REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Rev. Anal. Numér. Théor. Approx., vol. 35 (2006) no. 2, pp. 207–219
ictp.acad.ro/jnaat

GENERATION OF NON-UNIFORM LOW-DISCREPANCY SEQUENCES IN THE MULTIDIMENSIONAL CASE

NATALIA ROŞCA*

Abstract. In this paper, we extend the results we obtained in an earlier paper, from the one-dimensional case to the s-dimensional case. We propose two inversion type methods for generating G-distributed low-discrepancy sequences in $[0,1]^s$, where G is an arbitrary distribution function. Our methods are based on the approximation of the inverses of the marginal distribution functions using linear Lagrange interpolation or cubic Hermite interpolation. We also determine upper bounds for the G-discrepancy of the sequences we generate using the proposed methods.

MSC 2000. 65C05, 65C10, 11K38.
Keywords. Discrepancy, uniformly distributed sequences, G-discrepancy, G-distributed sequences, low-discrepancy sequences, Quasi-Monte Carlo integration, interpolation.

1. INTRODUCTION

We consider an s-dimensional continuous distribution on $[0,1]^s$, with distribution function G and density function g (g is nonnegative and $\int_{[0,1]^s} g(u)\mathrm{d}u = 1$). We are interested in generating G-distributed low-discrepancy sequences in $[0,1]^s$. We first recall some useful notions and results.

Definition 1 (discrepancy). *Let $P = (x_k)_{k\in\mathbb{N}^*}$ be a sequence of points in $[0,1]^s$. The discrepancy of the first N terms of sequence P is defined as*

$$D_N(x_1,\ldots,x_N) = \sup_{J\subseteq[0,1]^s} \left|\tfrac{1}{N}A_N(J,P) - \lambda_s(J)\right|,$$

where the supremum is calculated over all subintervals J of $[0,1]^s$ of the form $\prod_{i=1}^s [a_i,b_i]$; λ_s is the s-dimensional Lebesgue measure; $A_N(J,P)$ counts the number of elements of the set $(x_1,\ldots,x_N)$, falling into the interval J, i.e.

$$A_N(J,P) = \sum_{k=1}^{N} 1_J(x_k).$$

1_J is the characteristic function of J.

---

*"Babeş-Bolyai" University, Faculty of Mathematics and Computer Science, Str. Kogălniceanu 1, Cluj-Napoca, Romania, e-mail: natalia@math.ubbcluj.ro.

The sequence P is called uniformly distributed if $D_N(x_1, \ldots, x_N) \to 0$ *when* $N \to \infty$.

The uniformly distributed sequence P is said to be a low-discrepancy sequence if we have

$$D_N(x_1, \ldots, x_N) = \mathcal{O}\big((\log N)^s/N\big) \quad \text{for all } N \geq 2.$$

The discrepancy can be viewed as a measure for the deviation from the uniform distribution. Uniformly distributed low-discrepancy sequences are constructed in [4], [5], [6] and [10]. An overview on discrepancy and uniformly distributed low-discrepancy sequences is provided in [10]. The definition of discrepancy can be generalized in a straightforward way.

Definition 2 (G-discrepancy). *Consider an s-dimensional continuous distribution on* $[0,1]^s$, *with distribution function G. Let λ_G be the probability measure induced by G. Let $P = (x_k)_{k \in \mathbb{N}^*}$ be a sequence of points in* $[0,1]^s$. *The G-discrepancy of the first N terms of sequence P is defined as*

$$D_{N,G}(x_1, \ldots, x_N) = \sup_{J \subseteq [0,1]^s} \left| \tfrac{1}{N} A_N(J,P) - \lambda_G(J) \right|,$$

where the supremum is calculated over all subintervals J of $[0,1]^s$ *of the form* $\prod_{i=1}^s [a_i, b_i]$.

The sequence P is called G-distributed if $D_{N,G}(x_1, \ldots, x_N) \to 0$ *when* $N \to \infty$.

The G-distributed sequence P is said to be a low-discrepancy sequence if we have

$$D_{N,G}(x_1, \ldots, x_N) = \mathcal{O}\big((\log N)^s/N\big) \quad \text{for all } N \geq 2.$$

G-distributed low-discrepancy sequences are used in Quasi-Monte Carlo (QMC) integration, to approximate $\int_{[0,1]^s} f(x)\mathrm{d}G(x)$, where $f : [0,1]^s \to \mathbb{R}$. The integral $\int_{[0,1]^s} f(x)\mathrm{d}G(x)$ is approximated by $\frac{1}{N}\sum_{k=1}^N f(x_k)$, where $(x_k)_{k \in \mathbb{N}^*}$ is a G-distributed low-discrepancy sequence in $[0,1]^s$. If f is a function with finite variation in the sense of Hardy and Krause, then an upper bound for the error of approximation in QMC integration is given by the non-uniform Koksma-Hlawka inequality (see [1] or [11]). A detailed discussion on the variation in the sense of Hardy and Krause is given in [12].

Theorem 3 (non-uniform Koksma-Hlawka inequality). [1], [11]. *Let $f : [0,1]^s \to \mathbb{R}$ be a function of bounded variation in the sense of Hardy and Krause. Consider a distribution on* $[0,1]^s$, *with distribution function G. Then, for any* $x_1, \ldots, x_N \in [0,1]^s$, *we have*

$$\left| \frac{1}{N} \sum_{k=1}^N f(x_k) - \int_{[0,1]^s} f(x)\mathrm{d}G(x) \right| \leq V_{HK}(f) D^*_{N,G}(x_1, \ldots, x_N), \tag{1}$$

where $V_{HK}(f)$ is the variation of f in the sense of Hardy and Krause.

In the following, we are concerned with methods for generating G-distributed low-discrepancy sequences in $[0,1]^s$. The methods in this paper use the one-dimensional marginal distributions defined below.

DEFINITION 4. *Consider an s-dimensional continuous distribution on $[0,1]^s$, with density function g. For a point $u = (u^{(1)}, \dots, u^{(s)}) \in [0,1]^s$, the marginal density functions g_l, $l = 1, \dots, s$, are defined by*

(2)
$$g_l(u^{(l)}) = \\ = \underbrace{\int \dots \int}_{[0,1]^{s-1}} g(t^{(1)}, \dots, t^{(l-1)}, u^{(l)}, t^{(l+1)}, \dots t^{(s)}) \mathrm{d}t^{(1)} \dots \mathrm{d}t^{(l-1)} \mathrm{d}t^{(l+1)} \dots \mathrm{d}t^{(s)},$$

and the marginal distribution functions G_l, $l = 1, \dots, s$, are defined by

$$G_l(u^{(l)}) = \int_0^{u^{(l)}} g_l(t)\mathrm{d}t. \tag{3}$$

In this paper, we consider s-dimensional continuous distributions on $[0,1]^s$, with $g(u) = \prod_{l=1}^s g_l(u^{(l)})$, $\forall u = (u^{(1)}, \dots, u^{(s)}) \in [0,1]^s$. We assume that G_l, $l = 1, \dots, s$, are invertible on $[0,1]$.

THEOREM 5. [11]. *Let $\alpha = (\alpha_1, \dots, \alpha_N)$ be a set of points in $[0,1]^s$, with $\alpha_k = (\alpha_k^{(1)}, \dots, \alpha_k^{(s)})$, $k = 1, \dots, N$. Consider an s-dimensional continuous distribution on $[0,1]^s$, with distribution function G and density function g. Let g_l, $l = 1, \dots, s$, be the marginal density functions and assume that $g(u) = \prod_{l=1}^s g_l(u^{(l)})$, $\forall u = (u^{(1)}, \dots, u^{(s)}) \in [0,1]^s$. Furthermore, let G_l, $l = 1, \dots, s$, be the marginal distribution functions and assume that they are invertible on $[0,1]$. Construct the set of points $\beta = (\beta_1, \dots, \beta_N)$ in $[0,1]^s$, with $\beta_k = (\beta_k^{(1)}, \dots, \beta_k^{(s)})$, $k = 1, \dots, N$, by $\beta_k^{(1)} = G_1^{-1}(\alpha_k^{(1)})$, $\beta_k^{(2)} = G_2^{-1}(\alpha_k^{(2)})$, $\dots$, $\beta_k^{(s)} = G_s^{-1}(\alpha_k^{(s)})$, sequentially. Then the G-discrepancy of the constructed set of points is given by*

$$D_{N,G}(\beta_1, \dots, \beta_N) = D_N(\alpha_1, \dots, \alpha_N).$$

Based on Theorem 5, it follows that in order to generate G-distributed low-discrepancy sequences in $[0,1]^s$, we may proceed as follows. First, we consider a uniformly distributed low-discrepancy sequence $(x_k)_{k \in \mathbb{N}^*}$ in $[0,1]^s$ and then, we construct the sequence $(y_k)_{k \in \mathbb{N}^*}$, with $y_k = (y_k^{(1)}, \dots, y_k^{(s)})$, where

$$y_k^{(1)} = G_1^{-1}(x_k^{(1)}), \quad y_k^{(2)} = G_2^{-1}(x_k^{(2)}), \; \dots \; , y_k^{(s)} = G_s^{-1}(x_k^{(s)}), \quad k \in \mathbb{N}^*. \tag{4}$$

The constructed sequence $(y_k)_{k \in \mathbb{N}^*}$ is a G-distributed low-discrepancy sequence in $[0,1]^s$.

This modality of generating G-distributed low-discrepancy sequences in $[0,1]^s$ can be applied only if the analytical expressions of functions G_l^{-1}, $l = 1, \dots, s$, are known.

Very often, the inverse functions G_l^{-1}, $l = 1, \dots, s$, are not given analytically. In this case, one may use approximations of the inverse functions G_l^{-1}, $l = 1, \dots, s$. Several ways of approximating these functions are proposed in the literature.

The method proposed in Hlawka [8] is based on the following result.

Theorem 6. [8]. *Consider an s-dimensional continuous distribution on $[0,1]^s$, with distribution function G and density function $g(u) = \prod_{j=1}^{s} g_j(u^{(j)})$, $\forall u = (u^{(1)}, \dots, u^{(s)}) \in [0,1]^s$. Assume that $g_j(t) \neq 0$, for almost every $t \in [0,1]$ and for all $j = 1, \dots, s$. Furthermore, assume that g_j, $j = 1, \dots, s$, are continuous on $[0,1]$. Denote by $M_g = \sup_{u\in[0,1]^s} g(u)$. Let $(x_1, \dots, x_N)$ be a set of points in $[0,1]^s$. Generate the set of points $(y_1, \dots, y_N)$, with*

$$y_k^{(j)} = \frac{1}{N} \sum_{r=1}^{N} [1 + x_k^{(j)} - G_j(x_r^{(j)})] = \frac{1}{N} \sum_{r=1}^{N} 1_{[0,x_k^{(j)}]}(G_j(x_r^{(j)})), \tag{5}$$

for all $k = 1, \dots, N$ and all $j = 1, \dots, s$, where $[a]$ denotes the integer part of a. Then the generated set of points has a G-discrepancy of

$$D_{N,G}(y_1, \dots, y_N) \le (2 + 6sM_g) D_N(x_1, \dots, x_N). \tag{6}$$

In Theorem 6, the assumption that $g_j(t) \neq 0$, for almost every $t \in [0,1]$, implies that G_j is strictly increasing on $[0,1]$, for all $j = 1, \dots, s$. Based on this, and on the fact that $G_j(0) = 0$, $G_j(1) = 1$ and G_j is continuous on $[0,1]$ (as it is absolutely continuous), it follows that G_j is invertible on $[0,1]$, for all $j = 1, \dots, s$. The assumption that g_j, $j = 1, \dots, s$, are continuous on $[0,1]$ implies that $M_g < \infty$.

The disadvantage of the Hlawka method is that the resulting set of points is generated on a grid with spacing $1/N$. This implies that, when adding some points, all the other points have to be regenerated.

Hartinger and Kainhofer [7] avoid the grid structure in the Hlawka method. They propose a transformation and bound the G-discrepancy of the transformed set of points as follows.

Theorem 7. [7]. *Let $(x_1, \dots, x_N)$ be a set of points in $[0,1]^s$ and $P = (z_1, \dots, z_N)$ be a set of points in $[0,1]$. Let G be a distribution function on $[0,1]^s$, with bounded, continuous density $g(u) = \prod_{l=1}^{s} g_l(u^{(l)})$, $\forall u = (u^{(1)}, \dots, u^{(s)}) \in [0,1]^s$, and $g_l(u^{(l)}) \le M < \infty$ for all l. Furthermore, assume that the marginal distribution functions G_l, $l = 1, \dots, s$, are invertible on $[0,1]$. Define for $k = 1, \dots, N$ and $l = 1, \dots, s$ the values*

$$z_k^{(l)-} = \max_{\mathcal{A}=\{z_n\in P | G_l(z_n) \le x_k^{(l)}\}} z_n \qquad \text{and} \qquad z_k^{(l)+} = \min_{\mathcal{B}=\{z_n\in P | x_k^{(l)} \le G_l(z_n)\}} z_n.$$

Set $z_k^{(l)-} = 0$ if $\mathcal{A} = \emptyset$ and $z_k^{(l)+} = 1$ if $\mathcal{B} = \emptyset$.

Then the G-discrepancy of any transformed set of points $(y_1, \dots, y_N)$, with the property that $y_k^{(l)} \in (z_k^{(l)-}, z_k^{(l)+}]$ for all $1 \le k \le N$ and all $1 \le l \le s$, is *bounded by*

$$D_{N,G}(y_1,\dots,y_N) \le D_N(x_1,\dots,x_N) + (1+2M)^s D_N(z_1,\dots,z_N). \tag{7}$$

In the method of Hartinger and Kainhofer, any value $y_k^{(l)} \in \left(z_k^{(l)-}, z_k^{(l)+}\right]$ can be considered as $G_l^{-1}(x_k^{(l)})$. They do not analyze the possibility of approximating G_l^{-1} using interpolation methods.

2. AN INVERSION TYPE METHOD USING LAGRANGE INTERPOLATION

In the following, we propose an inversion type method for generating G-distributed low-discrepancy sequences in $[0,1]^s$. The method is based on the approximation of the inverse functions G_l^{-1}, $l = 1,\dots,s$, using linear Lagrange interpolation. We also determine upper bounds for the G-discrepancy of the sequence we generate using the proposed method. The method extends the results we obtained in an earlier paper [13], from the one-dimensional case to the s-dimensional case.

The following results are needed to prove the main results of this section.

Lemma 8. [9]. *Let $P_1 = (u_1,\dots,u_N)$ and $P_2 = (v_1,\dots,v_N)$ be two sets of points in $[0,1]^s$. If for all $1 \le l \le s$ and all $1 \le k \le N$ the condition*

$$\left| u_k^{(l)} - v_k^{(l)} \right| \le \varepsilon_l \tag{8}$$

holds for some values ε_l, we get the following bound on the difference of the discrepancies

$$\left| D_N(u_1,\dots,u_N) - D_N(v_1,\dots,v_N) \right| \le \prod_{l=1}^{s} (1+2\varepsilon_l) - 1. \tag{9}$$

Proposition 9. [13]. *Let $(z_1, z_2, \dots, z_N)$ be a set of points in $[0,1]$, with $z_0 := 0 \le z_1 < z_2 < \ldots < z_N \le 1 =: z_{N+1}$. The following inequality holds*

$$|z_n - z_{n+1}| \le D_N(z_1,\dots,z_N), \qquad n = 0,\dots,N. \tag{10}$$

We have the following main result, in which we bound the G-discrepancy of the set of points that will be generated.

Theorem 10. *Let $(x_1,\dots,x_N)$ be a set of points in $[0,1]^s$ and $(z_1, z_2, \dots, z_N)$ be a set of points in $[0,1]$, with $0 \le z_1 < z_2 < \ldots < z_N \le 1$. We define $z_0 = 0$ and $z_{N+1} = 1$. Consider an s-dimensional continuous distribution on $[0,1]^s$, with density function g and distribution function G. Let g_l, $l = 1,\dots,s$, be the marginal density functions and assume that $g(u) = \prod_{l=1}^{s} g_l(u^{(l)})$, $\forall u = (u^{(1)},\dots,u^{(s)}) \in [0,1]^s$. Furthermore, assume that $\sup_{t\in[0,1]} g_l(t) \le M < \infty$, $\forall l = 1,\dots,s$, and $g_l(t) \ne 0$, $\forall t \in [0,1], \forall l = 1,\dots,s$. Let G_l, $l = 1,\dots,s$, be the marginal distribution functions. For $k = 1,\dots,N$ and $l = 1,\dots,s$, determine the interval $\left(z_k^{(l)-}, z_k^{(l)+}\right]$ of the form $(z_n, z_{n+1}]$, $n \in \{0,1,\dots,N\}$, such that*

$$G_l(z_k^{(l)-}) < x_k^{(l)} \le G_l(z_k^{(l)+}).$$

Generate the set of points $(y_1, \ldots, y_N)$ *in* $[0,1]^s$*, given by*

$$y_k^{(l)} = \frac{x_k^{(l)} - G_l(z_k^{(l)+})}{G_l(z_k^{(l)-}) - G_l(z_k^{(l)+})} z_k^{(l)-} + \frac{x_k^{(l)} - G_l(z_k^{(l)-})}{G_l(z_k^{(l)+}) - G_l(z_k^{(l)-})} z_k^{(l)+}, \tag{11}$$

for all $k = 1, \ldots, N$ *and all* $l = 1, \ldots, s$.

If $G_l \in C^2[0,1]$ *and* $\|\frac{g_l'}{g_l^3}\|_\infty \leq L$ *for all* $1 \leq l \leq s$*, then the* G*-discrepancy of the constructed set of points* $(y_1, y_2, \ldots, y_N)$ *is bounded by*

$$D_{N,G}(y_1, \ldots, y_N) \leq D_N(x_1, \ldots, x_N) + \left((1+M^3L)^s - 1\right) D_N(z_1, \ldots, z_N). \tag{12}$$

Proof. The proof proceeds in the following two steps.

Step 1. For each $l = 1, \ldots, s$, we consider the one-dimensional projections $(x_1^{(l)}, \ldots, x_N^{(l)})$.

The assumption that $g_l(t) \neq 0$, $\forall t \in [0,1]$, implies that G_l is strictly increasing on $[0,1]$. From the expression (3) of the marginal distribution function G_l, it follows that $G_l(0) = 0$, $G_l(1) = 1$ and G_l is continuous on $[0,1]$ (as it is absolutely continuous). Based on these considerations, it follows that G_l is invertible on $[0,1]$.

In order to approximate $G_l^{-1}(x_k^{(l)})$, $k = 1, \ldots, N$, we determine the interval $(z_k^{(l)-}, z_k^{(l)+}]$, as described in the statement of the theorem. As $G_l(z_k^{(l)-}) < x_k^{(l)} \leq G_l(z_k^{(l)+})$, it follows that $G_l^{-1}(x_k^{(l)}) \in (z_k^{(l)-}, z_k^{(l)+}]$. We will approximate $G_l^{-1}(x_k^{(l)})$ with a value $y_k^{(l)} \in (z_k^{(l)-}, z_k^{(l)+}]$, which is calculated using a linear Lagrange interpolation of G_l^{-1} with nodes $G_l(z_k^{(l)-})$ and $G_l(z_k^{(l)+})$. The interpolation formula is

$$G_l^{-1} = L_1 G_l^{-1} + R_1 G_l^{-1},$$

where $L_1 G_l^{-1}$ is the Lagrange interpolation polynomial of degree 1 and $R_1 G_l^{-1}$ is the remainder. The values of G_l^{-1} at the nodes are $G_l^{-1}(G_l(z_k^{(l)-})) = z_k^{(l)-}$ and $G_l^{-1}(G_l(z_k^{(l)+})) = z_k^{(l)+}$. Using the expression of the Lagrange interpolation polynomial, we obtain

$$(L_1 G_l^{-1})(x_k^{(l)}) = \frac{x_k^{(l)} - G_l(z_k^{(l)+})}{G_l(z_k^{(l)-}) - G_l(z_k^{(l)+})} z_k^{(l)-} + \frac{x_k^{(l)} - G_l(z_k^{(l)-})}{G_l(z_k^{(l)+}) - G_l(z_k^{(l)-})} z_k^{(l)+} = y_k^{(l)}.$$

As a consequence, we approximate $G_l^{-1}(x_k^{(l)})$ with $y_k^{(l)}$. The approximation error (see [14]) is bounded by

$$|G_l^{-1}(x_k^{(l)}) - y_k^{(l)}| = |R_1(G_l^{-1})(x_k^{(l)})| \leq \frac{|u(x_k^{(l)})|}{2!} \|(G_l^{-1})''\|_\infty, \tag{13}$$

where

$$|u(x_k^{(l)})| = |(x_k^{(l)} - G_l(z_k^{(l)-}))(x_k^{(l)} - G_l(z_k^{(l)+}))|,$$

and

$$(G_l^{-1})'' = -\frac{g_l'(G_l^{-1})}{(g_l(G_l^{-1}))^3}, \qquad ||(G_l^{-1})''||_\infty = \left\| \frac{g_l'}{g_l^3} \right\|_\infty. \tag{14}$$

As $G_l(z_k^{(l)-}) < x_k^{(l)} \leq G_l(z_k^{(l)+})$, we obtain

$$|x_k^{(l)} - G_l(z_k^{(l)-})| \leq |G_l(z_k^{(l)+}) - G_l(z_k^{(l)-})| = \left| \int_{z_k^{(l)-}}^{z_k^{(l)+}} g_l(t)\mathrm{d}t \right| \leq M|z_k^{(l)+} - z_k^{(l)-}|. \tag{15}$$

Since $(z_k^{(l)-}, z_k^{(l)+}]$ is an interval of type $(z_n, z_{n+1}]$, we apply Proposition 9 and we get

$$|z_k^{(l)+} - z_k^{(l)-}| = |z_{n+1} - z_n| \leq D_N(z_1, \dots, z_N).$$

Relation (15) becomes

$$|x_k^{(l)} - G_l(z_k^{(l)-})| \leq MD_N(z_1, \dots, z_N).$$

Similarly, we get

$$|x_k^{(l)} - G_l(z_k^{(l)+})| \leq MD_N(z_1, \dots, z_N).$$

It follows that

$$|u(x_k^{(l)})| = |(x_k^{(l)} - G_l(z_k^{(l)-}))(x_k^{(l)} - G_l(z_k^{(l)+}))| \leq M^2 D_N^2(z_1, \dots, z_N). \tag{16}$$

Using (16) and (14), relation (13) becomes

$$|G_l^{-1}(x_k^{(l)}) - y_k^{(l)}| \leq \frac{M^2 D_N^2(z_1, \dots, z_N)}{2} \left\| \frac{g_l'}{g_l^3} \right\|_\infty \leq \frac{M^2 D_N^2(z_1, \dots, z_N)}{2} L. \tag{17}$$

On the other hand, we have

$$\begin{aligned} |G_l(y_k^{(l)}) - x_k^{(l)}| &= |G_l(y_k^{(l)}) - G_l(G_l^{-1}(x_k^{(l)}))| = \left| \int_{G_l^{-1}(x_k^{(l)})}^{y_k^{(l)}} g_l(t)\mathrm{d}t \right| \\ &\leq M|G_l^{-1}(x_k^{(l)}) - y_k^{(l)}|. \end{aligned} \tag{18}$$

Substituting (17) into (18) and using $(D_N)^2 \leq D_N$, as $D_N \leq 1$, we obtain

$$|G_l(y_k^{(l)}) - x_k^{(l)}| \leq \frac{M^3 D_N^2(z_1, \dots, z_N)}{2} L \leq \frac{M^3 D_N(z_1, \dots, z_N)}{2} L. \tag{19}$$

Step 2. We know that condition (19) is verified for all $l = 1, \dots, s$ and all $k = 1, \dots, N$. We apply Lemma 8, with $\varepsilon_l = \frac{M^3 L D_N(z_1, \dots, z_N)}{2} = \varepsilon$, $P_1 = (v_1, \dots, v_N)$, with $v_k = (G_l(y_k^{(l)}))_{l=1\dots,s}$, $k = 1, \dots, N$ and $P_2 = (x_1, \dots, x_N)$, with $x_k = (x_k^{(l)})_{l=1\dots,s}$, $k = 1, \dots, N$. We obtain

$$|D_N(P_1) - D_N(P_2)| \leq \prod_{l=1}^{s} (1 + 2\varepsilon_l) - 1 = (1 + 2\varepsilon)^s - 1. \tag{20}$$

Next, we apply Theorem 5 and we get $D_N(P_1) = D_{N,G}(y_1, \dots, y_N)$. Formula (20) becomes

$$D_{N,G}(y_1, \dots, y_N) \le D_N(x_1, \dots, x_N) + (1 + 2\varepsilon)^s - 1.$$

We expand the binomial term and we use $(D_N)^i \le D_N$, as $D_N \le 1$. We get

$$\begin{aligned} D_{N,G}(y_1, \dots, y_N) &\le D_N(x_1, \dots, x_N) + \left(1 + M^3 L D_N(z_1, \dots, z_N)\right)^s - 1 \\ &\le D_N(x_1, \dots, x_N) + \sum_{k=1}^{s} C_s^k \left(M^3 L D_N(z_1, \dots, z_N)\right)^k \\ &\le D_N(x_1, \dots, x_N) + D_N(z_1, \dots, z_N) \sum_{k=1}^{s} C_s^k \left(M^3 L\right)^k \\ &= D_N(x_1, \dots, x_N) + ((1 + M^3 L)^s - 1) D_N(z_1, \dots, z_N). \end{aligned}$$

□

Based on Theorem 10, it follows that for generating G-distributed low-discrepancy sequences in $[0,1]^s$, we may proceed as follows. First, we consider a uniformly-distributed low-discrepancy sequence $(x_k)_{k\in\mathbb{N}^*}$ in $[0,1]^s$. Then, we construct the sequence $(y_k)_{k\in\mathbb{N}^*}$, given by (11). The sequence $(y_k)_{k\in\mathbb{N}^*}$ is a G-distributed low-discrepancy sequence in $[0,1]^s$, as we prove in the following result.

THEOREM 11. *Let $(x_k)_{k\in\mathbb{N}^*}$ be a uniformly distributed low-discrepancy sequence in $[0,1]^s$ and $(z_k)_{k\in\mathbb{N}^*}$ be an increasing uniformly distributed low-discrepancy sequence in $[0,1]$. Consider a distribution on $[0,1]^s$, with distribution function G and density function g, that verify the conditions in Theorem* 10. *Construct the sequence $(y_k)_{k\in\mathbb{N}^*}$, with $y_k = (y_k^{(1)}, \dots, y_k^{(s)})$, $k \in \mathbb{N}^*$, where $y_k^{(l)}$ is given by* (11), *for all $k \in \mathbb{N}^*$ and all $l = 1, \dots, s$. Then $(y_k)_{k\in\mathbb{N}^*}$ is a G-distributed low-discrepancy sequence in $[0,1]^s$.*

Proof. As $(x_k)_{k\in\mathbb{N}^*}$ and $(z_k)_{k\in\mathbb{N}^*}$ are uniformly distributed sequences in $[0,1]^s$ and $[0,1]$, respectively, we have $\lim_{N\to\infty} D_N(x_1, \dots, x_N) = 0$ and $\lim_{N\to\infty} D_N(z_1, \dots, z_N) = 0$. Using formula (12), we get

$$\begin{aligned} &\lim_{N\to\infty} D_{N,G}(y_1, \dots, y_N) \le \\ &\quad \le \lim_{N\to\infty} D_N(x_1, \dots, x_N) + \left((1 + M^3 L)^s - 1\right) \lim_{N\to\infty} D_N(z_1, \dots, z_N) = 0. \end{aligned} \tag{21}$$

This implies $\lim_{N\to\infty} D_{N,G}(y_1, \dots, y_N) = 0$. According to Definition 2, $(z_k)_{k\in\mathbb{N}^*}$ is a G-distributed sequence in $[0,1]^s$.

Furthermore, as $(x_k)_{k\in\mathbb{N}^*}$ and $(z_k)_{k\in\mathbb{N}^*}$ are low-discrepancy sequences in $[0,1]^s$ and $[0,1]$, respectively, we have

$$D_N(x_1, \dots, x_N) = \mathcal{O}\big((\log N)^s/N\big), \quad D_N(z_1, \dots, z_N) = \mathcal{O}\big((\log N)/N\big),$$

for all $N \ge 2$.

Using (12), we get $D_{N,G}(y_1,\dots,y_N) = \mathcal{O}\big((\log N)^s/N\big)$ for all $N \geq 2$. According to Definition 2, $(y_k)_{k\in\mathbb{N}^*}$ is a low-discrepancy sequence in $[0,1]^s$. $\square$

In our method, by adding a new point, the elements already generated remain unchanged, which represents an advantage. We notice that the analytical expressions of functions G_l, g_l, $g_l^{'}$, $l = 1,\dots,s$, have to be known. These functions have to verify the conditions in Theorem 10. The generation of a G-distributed low-discrepancy sequence in $[0,1]^s$ is described in Algorithm 12.

Algorithm 12. *An inversion type method based on the approximation of functions G_l^{-1}, $l = 1,\dots,s$, using linear Lagrange interpolation*

Input data:

- the integer N;
- the set of points $(x_1,\dots,x_N)$, consisting of the first N terms of a uniformly distributed low-discrepancy sequence $(x_k)_{k\in\mathbb{N}^*}$ in $[0,1]^s$;
- the set of points $(z_1,z_2,\dots,z_N)$, consisting of the first N terms of a uniformly distributed low-discrepancy sequence in $[0,1]$;
- the marginal distribution functions G_l, $l = 1,\dots,s$;

Step 1. Sort increasingly the elements of the set of points $(z_1,\dots,z_N)$.
Step 2.
for $l = 1,\dots,s$ **do**
  for $k = 1,\dots,N$ **do**
    Compute the values $z_k^{(l)-}$ şi $z_k^{(l)+}$, as described in Theorem 10;
    Compute the value $y_k^{(l)}$, given by formula (11).
  end for
end for
Output data: the set of points $(y_1,\dots,y_N)$, with $y_k = \big(y_k^{(1)},\dots,y_k^{(s)}\big)$, $k = 1,\dots,N$, consisting of the first N terms of a G-distributed low-discrepancy sequence in $[0,1]^s$.

3. AN INVERSION TYPE METHOD USING HERMITE INTERPOLATION

Next, we propose an inversion type method that is based on the approximation of the inverse functions G_l^{-1}, $l = 1,\dots,s$, using cubic Hermite interpolation. The method can be used to generate G-distributed low-discrepancy sequences in $[0,1]^s$. We obtain the following main result.

Theorem 13. *Consider the same hypotheses as in Theorem* 10. *For $k = 1,\dots,N$ and $l = 1,\dots,s$, determine the interval $\big(z_k^{(l)-}, z_k^{(l)+}\big]$ of the form $(z_n,z_{n+1}]$, $n \in \{0,1,\dots,N\}$, such that*

$$G_l\big(z_k^{(l)-}\big) < x_k^{(l)} \leq G_l\big(z_k^{(l)+}\big).$$

Generate the set of points $(y_1,\dots,y_N)$ in $[0,1]^s$, given by

$$(22)\quad y_k^{(l)} = h_{00}\big(x_k^{(l)}\big) z_k^{(l)-} + h_{10}\big(x_k^{(l)}\big) z_k^{(l)+} + h_{01}\big(x_k^{(l)}\big)\frac{1}{g_l\big(z_k^{(l)-}\big)} + h_{11}\big(x_k^{(l)}\big)\frac{1}{g_l\big(z_k^{(l)+}\big)},$$

where

$$\begin{aligned}
h_{00}(x_k^{(l)}) &= \frac{\left(x_k^{(l)}-G_l\left(z_k^{(l)+}\right)\right)^2}{\left(G_l\left(z_k^{(l)-}\right)-G_l\left(z_k^{(l)+}\right)\right)^2}\left(1-2\frac{x_k^{(l)}-G_l\left(z_k^{(l)-}\right)}{G_l\left(z_k^{(l)-}\right)-G_l\left(z_k^{(l)+}\right)}\right),\\
h_{10}(x_k^{(l)}) &= \frac{\left(x_k^{(l)}-G_l\left(z_k^{(l)-}\right)\right)^2}{\left(G_l\left(z_k^{(l)+}\right)-G_l\left(z_k^{(l)-}\right)\right)^2}\left(1-2\frac{x_k^{(l)}-G_l\left(z_k^{(l)+}\right)}{G_l\left(z_k^{(l)+}\right)-G_l\left(z_k^{(l)-}\right)}\right),\\
h_{01}(x_k^{(l)}) &= \frac{\left(x_k^{(l)}-G_l\left(z_k^{(l)-}\right)\right)\left(x_k^{(l)}-G_l\left(z_k^{(l)+}\right)\right)^2}{\left(G_l\left(z_k^{(l)-}\right)-G_l\left(z_k^{(l)+}\right)\right)^2},\\
h_{11}(x_k^{(l)}) &= \frac{\left(x_k^{(l)}-G_l\left(z_k^{(l)+}\right)\right)\left(x_k^{(l)}-G_l\left(z_k^{(l)-}\right)\right)^2}{\left(G_l\left(z_k^{(l)+}\right)-G_l\left(z_k^{(l)-}\right)\right)^2},
\end{aligned}$$

for all $k = 1, \ldots, N$ *and all* $l = 1, \ldots, s$.

If $G_l \in C^4[0,1]$ *and*

$$\left\| \frac{g_l''' g_l^2 - 10 g_l'' g_l' g_l + 15 g_l'^3}{g_l^7} \right\|_\infty \le L$$

for all $1 \le l \le s$*, then the* G*-discrepancy of the constructed set of points* $(y_1, \ldots, y_N)$ *is bounded by*

$$D_{N,G}(y_1,\ldots,y_N) \le D_N(x_1,\ldots,x_N) + \left(\left(1+\tfrac{M^5 L}{12}\right)^s - 1\right) D_N(z_1,\ldots,z_N). \tag{23}$$

Proof. We follow the same steps as in Theorem 10. We give only the differences.

Step 1. We consider the one-dimensional projections $\left(x_1^{(l)}, \ldots, x_N^{(l)}\right)$, for each $l = 1, \ldots, s$. We will approximate $G_l^{-1}(x_k^{(l)})$, $k = 1, \ldots, N$, with a value $y_k^{(l)} \in \left(z_k^{(l)-}, z_k^{(l)+}\right]$, which is calculated using a cubic Hermite interpolation of G_l^{-1} with double nodes $G_l(z_k^{(l)-})$ and $G_l(z_k^{(l)+})$. The values of G_l^{-1} and $(G_l^{-1})'$ at the nodes are $G_l^{-1}(G_l(z_k^{(l)-})) = z_k^{(l)-}$, $G_l^{-1}(G_l(z_k^{(l)+})) = z_k^{(l)+}$, $(G_l^{-1})'(G_l(z_k^{(l)-})) = \frac{1}{g_l\left(z_k^{(l)-}\right)}$ and $(G_l^{-1})'(G_l(z_k^{(l)+})) = \frac{1}{g_l\left(z_k^{(l)+}\right)}$. The Hermite interpolation formula is

$$G_l^{-1} = H_3 G_l^{-1} + R_3 G_l^{-1},$$

where $H_3 G_l^{-1}$ is the Hermite interpolation polynomial of degree 3 and $R_3 G_l^{-1}$ is the remainder. Using the expression of the Hermite polynomial with double nodes (see [2]), it can be proved, after some calculus, that

$$(H_3 G_l^{-1})(x_k^{(l)}) = y_k^{(l)}.$$

As a consequence, we approximate $G_l^{-1}(x_k^{(l)})$ with $y_k^{(l)}$. The bound for the approximation error (see [2]) is

$$|R_3(G_l^{-1})(x_k^{(l)})| = |G_l^{-1}(x_k^{(l)}) - y_k^{(l)}| \le \frac{|u(x_k^{(l)})|}{4!} \|(G_l^{-1})^{(4)}\|_\infty, \tag{24}$$

where

$$\left|u(x_k^{(l)})\right| = \left|(x_k^{(l)} - G_l(z_k^{(l)-}))^2(x_k^{(l)} - G_l(z_k^{(l)+}))^2\right| \le M^4 D_N^4(z_1,\dots,z_N),$$

and

$$\left\|(G_l^{-1})^{(4)}\right\|_\infty = \left\|\frac{g_l'''g_l^2 - 10g_l''g_l'g_l + 15g_l'^3}{g_l^7}\right\|_\infty \le L.$$

Relation (24) becomes

$$|G_l^{-1}(x_k^{(l)}) - y_k^{(l)}| \le \frac{M^4 D_N^4(z_1,\dots,z_N)}{4!}L.$$

As in Theorem 10 and using $(D_N)^4 \le D_N$, since $D_N \le 1$, we obtain

$$|G_l(y_k^{(l)}) - x_k^{(l)}| \le \frac{M^5 L}{24} D_N^4(z_1,\dots,z_N) \le \frac{M^5 L}{24} D_N(z_1,\dots,z_N). \tag{25}$$

Step 2. We know that condition (25) is verified for all $l = 1,\dots,s$ and all $k = 1,\dots,N$. We apply Lemma 8, with $\varepsilon_l = \frac{M^5 L D_N(z_1,\dots,z_N)}{24} = \varepsilon$, $P_1 = (v_1,\dots,v_N)$, with $v_k = \left(G_l(y_k^{(l)})\right)_{l=1\dots,s}$, $k = 1,\dots,N$ and $P_2 = (x_1,\dots,x_N)$, with $x_k = \left(x_k^{(l)}\right)_{l=1\dots,s}$, $k = 1,\dots,N$. We obtain

$$|D_N(P_1) - D_N(P_2)| \le \prod_{l=1}^{s}(1 + 2\varepsilon_l) - 1 = (1+2\varepsilon)^s - 1. \tag{26}$$

From Theorem 5, we get $D_N(P_1) = D_{N,G}(y_1,\dots,y_N)$. Relation (26) becomes

$$D_{N,G}(y_1,\dots,y_N) \le D_N(x_1,\dots,x_N) + (1+2\varepsilon)^s - 1.$$

We expand the binomial term and we use $(D_N)^i \le D_N$, as $D_N \le 1$. We obtain

$$D_{N,G}(y_1,\dots,y_N) \le D_N(x_1,\dots,x_N) + \Big((1 + \tfrac{M^5 L}{12})^s - 1\Big) D_N(z_1,\dots,z_N).$$

□

Based on Theorem 13, it follows that in order to generate G-distributed low-discrepancy sequences in $[0,1]^s$, we may proceed as follows. First, we consider a uniformly-distributed low-discrepancy sequence $(x_k)_{k\in\mathbb{N}^*}$ in $[0,1]^s$. Then, we construct the sequence $(y_k)_{k\in\mathbb{N}^*}$, given by (22). The sequence $(y_k)_{k\in\mathbb{N}^*}$ is a G-distributed low-discrepancy sequence in $[0,1]^s$, as we show in the result bellow.

THEOREM 14. *Let $(x_k)_{k\in\mathbb{N}^*}$ be a uniformly distributed low-discrepancy sequence in $[0,1]^s$ and $(z_k)_{k\in\mathbb{N}^*}$ be an increasing uniformly distributed low-discrepancy sequence in $[0,1]$. Consider a continuous distribution on $[0,1]^s$, with distribution function G and density function g, that verify the conditions in Theorem 13. Construct the sequence $(y_k)_{k\in\mathbb{N}^*}$, with $y_k = (y_k^{(1)},\dots,y_k^{(s)})$, $k \in \mathbb{N}^*$, where $y_k^{(l)}$ is given by (22), for all $k \in \mathbb{N}^*$ and all $l = 1,\dots,s$. Then $(y_k)_{k\in\mathbb{N}^*}$ is a G-distributed low-discrepancy sequence in $[0,1]^s$.*

Proof. The proof is similar to the one given in Theorem 11. □

In our method, the analytical expressions of functions G_l, g_l, $g_l^{'}$, $g_l^{''}$, $g_l^{'''}$, $l = 1, \ldots, s$, have to be known. These functions must verify the conditions in Theorem 13. The generation of a G-distributed low-discrepancy sequence in $[0,1]^s$ is described in Algorithm 15.

ALGORITHM 15. *An inversion type method based on the approximation of functions G_l^{-1}, $l = 1, \ldots, s$, using cubic Hermite interpolation*

We use the same *Input data* as in Algorithm 12.
Step 1. Sort increasingly the elements of the set of points $(z_1, \ldots, z_N)$.
Step 2.
for $l = 1, \ldots, s$ **do**
  for $k = 1, \ldots, N$ **do**
    Compute the values $z_k^{(l)-}$ şi $z_k^{(l)+}$, as described in Theorem 13;
    Compute the value $y_k^{(l)}$, given by formula (22).
  end for
end for
Output data: the set of points $(y_1, \ldots, y_N)$, with $y_k = (y_k^{(1)}, \ldots, y_k^{(s)})$, $k = 1, \ldots, N$, consisting of the first N terms of a G-distributed low-discrepancy sequence in $[0,1]^s$.

We conclude that the proposed inversion type methods using linear Lagrange interpolation or cubic Hermite interpolation generate G-distributed low-discrepancy sequences in $[0,1]^s$. The proposed methods are recommended when the inverses of the marginal distribution functions cannot be given explicitly in analytical form.

REFERENCES

[1] CHELSON, P., *Quasi-Random Techniques for Monte Carlo Methods*, Ph.D Dissertation, The Claremont Graduate School, 1976.

[2] COMAN, GH., *Numerical Analysis*, Ed. Libris, Cluj-Napoca, 1995 (in Romanian).

[3] DEAK, I., *Random Number Generators and Simulation*, Akademiai Kiado, Budapest, 1990.

[4] FAURE, H., *Discrépances de suites associées à un système de numération* (en dimension un), Bull. Soc. Math. France, 109 (1981), 143–182.

[5] FAURE, H., *Discrépances de suites associées à un système de numération* (en dimension s), Acta Arith., 41 (1982), 337–351.

[6] HALTON, J.H., *On the efficiency of certain quasi-random sequences of points in evaluating multidimensional integrals*, Numer. Math., 2 (1960), 84–90.

[7] HARTINGER, J. and KAINHOFER, R., *Non-Uniform Low-Discrepancy Sequence Generation and Integration of Singular Integrands*, Proceedings of Monte Carlo and Quasi-Monte Carlo Methods 2004, H. Niederreiter, eds., Springer-Verlag, Berlin, 2006, 163–180.

[8] HLAWKA, E., *Gleichverteilung und Simulation*, Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II, 206 (1997), 183–216.

[9] Hlawka, E. and Mück, R., *Über Eine Transformation von gleichverteilten Folgen*, II, Computing, 9 (1972), 127–138.

[10] Niederreiter, H., *Random number generation and Quasi-Monte Carlo methods*, Society for Industrial and Applied Mathematics, Philadelphia, 1992.

[11] Okten, G., *Error Reduction Techniques in Quasi-Monte Carlo Integration*, Math. Comput. Modelling, Vol. **30**, Nos. 7–8 (1999), 61–69.

[12] Owen, A.B., *Multidimensional variation for Quasi-Monte Carlo*, Technical report, Standford University, 2004.

[13] Roşca, N., *Generation of Non-Uniform Low-Discrepancy Sequences in Quasi-Monte Carlo Integration* (in dimension one), Stud. Univ. Babeş-Bolyai Math., Vol. **L**, no. 2 (2005), 77–90.

[14] Stancu, D.D., Coman, Gh., Agratini, O. and Trîmbiţaş, R., *Numerical Analysis and Approximation Theory*, Vol. **I**, Presa Universitara Clujeana, Cluj-Napoca, 2001 (in Romanian).

Received by the editors: February 3, 2006.