ON QUASI-MONTE CARLO SIMULATION
OF STOCHASTIC DIFFERENTIAL EQUATIONS

NORBERT HOFMANN AND PETER MATHÉ

ABSTRACT. In a number of problems of mathematical physics and other fields stochastic differential equations are used to model certain phenomena. Often the solution of those problems can be obtained as a functional of the solution of some specific stochastic differential equation. Then we may use the idea of weak approximation to carry out numerical simulation.

We analyze some complexity issues for a class of linear stochastic differential equations (Langevin type), which can be given by

$$dX_t = -\alpha X_t dt + \beta(t)dW_t, \quad X_0 := 0,$$

where $\alpha > 0$ and $\beta \colon [0,T] \to \mathbb{R}$. It turns out that for a class of input data which are not more than Lipschitz continuous the explicit Euler scheme gives rise to an optimal (by order) numerical method.

Then we study numerical phenomena which occur when switching from (real) Monte Carlo simulation to quasi–Monte Carlo simulation, which is the case when we carry out the simulation on computers. It will easily be seen that completely uniformly distributed sequences yield good substitutes for random variates, while not all uniformly distributed (mod 1) sequences are suited. In fact we provide necessary conditions on a sequence in order to serve quasi–Monte Carlo purposes. This condition is expressed in terms of the measure of well-distributions. Numerical examples complement the theoretical analysis.

1. INTRODUCTION

The purpose of this paper is the study of weak solutions of Ito stochastic differential equations which are usually written as

$$(1.1) \qquad X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s,$$

where the first summand in (1.1) is the initial value, the second one represents the drift and the third one, which shall be understood as integral in the sense of Ito, driven by some Brownian motion $\{W_s, \quad 0 \le s \le T\}$, the diffusion of the stochastic process which solves equation (1.1). Here the functions a and σ have to fulfill certain regularity conditions in order to ensure existence and uniqueness of solutions.

The numerical treatment of stochastic differential equations consists of two major branches. First, in case the driving Brownian motion is given, one may attempt to solve the corresponding equation by use of some (deterministic) scheme. The corresponding result will then be called strong solution. In a second approach one may

---

think of the given stochastic differential equation as a defining relation and some functional of the solution has a certain meaning only. The latter situation is met often in connection with partial differential equations, when we may represent the solution of some boundary value problem in terms of a functional of some stochastic differential equation. Thus we aim at approximating the distribution of a stochastic differential equation rather than finding solutions to some particular realization. In this context one may speak of weak solutions of stochastic differential equations. If this is the case, then we may simulate the given stochastic differential equation; thus we are free to propose random mechanisms which are distributed according to the law of the solution of the stochastic differential equation. The natural way is to discretize the stochastic differential equation in time. Then, knowing that the increments of the Brownian motion are independent random normal variables with known variance, a numerical scheme is applied to find approximate values of the sample path step by step at the discretization points. It is well known that we do not need to simulate normal variates if we are only interested in weak solutions. Random variables which are distributed uniformly on the unit interval may serve the same purpose. It is even more surprising that discrete variables may be used, see [4, 6, 10].

In practice, when simulating stochastic differential equations at the computer, pseudo–random numbers are used instead of random ones. Pseudo–random numbers are deterministic and numerical methods based on the use of such deterministic analogs are often called quasi–Monte Carlo methods. The mathematical theory of quasi–Monte Carlo simulation is well developed for numerical integration. The heart of this theory is the Koksma–Hlawka–inequality, which provides an error estimate for a quadrature formula based on quasi–random numbers in terms of the discrepancy of this sequence of numbers. Especially it follows that those sequences fit best which are of low discrepancy. The error estimate for a carefully chosen low–discrepancy sequence is even superior to the one provided by crude Monte Carlo; for a detailed discussion see [11], a more recent exposition is [12].

The purpose of this paper can roughly be expressed in the question: Does this carry over to the simulation of stochastic differential equations? In other words, may we use low–discrepancy sequences instead of random variables, and if this is the case, may we improve the error estimate as this was true for the integration problem? It is well known, that we may proceed from numerical integration to the solution of Fredholm integral equations, see [3].

Below we are going to prove some negative results by providing necessary conditions a deterministic sequence of numbers has to fulfill in order to serve for a quasi–Monte Carlo simulation. Especially we prove that low–discrepancy points must not be used. This will also be transparent by simulation results given below.

Moreover, we show that restricting uniformly (mod 1) distributed sequences to those which are completely uniformly distributed (CUD), yields sequences which may serve for quasi–Monte Carlo simulation.

The next section formalizes the previous considerations.

2. Problem formulation

The mathematical analysis of the quasi–Monte Carlo simulation will be carried out for the following family of stochastic differential equations. Let $\{W_t, \quad 0 \leq t \leq T\}$ be a standard Brownian motion and consider the (one–dimensional) stochastic

differential equation

$$(2.1) \qquad dX_t = -\alpha X_t dt + \beta(t) dW_t, \quad X_0 := 0,$$

for parameters $\alpha > 0$ and functions $\beta \colon [0, T] \to \mathbb{R}$. Thus equations (2.1) are obtained from (1.1) by specifying

$$a(s, x) := \alpha x, \quad x \in \mathbb{R} \quad \text{and} \quad \sigma(s, x) := \beta(s), \quad 0 \le s \le t.$$

These stochastic differential equations are of Langevin type, the solution of which can be given analytically; namely, see e.g. [5], we have

$$(2.2) \qquad X_t = \int_0^t e^{-\alpha(t-s)} \beta(s) dW_s, \quad t > 0.$$

We aim at approximating the second moment at the final time point T, i.e.,

$$\left(\mathbf{E}|X_T|^2\right)^{1/2} = \left(\int_0^T e^{-2\alpha(T-s)} |\beta(s)|^2 ds\right)^{1/2}.$$

We assume $T = 1$ for simplicity. The exact solution is a function, say $S(\alpha, \beta)$, of the data (α, β); see equation (2.2), that is

$$S(\alpha, \beta) := \left(\int_0^1 e^{-2\alpha(1-s)} |\beta(s)|^2 ds\right)^{1/2}.$$

The description of the mathematical problem will be complete after specifying the possible input. The problem elements under consideration, i.e., the class of equations (2.1) is described by restrictions on the values of α and functions $\beta \colon [0, 1] \to \mathbb{R}$. For $a, b, L > 0$, we let

$$X(a, b, L) := \left\{ (\alpha, \beta), \quad 0 < \alpha \le a, \quad \max_{0 \le s \le 1} |\beta(s)| \le b, \ |\beta(s) - \beta(t)| \le L|s - t| \right\},$$

hence the crucial restriction is the Lipschitz continuity of β, the class of possible problem elements. If we equip $X(a, b, L)$ with metric

$$\rho((\alpha, \beta), (\alpha', \beta')) := |\alpha - \alpha'| + \sup_{s \in [0,1]} |\beta(s) - \beta'(s)|,$$

then the set of problem elements admits a countable everywhere dense subset (it is separable).

The mapping $S \colon X(a, b, L) \to \mathbb{R}$ is sometimes called solution operator. It is readily checked that this operator acts continuously on $(X(a, b, L), \rho)$.

Next we specify the admissible numerical schemes to treat stochastic differential equations from type (2.1). To be precise let

$$(2.3) \quad \mathfrak{U}(X(a, b, L), \mathbb{R}) := \{u(\alpha, \beta) = \varphi(\alpha, \beta(t_1), \ldots, \beta(t_n)),$$
$$n \in \mathbb{N}, \ t_1, \ldots, t_n \in [0, 1], \ \varphi \colon \mathbb{R}^{n+1} \to \mathbb{R}\}.$$

The number n of function evaluations supplemented by the evaluation of α gives rise to saying that the method u in the braces in (2.3) has cardinality $\operatorname{card}(u) = n + 1$. Thus we allow any method based on evaluation of β at some time discretization points. The final methods, which shall be Monte Carlo methods are derived from this class $\mathfrak{U}(X(a, b, L), \mathbb{R})$ in Section 3.

Below we shall analyze the Euler scheme with respect to an equidistant time discretization, N steps of step size $1/N$ each, in more detail. The definition will

be given for any realization of the Brownian motion in (2.1). That is, we obtain a recursive description of the approximation $\bar{Y}_n, \quad n = 0, \ldots, N-1$ as

$$(2.4) \qquad \bar{Y}_{n+1} = \bar{Y}_n - \frac{\alpha}{N}\bar{Y}_n + \beta(\frac{n}{N})\frac{1}{\sqrt{N}}\gamma_n, \quad n = 0, \ldots, N-1, \quad \bar{Y}_0 = 0,$$

where $\frac{1}{\sqrt{N}}\gamma_0, \ldots, \frac{1}{\sqrt{N}}\gamma_{N-1}$ are the increments $\frac{1}{\sqrt{N}}\gamma_k := W_{\frac{k+1}{N}} - W_{\frac{k}{N}}$ of the Brownian motion $\{W_s, \quad 0 \le s \le 1\}$. For the purpose of simulation we observe that the random variables $\gamma_k, \quad k = 0, \ldots, N-1$, are independent with common standard normal distribution. It is also well known, that these random variables may be replaced by any other random variables having mean 0 and variance 1, without spoiling the asymptotic behavior of the approximation. So we may replace $\gamma_0, \ldots, \gamma_{N-1}$ by $\sqrt{3}(1-2\xi_0), \ldots, \sqrt{3}(1-2\xi_{N-1})$, with independent uniformly distributed on $[0,1]$ random variables $\xi_0, \ldots, \xi_{N-1}$. But, since we aim at studying quasi–Monte Carlo simulation we will replace the $\xi_0, \ldots, \xi_{N-1}$ by (deterministically from $[0,1)$ chosen) points $\omega_0, \ldots, \omega_{N-1}$ and arrive at

$$(2.5) \qquad Y_{n+1} = Y_n - \frac{\alpha}{N}Y_n + \beta(\frac{n}{N})\frac{\sqrt{3}}{\sqrt{N}}(1 - 2\omega_n), \quad n = 0, \ldots, N-1, \quad Y_0 = 0.$$

For each realization of the $\gamma_0, \ldots, \gamma_{N-1}$ in equation (2.4) (or the respective choice of points $\omega_0, \ldots, \omega_{N-1}$ in equation (2.5)) the corresponding Y_N and $\bar{Y}_N$ are hopefully distributed close to the actual distribution of X_1.

In the simple situation of the model equations of the form (2.1) we can provide an explicit representation of the final approximants $\bar{Y}_N$ and Y_N, respectively. We have

$$(2.6) \qquad \bar{Y}_N = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \left(1 - \frac{\alpha}{N}\right)^{N-1-k} \beta(\frac{k}{N})\gamma_k$$

and correspondingly

$$Y_N = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \left(1 - \frac{\alpha}{N}\right)^{N-1-k} \beta(\frac{k}{N})\sqrt{3}(1 - 2\omega_k).$$

For the study of the quasi–Monte Carlo simulation we assume that we are given an infinite sequence of points $(\omega_i)_{i=0}^{\infty} \subset [0,1)$ from the very beginning and that we sequentially take elements whenever they are needed. Thus we are looking for multi–purpose sequences, working for any number of time steps chosen. The analysis would be different, if we would fit the sequence to a prescribed time discretization.

To complete the numerical scheme we will simulate the second moment by the sample mean of M independent copies of X_1, in our case,

$$(2.7) \qquad \mathbf{E}|X_1|^2 \sim \frac{1}{M} \sum_{j=1}^{M} |\bar{Y}_N^j|^2,$$

or for the quasi–Monte Carlo simulation

$$\mathbf{E}|X_1|^2 \sim \frac{1}{M} \sum_{j=1}^{M} |Y_N^j|^2,$$

respectively, where the j–th copy of Y_N shall be obtained by replacing ω_k by $\omega_k^{(j)}$, $k = 0, \ldots, N - 1$, which are obtained from the original sequence $(\omega_i)_{i=0}^{\infty}$ by letting $\omega_k^{(j)} := \omega_{(j-1)N+k}$, which means

$$(2.8) \quad \frac{1}{M} \sum_{j=1}^{M} |Y_N^j|^2 = \frac{1}{M} \sum_{j=1}^{M} |\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \left(1 - \frac{\alpha}{N}\right)^{N-1-k} \beta(\frac{k}{N}) \sqrt{3}(1 - 2\omega_k^{(j)})|^2$$

represents the mean of M trajectories from the Euler scheme $Y_N := Y_N(\alpha, \beta)$ with N equidistant time points in the quasi–Monte Carlo simulation.

This is made precise below.

3. Monte Carlo simulation

Most texts on the Monte Carlo simulation to obtain weak approximations to equations (1.1) provide schemes of different order of weak approximation. In general there is a lack of optimality issues, i.e., there are no indications on lower estimates. In the present context, that is, dealing with such simple equations as (2.1), the optimal order on Lipschitz continuous input can be given. To do this we have to describe the class of possible input functions as well as admissible numerical schemes.

Since possible input was described above we shall turn to the admissible stochastic numerical schemes, which in fact are Monte Carlo methods, each realization of which shall belong to class $\mathfrak{U}(X(a, b, L), \mathbb{R})$ introduced by (2.3). Thus such a Monte Carlo method shall be a random element taking values in $\mathfrak{U}(X(a, b, L), \mathbb{R})$. It is important to note that we have only deterministic partial information on β which comes from the deterministic time discretization. Moreover we assume that for any given data (α, β) the mapping $\omega \to u_\omega(\alpha, \beta)$ is measurable. To complete the description of the Monte Carlo scheme we note that any random mechanism driving $\omega \to u_\omega(\alpha, \beta)$ is allowed. This will be given by a probability space $[\Omega, \mathcal{F}, P]$.

To be precise, a Monte Carlo method $\mathcal{P}$ of cardinality n is given as a triple $\mathcal{P} = ([\Omega, \mathcal{F}, P], u, n)$ where

1. $[\Omega, \mathcal{F}, P]$ is a probability space.
2. For any data $(\alpha, \beta) \in X(a, b, L)$ the mapping $\omega \to u_\omega(\alpha, \beta) \in \mathfrak{U}(X(a, b, L), \mathbb{R})$ is a real random variable where we assume that any evaluation of the data is deterministic.
3. Each realization u_ω has cardinality at most n.

Remark 3.1. The introduction of Monte Carlo methods as given above is presented in [9]. It turns out to be rich enough to include all known methods, not only within the treatment of stochastic differential equations, but in a fairly general context. However, within the present context information of the data is assumed to be deterministic, which corresponds to a deterministic choice of time discretization. This is customary for treating stochastic differential equations.

Let us also mention that we allow fixed cardinality only. In a more general framework varying cardinality would also be of interest. We shall not turn to that problem.

The (global) error of such a Monte Carlo method $\mathcal{P}$ (for the given solution operator S) shall be measured by

$$e(S, \mathcal{P}) := \sup_{(\alpha, \beta) \in X(a,b,L)} \left(\int_{\Omega} |S(\alpha, \beta) - \mathrm{u}_{\omega}(\alpha, \beta)|^2 dP(\omega) \right)^{1/2}.$$

This is the usual mean squared error customary for Monte Carlo methods. Our main goal is to minimize this Monte Carlo error with respect to a choice of Monte Carlo methods $\mathcal{P}$ with prescribed cardinality $\mathrm{card}(\mathcal{P})$ at most $n-1$, thus we study

$$e_n^{mc}(S, \mathfrak{U}) := \inf \{ e(S, \mathcal{P}), \quad \mathrm{card}(\mathcal{P}) \le n-1 \}.$$

We are now in a position to provide the optimal rate of convergence within the present context of equations (2.1).

Theorem 3.2. *There are constants $C > c > 0$, such that*

$$c\frac{1}{n} \le e_n^{mc}(S, \mathfrak{U}) \le C\frac{1}{n}.$$

Proof. The upper bound shall be obtained by using the Euler scheme with N equidistant time points as provided by the right–hand side in (2.7), which is easily seen to be a Monte Carlo method, obtained from a basic one, say $\bar{\mathcal{P}}$ by independent sampling. This Monte Carlo method $\bar{\mathcal{P}}$, provided in equation (2.4), can be given by $\bar{\mathcal{P}} = ([\Omega, \mathcal{F}, P], \mathrm{u}, N)$ as

1. $\Omega = \mathbb{R}^N$ equipped with the product σ–algebra $\mathcal{F}^N$ and the corresponding product probability $\mathcal{N}(0,1)^N$, describing the choice of N independent standard normal variates.
2. For given (α, β) the mapping u is given from equation (2.6).
3. The equidistant design consists of N points $t_k := \frac{k}{N}$, $k = 0, \ldots, N-1$.

Using the notation from the previous Section 2 the error of $\mathcal{P}$ can be decomposed into the bias and the statistical error to arrive at

$$(3.1) \qquad \left(\mathbf{E} \left| \left(\mathbf{E}|X_1|^2 \right)^{1/2} - \left(\frac{1}{M} \sum_{j=1}^{M} |\bar{Y}_N^j|^2 \right)^{1/2} \right|^2 \right)^{1/2}$$

$$\le \left| \left(\mathbf{E}|X_1|^2 \right)^{1/2} - \left(\mathbf{E}|\bar{Y}_N|^2 \right)^{1/2} \right|$$

$$+ \left(\mathbf{E} \left| \left(\mathbf{E}|\bar{Y}_N|^2 \right)^{1/2} - \left(\frac{1}{M} \sum_{j=1}^{M} |\bar{Y}_N^j|^2 \right)^{1/2} \right|^2 \right)^{1/2}.$$

The first summand in (3.1) is controlled by the step size of the Euler scheme and is known to be of the order $1/N$. We are left to show that this bound is uniform over input data from $X(a, b, L)$. First we observe that Y_N also admits a representation as an Ito integral of some step function f_N which is piecewise constant on intervals $[k/N, (k+1)/N)$, taking values $(1 - \frac{\alpha}{N})^{N((N-1)/N-k/N)} \beta(\frac{k}{N})$ there. Let us for a moment denote the actual integrand in equation (2.2) by

$$f(s) = e^{-\alpha(1-s)} \beta(s).$$

With this notation an application of the triangle inequality provides

$$\left| \left(\mathbf{E}|X_1|^2 \right)^{1/2} - \left(\mathbf{E}|\bar{Y}_N|^2 \right)^{1/2} \right| \le \left(\mathbf{E} \left| \int_0^1 (f(s) - f_N(s)) dW_s \right|^2 \right)^{1/2}.$$

Employing the basic isometry for Ito integrals we obtain

$$\big| \left(\mathbf{E}|X_1|^2\right)^{1/2} - \left(\mathbf{E}|\bar{Y}_N|^2\right)^{1/2} \big|$$

$$\leq \max_{k} \max_{s \in [k/N, (k+1)/N]} |f(s) - f_N(s)|$$

(3.2)

$$\leq \max_{k} \max_{s \in [k/N, (k+1)/N]} |e^{-\alpha(1-s)}\beta(s) - e^{-\alpha(1-s)}\beta(k/N)|$$

(3.3)

$$+ \max_{k} \max_{s \in [k/N, (k+1)/N]} |e^{-\alpha(1-s)}\beta(k/N) - \left(1 - \frac{\alpha}{N}\right)^{N\left(\frac{N-1}{N} - \frac{k}{N}\right)} \beta(k/N)|.$$

From this standard manipulation we obtain an upper estimate of the form

$$\big| \left(\mathbf{E}|X_1|^2\right)^{1/2} - \left(\mathbf{E}|\bar{Y}_N|^2\right)^{1/2} \big| \leq \frac{L}{N} + C\frac{ba}{N}$$

for some universal constant $C < \infty$ and $N \geq N_0$, where the first summand bounds (3.2) and the second the expression in (3.3).

The second summand in (3.1), which represents the Monte Carlo error of the crude sampling is bounded by the dispersion of the integrand, which is $\left(\mathbf{E}|\bar{Y}_N|^2\right)^{1/2}$, divided by $\sqrt{M}$. This implies

(3.4) $$\left(\mathbf{E}\big| \left(\mathbf{E}|X_1|^2\right)^{1/2} - \left(\frac{1}{M}\sum_{j=1}^{M}|\bar{Y}_N^j|^2\right)^{1/2} \big|^2\right)^{1/2} \leq C(a,b,L)\left(\frac{1}{N} + \frac{1}{\sqrt{M}}\right)$$

for $N \geq N_0$.

Let us mention explicitly that the information on β does not change during the process of sampling, it is always at the time points previously chosen. The choice $N := n$ and $M := n^2$ in estimate (3.4) provides the required upper estimate in Theorem 3.2.

We turn to the lower bound. It is easily obtained for equations of the type (2.1), since there is an explicit representation of the solution in terms of $(\alpha, \beta) \in X(a, b, L)$. To this end let $\mathcal{P}$ be any Monte Carlo method of cardinality less than n. Let $(t_0, \ldots, t_{n-1})$ be the time discretization used by $\mathcal{P}$. We denote by $\Delta_k := [t_k, t_{k+1})$, $k = 0, \ldots, n-1$ (where we put $t_n = 1$), and let $d_k := t_{k+1} - t_k$, $k = 0, \ldots, n-1$. We assume without loss of generality that $\min_{k=0,\ldots,n-1} d_k > 0$.

We start our construction with some basic function $\varphi \colon [0, 1] \to \mathbb{R}^+$, given by

$$\varphi(x) := \begin{cases} \frac{x}{2L}, & \text{if } 0 \leq x < \frac{1}{2}, \\ \frac{1-x}{2L}, & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

Finally we put

$$\bar{\beta}(s) := \min\{1, 2bL\} d_k \varphi(\frac{s - t_k}{d_k}), \quad s \in \Delta_k, \quad k = 0, \ldots, n-1.$$

The integral of this function can easily be estimated from below via

$$\int_0^1 \bar{\beta}(s)ds \geq \frac{\min\{1, 2bL\}}{8L} \sum_{k=0}^{n-1} d_k^2 \geq \frac{\min\{1, 2bL\}}{8L} \frac{1}{n}.$$

Moreover it is seen that $(\frac{a}{2}, \bar{\beta}) \in X(a, b, L)$ and

$$S(\frac{a}{2}, \bar{\beta}) \geq e^{-a/2} \int_0^1 \bar{\beta}(s)ds \geq C(a, b, L)\frac{1}{n},$$

for some constant $C(a, b, L)$. Also the pair $(\frac{a}{2}, 0)$ belongs to $X(a, b, L)$, where 0 denotes the zero function. The main observation is that the method $\mathcal{P}$ cannot distinguish between both input data, they share the same information. This yields

$e(S, \mathcal{P})$

$$\geq \max \left\{ \left(\int |S(\frac{a}{2}, 0) - u_\omega(\frac{a}{2}, 0)|^2 dP(\omega) \right)^{1/2}, \left(\int |S(\frac{a}{2}, \bar{\beta}) - u_\omega(\frac{a}{2}, \bar{\beta})|^2 dP(\omega) \right)^{1/2} \right\}$$

$$\geq \max \left\{ \left(\int |u_\omega(\frac{a}{2}, 0)|^2 dP(\omega) \right)^{1/2}, \left(\int |S(\frac{a}{2}, \bar{\beta}) - u_\omega(\frac{a}{2}, 0)|^2 dP(\omega) \right)^{1/2} \right\}$$

$$\geq \frac{1}{2}S(\frac{a}{2}, \bar{\beta}) \geq \frac{1}{2}C(a, b, L)\frac{1}{n}$$

completing the proof of the lower bound and of the theorem.   □

Remark 3.3. It is worth discussing the above estimate from the complexity point of view. Suppose we choose an Euler scheme with N steps and do the simulation M times. In general, for equations of the type (1.1) this requires $n := N \cdot M$ computations of the functions a and σ. If we have for the general case of equations of type (1.1) an estimate similar to (3.4), then, optimizing the choice of N and M to minimize the error for given values n, we end up with $M \sim N^2$. In other words, the error can be bounded by $Cn^{-1/3}$ provided we are willing to allow n evaluations of the input data a and σ.

It is not known whether this is optimal for some more general class of data than $X(a, b, L)$. Thus the picture from Theorem 3.2 is not a final one. For the present purposes it is however sufficient.

4. QUASI–MONTE CARLO SIMULATION

We leave this brief outline of Monte Carlo simulation and turn to the main problem of the paper. Instead of doing (real) Monte Carlo simulation we switch to quasi–Monte Carlo simulation.

For later use we briefly review the main facts concerning the quasi–Monte Carlo simulation for the approximate computation of integrals, see [8]. The main notion is the discrepancy of a point set in some unit cube. Let s be any positive natural number. For brevity let us denote, given any $a \in (0, 1]^s$, $a = (a_0, \ldots, a_{s-1})$ by $[0, a)$ the cube

$$[0, a) := \{x = (x_0, \ldots x_{s-1}) \in [0, 1]^s, \quad x_i < a_i, \quad i = 0, \ldots, s - 1\}.$$

For any given collection of M points $(x^j)_{j=1}^M$ in $[0, 1)^s$ the quantity

$$D_M^* \left((x^j)_{j=1}^M \right) := \sup_{a \in (0,1]^s} |\frac{\# \{x^j, \quad x^j \in [0, a), \; j = 1, \ldots, M\}}{M} - a_0 \cdots a_{s-1}|$$

denotes the $*-discrepancy$.

Observe that we suppress the dependence on s, since this will be clear from the sequence under consideration.

This concept has far reaching applications. We recall the following

Definition 4.1. An infinite sequence of points is said to be *uniformly distributed* (mod 1), if the sequence of the $*$–discrepancies of the initial segments tends to 0.

Such sequences of points may be useful for the approximate computation of integrals. In fact, if a sequence $\left(x^j\right)_{j=1}^\infty \subset [0,1)^s$ is uniformly distributed (mod 1), then $\frac{1}{M}\sum_{j=1}^M f(x^j)$ tends to $\int_{[0,1]^s} f(x)dx$ for any Riemann integrable function on $[0,1]^s$. We note that the $*$–discrepancy can be seen as the uniform error bound for the integration of step functions $\chi_{[0,a)}$ by a quadrature formula of the type $\frac{1}{M}\sum_{j=1}^M \chi_{[0,a)}(x^j)$. This bound extends to functions of bounded variation (in the sense of Hardy–Krause). This is known as the Koksma–Hlawka inequality, mentioned in the introduction.

Koksma–Hlawka inequality. *For any given collection of M points $\left(x^j\right)_{j=1}^M$ in $[0,1)^s$ and any function $f\colon [0,1]^s \to \mathbb{R}$ of bounded variation, we have*

$$(4.1)\quad |\frac{1}{M}\sum_{j=1}^M f(x^j) - \int_{[0,1]^s} f(x)dx|$$

$$\leq \sum_{p=0}^{s-1}\sideset{}{^*}{\sum}_{1,\ldots,k;p} D_M^*(x^{p+1,\ldots,s-1})V^{(p)}(f(\ldots,1,\ldots,1)).$$

For the complete explanation of the symbols used above we refer to [8, Ch. 2, §5]. There one can find further details and variations of this inequality. Since we shall make use of the two–dimensional case below we shall state it explicitly:

For given M points $\left((\xi^j,\eta^j)\right)_{j=1}^M \subset [0,1)^2$ and integrable function $f\colon [0,1]^2 \to \mathbb{R}$ we have

$$(4.2)\quad |\frac{1}{M}\sum_{j=1}^M f(\xi^j,\eta^j) - \int_0^1\int_0^1 f(x,y)dxdy| \leq D_M^*\left((\xi^j)_{j=1}^M\right) V^{(1)}(f(x,1))$$

$$+ D_M^*\left((\eta^j)_{j=1}^M\right) V^{(1)}(f(1,y)) + D_M^*\left(((\xi^j,\eta^j))_{j=1}^M\right) V^{(2)}(f(x,y)).$$

Here $V^{(1)}$ denotes the (usual) variation of a function on $[0,1]$ (with respect to the variable as indicated). The symbol $V^{(2)}$ denotes the total variation (in the sense of Vitali), see [8, Ch. 2, §5].

First we shall return to the problem of approximating the solution $S(\alpha,\beta)^2$ by expressions (2.8). We propose the following

Definition 4.2. An infinite sequence $(\omega_i)_{i=0}^\infty \subset [0,1)$ is said to be *consistent* (for the solution operator S) if for all $(\alpha,\beta) \in X(a,b,L)$ we have

$$\lim_{N\to\infty}\lim_{M\to\infty} \frac{1}{M}\sum_{j=1}^M |Y_N^j(\alpha,\beta)|^2 \longrightarrow S(\alpha,\beta)^2.$$

(Here Y_N^j are defined as in (2.8).)

Thus the following question seems to be natural: Are uniformly (mod 1) sequences consistent? It turns out that this is not always the case.

But a slightly stronger property than uniform distribution (mod 1), namely being completely uniformly distributed (mod 1) (often abbreviated CUD), provides sequences which are consistent. The original introduction of such sequences has

probably been given by Korobov. The importance of the latter notion has been realized early when simulating Markov chains, see [1].

Let us recall the definition, which in fact is not the usual one. But as has been observed by Korobov in his dissertation, 1953, the present approach which focuses on Monte Carlo sampling is equivalent to the original one.

Definition 4.3. An infinite sequence $(\omega_i)_{i=0}^{\infty} \subset [0,1)$ is called *completely uniformly distributed* (CUD), if for all positive $s \in \mathbb{N}$ the sequence $\left(\omega_0^{(j)}, \ldots, \omega_{s-1}^{(j)}\right)_{j=1}^{\infty}$, which is defined by

$$\omega_k^{(j)} := \omega_{(j-1)s+k}, \quad k = 0, \ldots, s-1, \quad j \in \mathbb{N},$$

is uniformly distributed (mod 1) in the s–dimensional unit cube $[0,1)^s$.

Thus for any s, the choice of subsequent s–tuples yields a sequence of vectors uniformly distributed (mod 1). In the excellent treatment [7, 3.5 B], such sequences are called ∞–distributed. There one can find further discussion on the randomness and other properties. Also, indications on explicit constructions are given. A thorough comparison of various types of uniformly distributed sequences is provided in [2].

For more recent references and discussion we refer to [11, 12]. The following result is an immediate consequence of the definition.

Proposition 4.4. *CUD sequences are consistent for S on $X(a, b, L)$.*

Proof. It is readily checked that $\left(\mathbf{E}|\bar{Y}_N(\alpha, \beta)|^2\right)^{1/2}$ tends to $S(\alpha, \beta)$ if N tends to infinity. Thus it is enough to verify that for every fixed N the convergence

$$(4.3) \qquad \frac{1}{M} \sum_{j=1}^{M} |Y_N^j(\alpha, \beta)|^2 \longrightarrow \mathbf{E}|\bar{Y}_N(\alpha, \beta)|^2$$

for $M \to \infty$ is true. But it can be drawn from the representation in (2.8), that the left–hand side above is the average of the values of the function $f \colon [0,1]^N \to \mathbb{R}$ given by

$$f(x_0, \ldots, x_{N-1}) := |\frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \left(1 - \frac{\alpha}{N}\right)^{N-1-k} \beta(\frac{k}{N})\sqrt{3}(1 - 2x_k)|^2$$

at M points $(\omega_0^{(1)}, \ldots, \omega_{N-1}^{(1)}), \ldots, (\omega_0^{(M)}, \ldots, \omega_{N-1}^{(M)})$. Since the original sequence was supposed to be CUD, the sequence of points $\left(\omega_0^{(j)}, \ldots, \omega_{N-1}^{(j)}\right)_{j=1}^{\infty} \subset [0,1)^N$ is uniformly distributed (mod 1), such that the left–hand side in (4.3) converges as stated. $\qquad \square$

Remark 4.5. Looking at the above arguments we immediately recognize that CUD sequences may be used for quasi–Monte Carlo simulation whenever an explicit numerical scheme is used which is weakly convergent for some class of stochastic differential equations. For the appropriate setup we refer to [6, 9.7]

Since almost all (with respect to the uniform distribution on $[0,1]^{\mathbb{N}}$) sequences are CUD, see e.g. [8, Chapt. 3, Thm. 3.13], we also infer that almost all randomly chosen sequences are consistent for the solution operator S.

Next we turn to the problem of whether there are necessary conditions to be imposed on a sequence in order to be consistent for some class of problems. We start with the following

Proposition 4.6. *There are constants $C = C(a, b, L)$ and $N(a) \in \mathbb{N}$ such that for any infinite sequence $(\omega_i)_{i=0}^\infty \subset [0, 1)$ and for all $N \geq N(a)$, $M \in \mathbb{N}$ and $j = 1, \ldots, M$, we have*

$$|Y_N^j(\alpha, \beta)| \leq C\sqrt{N} \left\{ D_N^* \left(\left(\omega_k^{(j)} \right)_{k=0}^{N-1} \right) + D_N^* \left(\left((\frac{k}{N}, \omega_k^{(j)}) \right)_{k=0}^{N-1} \right) \right\}.$$

Proof. The proof is based on the Koksma–Hlawka inequality. Let $N(a) := 2(\lceil a \rceil + 1)$, such that $\frac{a}{N(a)} \leq \frac{1}{2}$. We fix any j, $j = 1, \ldots, M$ and rewrite $Y_N = Y_N^j$ as

$$Y_N = \sqrt{3N} \frac{1}{N} \sum_{k=0}^{N-1} \left(1 - \frac{\alpha}{N} \right)^{N((N-1)/N - k/N)} \beta(\frac{k}{N})(1 - 2\omega_k^{(j)}).$$

For $N \geq N(a)$ we apply the Koksma–Hlawka inequality to the function

$$f(x, y) := \left(1 - \frac{\alpha}{N} \right)^{N((N-1)/N - x)} \beta(x)(1 - 2y), \quad x, y \in [0, 1),$$

and obtain

(4.4)

$$\left| \frac{1}{N} \sum_{k=0}^{N-1} f(\frac{k}{N}, \omega_k^{(j)}) - \int_0^1 \int_0^1 f(x, y) dx dy \right|$$

$$\leq D_N^* \left((k/N)_{k=0}^{N-1} \right) V^{(1)}(f(x, 1)) + D_N^* \left(\left(\omega_k^{(j)} \right)_{k=0}^{N-1} \right) V^{(1)}(f(1, y))$$

$$+ D_N^* \left(\left((k/N, \omega_k^{(j)}) \right)_{k=0}^{N-1} \right) V^{(2)}(f(x, y))$$

$$\leq 4(L + b) \left\{ D_N^* \left((k/N)_{k=0}^{N-1} \right) + D_N^* \left(\left(\omega_k^{(j)} \right)_{k=0}^{N-1} \right) + D_N^* \left(\left((k/N, \omega_k^{(j)}) \right)_{k=0}^{N-1} \right) \right\}$$

$$\leq 8(L + b) \left\{ D_N^* \left(\left(\omega_k^{(j)} \right)_{k=0}^{N-1} \right) + D_N^* \left(\left((k/N, \omega_k^{(j)}) \right)_{k=0}^{N-1} \right) \right\},$$

where we used the easily established fact that the respective variations of f are all bounded by $2(L + b)/(1 - a/N(a)) \leq 4(L + b)$ to derive (4.4) and the known fact that the $*$–discrepancy of $0, \frac{1}{N}, \ldots, \frac{N-1}{N}$ is minimal and equal to $\frac{1}{N}$. Taking into account that the double integral of f evaluates to 0, the proof can be completed. $\square$

To proceed we need a result which relates the discrepancy of a certain point set in the unit square to the marginal coordinate sequence. The following lemma is a variation of an argument used in the proof of Theorem 2.2 in [8].

Lemma 4.7. *There is a natural number N_0, such that for all $N \geq N_0$ and point sets $(\omega_i)_{i=0}^{N-1} \subset [0, 1)$, there is an m, $1 \leq m \leq N$, for which*

$$mD_m^*((\omega_i)_{i=0}^{N-1}) \geq \frac{N}{4} D_N^*((i/N, \omega_i)_{i=0}^{N-1}).$$

Proof. For completeness we repeat the argument from [8] briefly. Since the sequence $ND_N^*((i/N, \omega_i)_{i=0}^{N-1})$ tends to infinity, see [8, Thm. 2.1], we can find N_0 for which

$ND_N^*((i/N,\omega_i)_{i=0}^{N-1}) \geq 4, \quad N \geq N_0.$ Since moreover the $*$–discrepancy is the supremum over cuboids, we can find $x, y \in (0, 1]$ for which

$$|\frac{\#\left\{(i/N,\omega_i), \quad i/N < x, \ \omega_i < y, \ i = 0, \dots, N-1\right\}}{N} - xy| \geq \frac{1}{2} D_N^*((i/N,\omega_i)_{i=0}^{N-1}).$$

Let for $N \geq N_0$ the number $m - 1$ be the greatest natural number smaller than xN. This yields

$$\#\left\{(i/N,\omega_i), \quad i/N < x, \ \omega_i < y, \ i = 0, \dots, N-1\right\}$$
$$= \#\left\{\omega_i, \quad \omega_i < y, \ i = 0, \dots, m-1\right\}.$$

Now we can conclude

$$|\frac{\#\left\{\omega_i, \quad \omega_i < y, \ i = 0, \dots, m-1\right\}}{m} - y|$$
$$= |\frac{\#\left\{(i/N,\omega_i), \quad i/N < x, \ \omega_i < y, \ i = 0, \dots, N-1\right\}}{N}\frac{N}{m} - y|$$
$$\geq |\frac{\#\left\{(i/N,\omega_i), \quad i/N < x, \ \omega_i < y, \ i = 0, \dots, N-1\right\}}{N}\frac{N}{m} - \frac{N}{m}xy|$$
$$- |\frac{N}{m}xy - y|$$
$$\geq \frac{N}{2m}D_N^*((i/N,\omega_i)_{i=0}^{N-1}) - y\frac{N}{m}|x - \frac{m}{N}|$$

(4.5) $\qquad \geq \frac{N}{2m}D_N^*((i/N,\omega_i)_{i=0}^{N-1}) - \frac{1}{m},$

where we used the definition of m to derive (4.5). Employing the definition of the $*$–discrepancy and multiplying by m, we obtain

(4.6) $\qquad\qquad mD_m^*((\omega_i)_{i=0}^{N-1}) \geq \frac{N}{2}D_N^*((i/N,\omega_i)_{i=0}^{N-1}) - 1.$

By our choice of N_0 estimate (4.6) yields

$$mD_m^*((\omega_i)_{i=0}^{N-1}) \geq \frac{N}{4}D_N^*((i/N,\omega_i)_{i=0}^{N-1}),$$

completing the proof of the lemma. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

In order to formulate the necessary condition we shall have to use a slightly stronger notion than discrepancy, which is related to the concept of well distributed sequences. Here we start with an infinite sequence $(\omega_i)_{i=0}^{\infty} \subset [0, 1)$ and put

(4.7) $\quad \tilde{D}_N\left((\omega_i)_{i=0}^{\infty}\right) := \sup_{k \in \mathbb{N}} \sup_{0 < x \leq 1} |\frac{\#\left\{\omega_{k+i}, \quad \omega_{k+i} < x, \ i = 0, \dots, N-1\right\}}{N} - x|.$

Definition 4.8. A sequence $(\omega_i)_{i=0}^{\infty} \subset [0, 1)$ is said to be *well distributed* (mod 1) if $\tilde{D}_N((\omega_i)_{i=0}^{\infty})$ tends to 0.

Since we are not aware of any name, we shall henceforth call $\tilde{D}_N\left((\omega_i)_{i=0}^{\infty}\right)$ the *measure of well distribution.*

It is known that there are sequences uniform (mod 1) which are not well distributed, see again [8, Ch. 1, §5]. The switch from uniformly (mod 1) distributed sequences to well distributed ones is plausible, since in the asymptotic analysis longer and longer segments from the original sequence may be disregarded. Now we turn to the main result of this section. It provides a necessary condition on sequences

to allow consistency for the problem S on $X(a, b, L)$. Recall that $C = C(a, b, L)$ denotes the constant from Proposition 4.6. We have

Theorem 4.9. *For any sequence $(\omega_i)_{i=0}^{\infty} \subset [0, 1)$ which is consistent for S on $X(a, b, L)$, we have*

$$(4.8) \qquad \limsup_{N \to \infty} \sqrt{N} \tilde{D}_N((\omega_i)_{i=0}^{\infty}) \geq \frac{b}{10C}.$$

Remark 4.10. The above lower bound can roughly be reformulated that $\tilde{D}_N$ must not decay too fast to ensure consistency. Remember that the minimal discrepancy of sequences in $[0, 1)$ is of the order $\frac{\log(N)}{N}$, see e.g. [8, Thm. 2.1], which is much faster than Theorem 4.9 permits. Let us emphasize that the necessary condition provided in Theorem 4.9 is also necessary for any class of input data containing some $X(a, b, L)$. This means that low-discrepancy sequences are not suited for the quasi–Monte Carlo simulation of stochastic differential equations. This is also supported by the computational results reported in Section 5.

Proof of Theorem 4.9. We first observe that

$$\max_{(\alpha, \beta) \in X(a, b, L)} |S(\alpha, \beta)| = |S(0, b)| \geq b.$$

Thus if $(\omega_i)_{i=0}^{\infty} \subset [0, 1)$ is any consistent sequence, then for $N \geq N_1 \geq N(a)$, where $N(a)$ is from Proposition 4.6, there is $M(N) \in \mathbb{N}$ such that we have

$$\frac{1}{M} \sum_{j=1}^{M} |Y_N^j(0, b)|^2 \geq \frac{b^2}{4}$$

for all $M \geq M(N)$. In view of the estimate provided in Proposition 4.6, we conclude that for $N \geq N_1$ and $M \geq M(N)$,

$$(4.9) \qquad \sqrt{N} \max_{j=1,\ldots,M} \left\{ D_N^* \left(\left(\omega_k^{(j)} \right)_{k=0}^{N-1} \right) + D_N^* \left(\left((k/N, \omega_k^{(j)}) \right)_{k=0}^{N-1} \right) \right\} \geq \frac{b}{2C}.$$

Now suppose to the contrary of (4.8) that, starting from some $N \geq N_2$ we have $\sqrt{N} \tilde{D}_N((\omega_i)_{i=0}^{\infty}) < \frac{b}{10C}$. The estimate (4.9) then implies for $N \geq \max\{N_1, N_2\}$ and $M \geq M(N)$ the inequality

$$\sqrt{N} \max_{j=1,\ldots,M} D_N^* \left(\left((k/N, \omega_k^{(j)}) \right)_{k=0}^{N-1} \right) \geq \frac{2b}{5C}.$$

Let for any $N \geq \max\{N_1, N_2\}$ the number $j(N)$ be such that the above maximum is attained. An application of Lemma 4.7 yields for all $N \geq \max\{N_0, N_1, N_2\}$ an $m \leq N$ for which

$$\sqrt{m} D_m^* \left(\left(\omega_k^{(j(N))} \right)_{k=0}^{N-1} \right) \geq \frac{1}{4} \sqrt{\frac{N}{m}} \sqrt{N} D_N^* \left(\left(k/N, \omega_k^{(j(N))} \right)_{k=0}^{N-1} \right)$$

$$\geq \frac{b}{10C} \sqrt{\frac{N}{m}}.$$

In terms of the original sequence $(\omega_i)_{i=0}^{\infty}$ we provided for any $N \geq \max\{N_0, N_1, N_2\}$ an $m(N) \leq N$ for which

$$(4.10) \qquad \sqrt{m(N)} \tilde{D}_{m(N)}((\omega_i)_{i=0}^{\infty}) \geq \frac{b}{10C} \sqrt{\frac{N}{m(N)}}.$$

The sequence $(m(N))_{N \in \mathbb{N}}$ cannot be bounded, since then the left–hand side in (4.10) would be bounded, while the right–hand side would tend to infinity. Thus we may extract an infinite sequence of numbers m_l, $l = 1, 2, \ldots$, converging to ∞, for which

$$\sqrt{m_l} \tilde{D}_{m_l}((\omega_i)_{i=0}^{\infty}) \geq \frac{b}{10C},$$

contradicting our initial assumption. This completes the proof of the theorem. $\square$

The theorem just proven has an immediate consequence on the measure of well distribution of CUD sequences.

Corollary 4.11. *There is a constant $C > 0$ such that for every CUD sequence $(\omega_i)_{i=0}^{\infty}$ we have*

$$\limsup_{N \to \infty} \sqrt{N} \tilde{D}_N((\omega_i)_{i=0}^{\infty}) \geq C.$$

But as can be seen right from the definition of CUD sequences, we have the stronger relation $\tilde{D}_N((\omega_i)_{i=0}^{\infty}) = 1$ for any such sequence $(\omega_i)_{i=0}^{\infty}$, since given N and letting $s = N$, the sequence must contain strings $(\omega_k, \omega_{k+1}, \ldots, \omega_{k+N-1})$ arbitrarily close to $(0, 0, \ldots, 0)$ for arbitrarily large k. We thank the referee for pointing out this simple argument to the authors.

It would indeed be interesting to have further results on the discrepancy of sequences which enjoy stronger properties than being uniformly distributed (mod 1). A rough explanation of the phenomenon presented above is that the properties of being independent and having low discrepancy are contradictory. This can be made more precise by introducing a corresponding measure of independence, related to the empirical coefficients of correlation as used in statistics.

5. NUMERICAL EXAMPLES

The theoretical results from the previous section are accomplished by some typical situations as met in computer simulations. We examine the most prominent examples of low-discrepancy sequences, the Kronecker–Weyl–sequence and the v. d. Corput sequence. It will be transparent, that the computational results will tend to 0 rather than approximating the exact solution. This could also be forecasted from Proposition 4.6.

Finally we exhibit the action of the pseudo–random–number generator as implemented in the Turbo Pascal compiler.

Example 5.1. The Kronecker–Weyl–sequence. The construction of the Kronecker–Weyl–sequence is sometimes called *Diophantine approximation* [8, Ch. 2, §3]. For a given real number $\zeta > 0$ we let

$$\omega_i := \{i\zeta\}, \quad i \in \mathbb{N},$$

where $\{x\} := x - \lfloor x \rfloor$ denotes the fractional part of any real number x. It is due to H. Weyl, that the resultant sequence $(\omega_i)_{i=0}^{\infty}$ is uniformly distributed (mod 1) (and also well distributed) if and only if ζ was irrational. The discrepancy and measure of well distribution depend on number theoretic properties of ζ. Especially, if ζ is a quadratic irrationality, hence has bounded partial quotients, then we have

$$D_N^*((\omega_i)_{i=0}^{N-1}) \leq \tilde{D}_N((\omega_i)_{i=0}^{\infty}) \leq C \frac{\log(N)}{N}$$

for some constant $C < \infty$. The fact that $\tilde{D}_N((\omega_i)_{i=0}^{\infty})$ has convergence no worse than the $*$–discrepancy is easiest seen from the Erdös–Turan inequality, see [8, Thm. 2.5]. Consequently, for $\zeta = \sqrt{2}$ the Kronecker–Weyl–sequence must fail for quasi–Monte Carlo simulation. This is demonstrated in Table 1.

TABLE 1. Numerical results obtained with the Kronecker-Weyl-sequence for $\alpha = \frac{1}{2}$ and different functions β

	$\beta(t) \equiv 1$	$\beta(t) = \mathrm{e}^{-t}$	$\beta(t) = t$	$\beta(t) = t^2$
$\mathbf{E}|X_1|^2$	0.6321205	0.2325441	0.2642411	0.1708934
approximate value $(N = 50, M = N^2)$	0.0412733	0.0154923	0.0208595	0.0180976
approximate value $(N = 100, M = N^2)$	0.0166281	0.0061776	0.0117100	0.0106130

Example 5.2. The v. d. Corput sequence. The following construction was first suggested by v. d. Corput. For any prime number $p \in \mathbb{N}$ let

$$\omega_i := \varphi_p(i), \quad i \in \mathbb{N}.$$

Here φ_p denotes the radical inverse function which is defined as follows. We put $\varphi_p(0) := 0$ and for $i \geq 1$ with a p–adic expansion $i = \sum_{j=0}^{s} a_j p^j$ we let $\varphi_p(i) := \sum_{j=0}^{s} a_j p^{-j-1}$, see also [8, Ch. 2, §3]. It is clear that the elements constructed this way belong to the unit interval.

It was proven by v. d. Corput, that we have

$$D_N^*((\omega_i)_{i=0}^{N-1}) \leq C \frac{\log(N)}{N}$$

for some constant $C < \infty$. A careful inspection of the proof of Theorem 3.5 in [8] allows to extend this estimate to the measure of well distribution. Thus the v. d. Corput sequence provides a further example of a low-discrepancy sequence which additionally has a low measure of well distribution, hence it must fail for quasi–Monte Carlo simulation. For $p = 2$ this is also demonstrated in Table 2.

TABLE 2. Numerical results obtained with the van der Corput sequence for $\alpha = \frac{1}{2}$ and different functions β

	$\beta(t) \equiv 1$	$\beta(t) = \mathrm{e}^{-t}$	$\beta(t) = t$	$\beta(t) = t^2$
$\mathbf{E}|X_1|^2$	0.6321205	0.2325441	0.2642411	0.1708934
approximate value $(N = 50, M = N^2)$	0.0112433	0.0042016	0.0100168	0.0087253
approximate value $(N = 100, M = N^2)$	0.0069880	0.0026074	0.0056135	0.0050400

Example 5.3. The prn–generator from Turbo Pascal. In the following table we establish the results with quasi–Monte Carlo simulation using the implemented pseudo–random number generator from Turbo Pascal. Table 3 indicates that this is a possible choice of point set, at least for moderate M and N.

TABLE 3. Numerical results obtained with the prn-generator from Turbo Pascal for $\alpha = \frac{1}{2}$ and different functions β

	$\beta(t) \equiv 1$	$\beta(t) = \mathrm{e}^{-t}$	$\beta(t) = t$	$\beta(t) = t^2$
$\mathbf{E}|X_1|^2$	0.6321205	0.2325441	0.2642411	0.1708934
approximate value $(N = 50, M = N^2)$	0.6203815	0.2307061	0.2614001	0.1642161
approximate value $(N = 100, M = N^2)$	0.6293551	0.2316389	0.2614481	0.1662125

Seemingly, the action of this prn–generator is correct. We are not aware of any theoretical result supporting this behavior. Since this prn–generator is a linear congruential one, it generates only a finite number of points, such that the asymptotic analysis we carried out does not apply.

6. CONCLUDING REMARKS

The above theoretical analysis as well as the computational results indicate that the program designer should be careful in using pseudo–random numbers for the simulation of stochastic differential equations. The (classical) analysis for the integration problem, based on the Koksma–Hlawka inequality, does not carry over to schemes as used for the simulation of stochastic differential equations. It would be interesting to provide also sufficient conditions, which can easily be checked, to evaluate given point sets with respect to their use in quasi–Monte Carlo simulation. This is especially important, since in general the exact solution of a given stochastic differential equation is not known, such that we need some a priori information on the validity of the computer simulations.

Moreover, we carried out a first analysis of the Monte Carlo simulation from the complexity theoretic point of view. Since by now there is a large zoo of available numerical schemes to solve stochastic differential equations (in the strong as well in the weak sense), it would be important to be able to decide which to choose, in dependence of the given class of input data, hence smoothness properties.

REFERENCES

[1] N. N. Chentsov. Pseudo–random numbers for modeling Markov chains. *Zh. Vychisl. Mat. i Mat. Fiz.*, 7:632 – 643, 1967.

[2] J. N. Franklin. Deterministic simulation of random processes. *Math. of Computation*, 17:28 – 59, 1963. MR **26:**7125

[3] E. Hlawka. Lösung von Integralgleichungen mittels zahlentheoretischer Methoden. *Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II*, 171:103 – 123, 1962. MR **27:**548

[4] N. Hofmann. Beiträge zur schwachen Approximation stochastischer Differentialgleichungen. Dissertation, HU Berlin, 1995.

[5] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113 of *Graduate Texts in Math.*, Springer, New York, Berlin, Heidelberg, 1988. MR **89c:**60096

[6] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23 of *Applications of Math.*, Springer, Berlin, Heidelberg, New York, 1992. MR **94b:**60069

[7] D. E. Knuth. *The Art of Computer Programming, Vol. 2/ Seminumerical Algorithms.* Addison–Wesley Publ. Co., Reading, Mass., 1969. MR **44:**3531

[8] L. Kuipers and H. Niederreiter. *Uniform Distribution of Sequences*. Wiley & Sons, New York, London, Sydney, Toronto, 1974. MR **54:**7415

[9] P. Mathé. Approximation theory of Monte Carlo methods. Habilitation thesis, 1994.

[10] G. N. Milstein. *Numerical Integration of Stochastic Differential Equations*, volume 313 of *Mathematics and its Appl.*, Kluwer Acad. Publ., Dordrecht, Boston, London, 1995. MR **96e:**65003

[11] H. Niederreiter. Quasi–Monte Carlo methods and pseudo–random numbers. *Bull. Amer. Math. Soc.*, 84(6):957 − 1041, 1978. MR **80d:**65016

[12] H. Niederreiter. *Random Number Generation and Quasi–Monte Carlo Methods*, volume 63 of *CBMS–NSF Region. Conf. Series in Appl. Math.* SIAM, Philadelphia, 1992. MR **93h:**65008

MATHEMATISCHES INSTITUT, UNIVERSITÄT ERLANGEN–NÜRNBERG, BISMARCKSTR. 1 1/2, D– 91054 ERLANGEN, GERMANY
E-mail address: `hofmann@mi.uni-erlangen.de`

WEIERSTRASS INSTITUTE FOR APPLIED ANALYSIS AND STOCHASTICS, MOHRENSTRASSE 39, D– 10117 BERLIN, GERMANY
E-mail address: `mathe@wias-berlin.de`