i)

$$\begin{aligned}
\frac{e^2(P_n; H(K_{s,B}))}{e^2(0; H(K_{s,B}))} &\le \frac{1}{N^2} \sum_{u:0\le r(u)\le R} \left(\frac{b+1}{2b}\right)^2 \gamma_{s,u} (2b \log_b bN)^{2|u|} \\
&\le \frac{c_B^* s 2^R}{N^2} (2b \log_b bN)^{2R} \left(\frac{b+1}{2b}\right)^2 \\
&\le \frac{s c_B^*}{N^2} (2b \log_b bN)^{2b} 2^b,
\end{aligned}$$

where the second inequality holds because the number of subsets u such that $r(u) \le R$ is bounded by $s2^R$. Therefore

$$\frac{e(P_N; H(K_{s,B}))}{e(0; H(K_{s,B}))} \le \frac{\sqrt{c_B^*}}{N} (2b)^b (\log_b bN)^b s^{1/2} 2^{b/2}. \tag{14}$$

Hence here we have QMC-tractability with s-exponent 1/2.

In the case (ii) where we use a stronger assumption on the weights $\gamma_{s,u}$, an argument similar to the one for (i) yields

$$\frac{e^2(P_n; H(K_{s,B}))}{e^2(0; H(K_{s,B}))} \le \frac{\mathcal{M}}{N^2} (2b \log_b bN)^{2b} \left(\frac{b+1}{2b}\right)^2.$$

and thus

$$\frac{e(P_N; H(K_{s,B}))}{e(0; H(K_{s,B}))} \leq \frac{\mathcal{M}^{1/2}}{N}(2b)^b(\log_b bN)^b, \tag{15}$$

and since the bound on the right-hand side of (15) does not depend on s, we clearly have QMC-strong tractability. The rest of the argument used for the anchored case can then be applied in the same way to prove the second statement for the case (ii). □

Remark 2. *Comparing with the corresponding results obtained for Niederreiter sequences in [13], in the anchored case we are able to achieve QMC-strong tractability with no further condition on the weights $\gamma_{s,u}$ other than being of finite range. To achieve the same with finite-order weights in [13], the condition*

$$\sup_{s=1,2,\ldots}\left(\frac{\sum_{u,|u|\leq q^*}\gamma_{s,u}\prod_{j\in u}(j\log_2(j+b))^2)}{\sum_{u,|u|\leq q^*}\gamma_{s,u}\prod_{j\in u}m_j}\right)<\infty$$

is used. In the unanchored case, our condition (12) for QMC-strong tractability is weaker than the one in [13], given by

$$\mathcal{M}=\sup_{s=1,2,\cdots}\sum_{0\leq r(u)\leq R}\gamma_{s,u}\prod_{j\in u}(j\log_2(j+b))^2<\infty.$$

Remark 3. *We note that in the anchored case, the initial error is bounded by $\sum_{u:r(u)\leq R}\gamma_{s,u}3^{-|u|}$, which grows linearly in s, assuming $\gamma_{s,u}3^{-|u|}$ is bounded. While this does not prevent strong QMC-tractability from holding, one could add conditions on the weights $\gamma_{s,u}$ to restrict the set of functions considered to those for which the initial error can remain bounded as s grows.*

4. Examples

In this section, we present a few numerical examples where the performance of our proposed construction is examined and compared against other alternatives. These alternative constructions are: 1) the Sobol' sequence with direction numbers from the RandQMC library [23]; 2) the extensible rank-1 construction [24] with generating vector found at `http://web.maths.unsw.edu.au/~fkuo/lattice/index.html` and based on order-2 weights; 3) the extensible Korobov lattice proposed in [25], with parameter $a = 14,471$.

Here is what we expect to see in these examples. First, the alternative constructions used in these examples are also designed to work well in high dimensions, so we do not necessarily expect our PGFS will do better than these. But we hope to demonstrate that it can be competitive with these other methods. In particular, we note that while our method yields better bounds than Niederreiter or Sobol' sequences for the classes of functions considered in this paper, the actual error may behave differently.

In these examples, constructions are compared by using an appropriate randomization and then computing the average absolute error of the resulting estimator. More precisely, let $\tilde{P}_{n,l}$ be the lth iid randomized copy of a point set P_N, for $l = 1, \ldots, L$. Then we compute

$$\hat{\epsilon}_N = \frac{1}{L} \sum_{l=1}^{L} |(Q_{N,s}(f; \tilde{P}_{N,l}) - I_s(f)|.$$

The randomization we chose is simply to add a random shift to P_N, which for point sets from the Sobol' sequence and PGFS is done on the digital expansion of each point $\boldsymbol{x}_i \in P_N$ (i.e., we use a digital shift), while for lattices we add the shift modulo 1 to each component of each point. All examples are done using $L = 50$ randomizations. We refer the reader to [26] for more information on randomized quasi-Monte Carlo methods.

If the integral $I_s(f)$ is not known analytically, we estimate it using the average value obtained from the 2^{22} first points of our extensible Korobov sequence over 50 iid random shifts. We chose to work with the average absolute error because although all our estimators are unbiased, the PGFS is not necessarily uniformly distributed [2], so we cannot guarantee, for a given randomized copy $\tilde{P}_{N,l}$, that the estimator $Q_{N,s}(f; \tilde{P}_{N,l})$ will converge to $I_s(f)$ when f is not included in the space of functions studied in this paper, as is the case for the problems considered in our second and third examples. This potential issue can be better detected by measuring the average absolute error than, say, the estimated variance.

We consider three examples. The first one is based on the test function

$$g(\boldsymbol{x}) = \prod_{j=1}^{s} (1 + c(x_j - 0.5))$$

introduced in [27]. Here however, we slightly modify this function so that it fits the type of functions for which a PGFS is expected to do well. More

precisely, we fix parameters k and s, and then use the function

$$g_{k,s}(\boldsymbol{x}) = \frac{1}{s-k+1} \sum_{l=1}^{s-k+1} g(x_l, \ldots, x_{l+k-1})$$

for different values of s and k.

This first example allows us to consider the case where we have a function, namely $g_{k,s}(\boldsymbol{x})$, in the unanchored Sobolev space $H(K_{s,B})$ with weights of finite-order k. To see why this holds, first notice that

$$g(\boldsymbol{x}) = 1 + \sum_{\emptyset \neq u \subseteq \{1,\ldots,s\}} c^{|u|} \prod_{j \in u} (x_j - 0.5)$$

and therefore

$$g_{k,s}(\boldsymbol{x}) = 1 + \frac{1}{s-k+1} \sum_{l=1}^{s-k+1} \left[\sum_{\emptyset \neq u \subseteq \{l,\ldots,l+k-1\}} c^{|u|} \prod_{j \in u} (x_j - 0.5) \right].$$

Hence we see that $g_{k,s}(\boldsymbol{x})$ can be written as a sum

$$g_{k,s}(\boldsymbol{x}) = \sum_{u : r(u) \leq k} g_{u,k,s}(\boldsymbol{x})$$

where each $g_{u,k,s}(\boldsymbol{x}) \in H(K_{u,s,B})$. Thus $g_{k,s}(\boldsymbol{x})$ is in the unanchored Sobolev space $H(K_{s,B})$, but with the additional restriction that $\gamma_{s,u} = 0$ if $r(u) > k$.

The results are presented in Figures 1 and 2. We used a PGFS in base $b = 97$ (and with $p = 42$ in (4)) for this example. It should be noted that since the figures use a log-log scale, there is more emphasis on smaller values of N.

As we can see in these examples, the extensible Korobov sequence is usually the one with the smallest error, but the PGFS is typically the next best one.

Our second example uses an Asian call option (see e.g., [26]). The corresponding function $f(\boldsymbol{x})$ is such that the price of the option $C(T, s, r, \sigma)$ satisfies

$$C(T, s, r, \sigma) = \mathrm{E}\left(e^{-rT} \max\left(\frac{1}{s} \sum_{j=1}^{s} S(t_j), 0 \right) \right) = \int_{I^s} f(\boldsymbol{x}) d\boldsymbol{x},$$

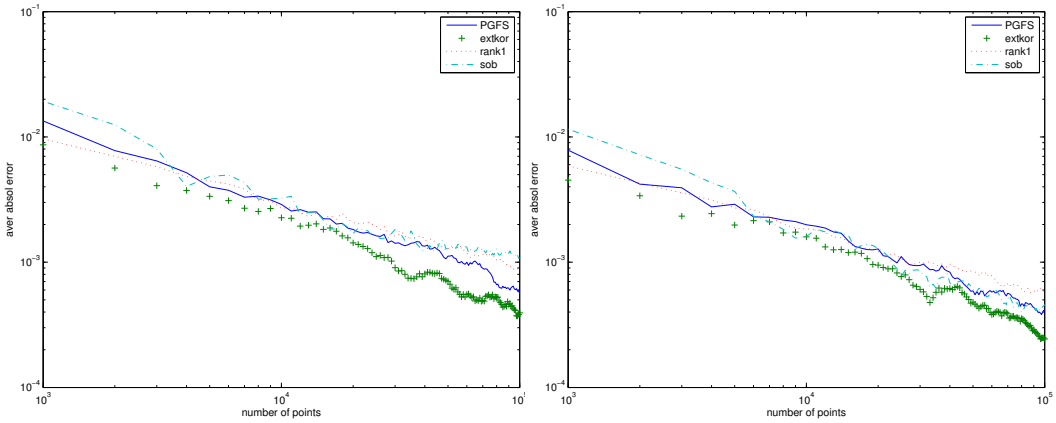

Figure 1: $g_{k,s}$ with $k = 20$ and $c = 1$; $s = 96$ (left), $s = 250$ (right)

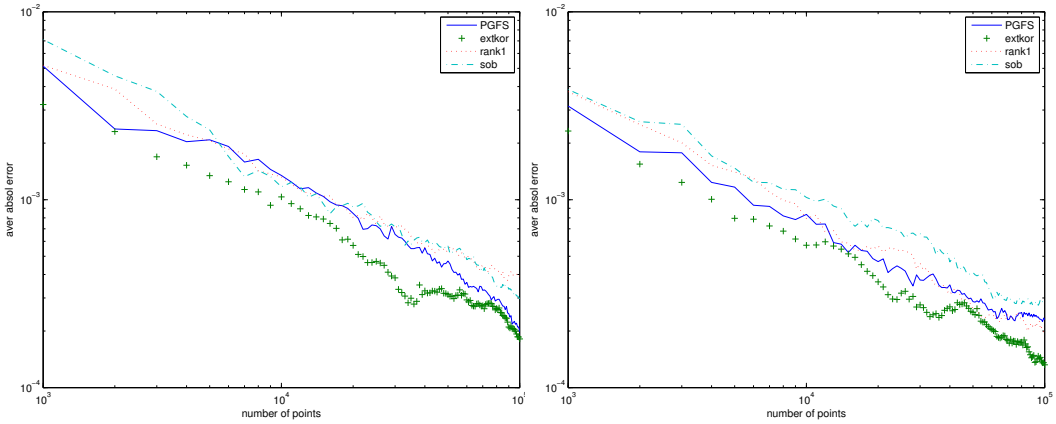

Figure 2: $g_{k,s}$ with $k = 20$ and $c = 1$; $s = 500$ (left), $s = 1000$ (right)

where $S(t_j)$ is the price of the underlying asset at time $t_j = jT/s$. Given the previous price $S(t_{j-1})$, $S(t_j)$ is assumed to have a lognormal distribution with parameters $((r - \sigma^2/2)/s, \sigma^2/s)$, and where T is the expiration time of the option, r is the risk-free rate, and σ is the volatility of the asset. Hence the function $f(\boldsymbol{x})$ in this case can be written as

$$f(\boldsymbol{x}) = e^{-rT} \max\left(0, \frac{1}{s}\sum_{j=1}^{s} \exp((r - \sigma^2/2)(jT/s) + \sum_{l=1}^{j} \Phi^{-1}(x_l)\sigma\sqrt{T/s})\right),$$

where $\Phi(\cdot)$ is the CDF of a standard Normal rv.

In this case, we do not have a function in one of the Sobolev spaces used

in the previous sections. However, using the method from [28], we were able to estimate a lower bound on the fraction of the function's variance that is explained by projections included in the subset that corresponds to those with a non-zero weight for functions with a fixed finite range. More precisely, we estimated a lower bound on

$$\gamma(2, 32, s) := \sum_{u:|u|\leq 2, r(u)\leq 32} S_u$$

for $s = 32, 64, 128, 256$. The results are provided in Table 1 (the numbers on the second row are the half-widths of a 95% confidence interval).

Table 1: Lower bound on fraction of variance explained by ANOVA components of order 2 and range smaller or equal to 32

	$s = 32$	$s = 64$	$s = 128$	$s = 256$
$\hat{\gamma}(2, 32, s)$	0.9735	0.9662	0.9105	0.8435
hw	0.0080	0.0056	0.0083	0.0152

While it is clear that $\gamma(2, 32, s)$ is decreasing with s, the results in Table 1 suggest that even when $s = 256$, we can approximate the corresponding function f fairly well by a sum of one- and two-dimensional components with a range restricted to 32. We thus expect that a PGFS (or any low-discrepancy point set with corresponding good projections over these subsets of indices of bounded range) will provide an estimator whose performance does not decrease too much with s. The results shown in Figures 3 and 4 seem to confirm this. We used a PGFS in base $b = 241$ and with $p = 122$ for this example. The results also show the performance of the naive construction mentioned at the end of Section 2, which consists in taking $A_j = I$ in (4). The error is much larger than for any of the other methods, and we also see a much larger sensitivity as to whether or not N is a power of b, with a large downward peak close to $b^2 = 58081$.

Our third example is based on a simple queueing system (see, e.g., [26]). Clients arrive according to a Poisson process with arrival rate of 1/minute, and receive service of length that is exponentially distributed with mean 55 seconds. All random variables in this model are assumed to be independent. We simulate the arrival of l clients and are interested in the expected number of clients who will have to wait for more than 10 minutes before being served.

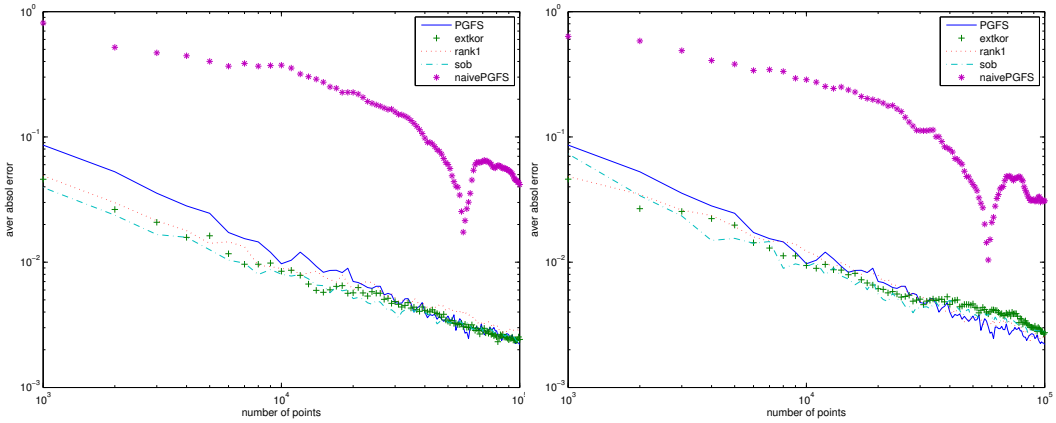

Figure 3: Asian call option with $K = 50$ and $s = 32$ (left), $s = 64$ (right)

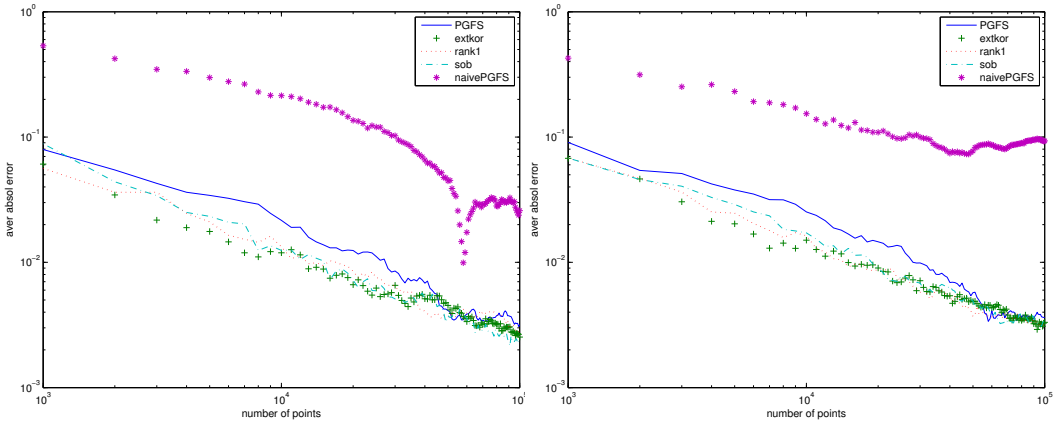

Figure 4: Asian call option with $K = 50$ and $s = 128$ (left), $s = 256$ (right)

The problem is thus $2l$-dimensional. We used a PGFS in base $b = 727$ and with $p = 396$.

We see in this example that the Sobol' sequence in dimension 10,000 is not performing so well for larger values of N. Our PGFS is competitive with the extensible lattice.

5. Conclusion

In this paper, we have presented a class of functions that is well suited for a construction introduced in [1]. This construction, which we call a periodized generalized Faure sequence here, has the advantage of providing

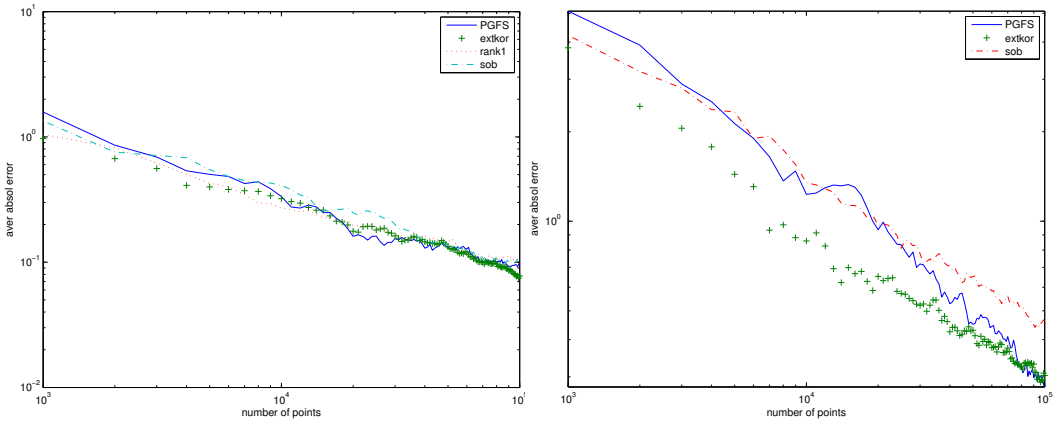

Figure 5: Queueing system with $l = 500$ clients (left) and $l = 5000$ clients (right)

perfect equidistribution over projections with a restricted range, while also being extensible in the dimension. This construction seems competitive with other popular ones, and is simple to implement.

Acknowledgements

I wish to thank Henri Faure for ongoing discussions on this topic and the two anonymous referees for their constructive comments. This work was supported by NSERC through grant #238959.

References

[1] C. Lemieux, H. Faure, New perspectives on $(0, s)$-sequences, in: P. L'Ecuyer, A. Owen (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2008, Springer-Verlag, 2009, pp. 113–130.

[2] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, Vol. 63 of SIAM CBMS-NSF Reg. Conf. Series Appl. Math., SIAM, Philadelphia, 1992.

[3] I. M. Sobol', On the distribution of points in a cube and the approximate evaluation of integrals, USSR Comp. Math. Math. Phys. 7 (1967) 86–112.

[4] H. Niederreiter, C. Xing, Low-discrepancy sequences and global function fields with many rational places, Finite Fields Appl. 2 (1996) 241–273.

[5] S. Tezuka, Polynomial arithmetic analogue of Halton sequences, ACM Trans. Model. Comput. Simul. 3 (1993) 99–107.

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

[7] H. Faure, Discrépance des suites associées à un système de numération (en dimension s), Acta Arith. 41 (1982) 337–351.

[8] S. Tezuka, A generalization of Faure sequences and its efficient implementation, Tech. Rep. RT0105, IBM Research, Tokyo Research Laboratory (1994).

[9] F. J. Hickernell, My dream quadrature rule, J. Complexity 19 (2003) 420–427.

[10] G. Larcher, Digital point sets: Analysis and applications, in: P. Hellekalek, G. Larcher (Eds.), Random and Quasi-Random Point Sets, Vol. 138 of Lecture Notes in Statistics, Springer, New York, 1998, pp. 167–222.

[11] X. Wang, Strong tractability of multivariate integration using quasi-Monte Carlo algorithms, Math. Comp. 72 (2002) 823–838.

[12] X. Wang, K.-T. Fang, The effective dimension and quasi-Monte Carlo integration, J. Complexity 19 (2003) 101–124.

[13] I. H. Sloan, X. Wang, H. Woźniakowski, Finite-order weights imply tractability of multivariate integration, J. Complexity 20 (2004) 46–74.

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

[15] I. H. Sloan, H. Woźniakowski, When are quasi-Monte Carlo algorithms efficient for high dimensional integrals?, J. Complexity 14 (1998) 1–33.

[16] I. H. Sloan, QMC integration – beating intractability by weighting the coordinate directions, in: K.-T. Fang, F. J. Hickernell, H. Niederreiter (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2000, Springer, New York, 2001, pp. 103–123.

[17] W. Hoeffding, A class of statistics with asymptotically normal distributions, Ann. Math. Stat. 19 (1948) 293–325.

[18] I. M. Sobol', Multidimensional Quadrature Formulas and Haar Functions, Nauka, Moskow, 1969, in Russian.

[19] R. E. Caflisch, W. Morokoff, A. B. Owen, Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension, J. Comp. Fin. 1 (1) (1997) 27–46.

[20] I. M. Sobol', Sensitivity estimates for nonlinear mathematical models, Math. Model. Comp. Exper. 1 (1993) 407–414, published in Russian in 1990.

[21] J. Dick, I. H. Sloan, X. Wang, H. Woźniakowski, Liberating the weights, J. Complexity 20 (2004) 593–623.

[22] C. Lemieux, P. L'Ecuyer, Selection criteria for lattice rules and other low-discrepancy point sets, Math. Comput. Sim. 55 (2001) 139–148.

[23] C. Lemieux, M. Cieslak, K. Luttmer, RandQMC user's guide: A package for randomized quasi-Monte Carlo methods in C, Tech. Rep. 2002-712-15, Department of Computer Science, University of Calgary (2002).

[24] R. Cools, F. Y. Kuo, D. Nuyens, Constructing embedded lattice rules for multivariate integration, SIAM Journal on Scientific Computing 28 (6) (2006) 2162–2188.

[25] H. S. Gill, C. Lemieux, A search for extensible Korobov rules, J. Complexity 23 (2007) 603–613.

[26] C. Lemieux, Monte Carlo and Quasi-Monte Carlo Sampling, Springer Series in Statistics, Springer, New York, 2009.

[27] I. M. Sobol', D. I. Asotsky, One more experiment on estimating high-dimensional integrals by quasi-Monte Carlo methods, Math. Comput. Simul. 62 (2003) 255–263.

[28] C. Lemieux, A. B. Owen, Quasi-regression and the relative importance of the ANOVA components of a function, in: K.-T. Fang, F. J. Hickernell, H. Niederreiter (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2000, Springer, New York, 2001, pp. 331–344.