JOURNAL OF COMPLEXITY **15**, 342–359 (1999)
Article ID jcom.1999.0509, available online at http://www.idealibrary.com on IDEAL®

Quasi-Monte Carlo Simulation of Diffusion

Christian Lécot[1] and Faysal El Khettabi[2]

Laboratoire de Mathématiques, Université de Savoie, 73376 Le Bourget-du-Lac Cedex, France

Received October 28, 1998

A particle method adapted to the simulation of diffusion problems is presented. Time is discretized into increments of length Δt. During each time step, the particles are allowed to random walk to any point by taking steps sampled from a Gaussian distribution centered at the current particle position with variance related to the

tion problems, the numerical results indicate that an improvement is achieved over standard random walk simulation. © 1999 Academic Press

1. INTRODUCTION

The most famous particle method used to solve diffusive problems is the Monte Carlo method: the diffusive term is modeled by random motions of the particles according to a suitable probability law [Cho73]. However the random choices introduce a large amount of noise leading to computations of high variance. A step towards improving the accuracy of the method is to replace the pseudo-random numbers by low-discrepancy sequences. Quasi-Monte Carlo methods are deterministic versions of Monte Carlo methods. In the last ten years their appeal has broadened significantly since it was found that in certain type of computational problems they systematically outperform Monte Carlo methods [DT97, Nie92, NHLZ98, NS95].

[1] E-mail: Christian.Lecot@univ-savoie.fr.
[2] E-mail: faysal@univ-savoie.fr.

0885-064X/99 $30.00

In the present work, we develop a particle method for solving the diffusion equation

$$\frac{\partial c}{\partial t}(\mathbf{x}, t) = \nabla \cdot (\mathbf{D}\, \nabla c)(\mathbf{x}, t), \qquad \mathbf{x} \in \mathbf{R}^s, \quad t > 0, \tag{1}$$

where the diagonal diffusivity tensor $\mathbf{D}$ has constant coefficients $D_i > 0$. We will focus on the initial-value problem

$$c(\mathbf{x}, 0) = c_0(\mathbf{x}), \qquad \mathbf{x} \in \mathbf{R}^s, \tag{2}$$

where

$$c_0(\mathbf{x}) \geqslant 0 \qquad \text{and} \qquad \int_{\mathbf{R}^s} c_0(\mathbf{x})\, d\mathbf{x} = 1.$$

We discretize time into intervals of length Δt. Discretizing c as a sum of δ-measures, we have

$$c^{(n)}(\mathbf{x}) := \frac{1}{N} \sum_{0 \leqslant j < N} \delta(\mathbf{x} - \mathbf{x}_j^{(n)}),$$

where the $\mathbf{x}_j^{(n)}$, $0 \leqslant j < N$, represent the location of N particles at time $n\, \Delta t$. At every time step, each particle is moved by a quasi-random displacement.

We recall from [Nie92] the fundamental concepts of quasi-random points. Let $s \geqslant 1$ be a fixed dimension. Then $I^s := [0, 1)^s$ is the s-dimensional unit cube and λ_s is the s-dimensional Lebesgue measure. Let Y be a point set consisting of $\mathbf{y}_0, ..., \mathbf{y}_{N-1} \in I^s$. For an arbitrary set $\mathscr{E} \subseteq I^s$ we define the *local discrepancy*

$$D(\mathscr{E}, Y) := \frac{1}{N} \sum_{0 \leqslant j < N} \chi_{\mathscr{E}}(\mathbf{y}_j) - \lambda_s(\mathscr{E}),$$

where $\chi_{\mathscr{E}}$ is the characteristic function of $\mathscr{E}$. If $\mathscr{J}$ is the family of all subintervals of I^s of the form $\prod_{i=1}^{s} [z_i, z_i')$, the *discrepancy* of the point set Y is defined by

$$D_N(Y) := \sup_{J \in \mathscr{J}} |D(J, Y)|.$$

The *star discrepancy* of the point set Y is defined by

$$D_N^*(Y) := \sup_{J \in \mathscr{J}^*} |D(J, Y)|,$$

where $\mathscr{I}^*$ is the family of all subintervals of I^s of the form $\prod_{i=1}^{s} [0, z_i)$. For a sequence Y of points in I^s, we write $D_N(Y)$ for the discrepancy and $D_N^*(Y)$ for the star discrepancy of the first N terms of Y. The most powerful known methods for the construction of low-discrepancy point sets and sequences are based on the theory of (t, m, s)-nets and (t, s)-sequences. For integers $b \geqslant 2$ and $0 \leqslant t \leqslant m$, a (t, m, s)-*net in base b* is a point set Y consisting of b^m points in I^s such that $D(J, Y) = 0$ for every interval $J \subset I^s$ of the form

$$J = \prod_{i=1}^{s} \left[\frac{a_i}{b^{d_i}}, \frac{a_i + 1}{b^{d_i}} \right),$$

with integers $d_i \geqslant 0$ and $0 \leqslant a_i < b^{d_i}$ for $1 \leqslant i \leqslant s$ and of measure $\lambda_s(J) = b^{t-m}$. Let $b \geqslant 2$ and $t \geqslant 0$ be integers. A sequence $\mathbf{y}_0, \mathbf{y}_1, \ldots$ of points in I^s is a (t, s)-*sequence in base b* if, for all integers $n \geqslant 0$ and $m > t$, the points $\mathbf{y}_j$ with $nb^m \leqslant j < (n+1)\, b^m$ form a (t, m, s)-net in base b. For many properties of (t, m, s)-nets and (t, s)-sequences along with their discrepancies, we refer to [Nie87, Nie88].

Quasi-random points, unlike pseudo-random points, are highly correlated with one another by design. It means that they cannot be blindly used in place of pseudo-random points without the risk of introducing biases into the results. For instance, it is shown in [HM97] that low-discrepancy sequences are not suited for the quasi-Monte Carlo simulation of stochastic differential equations. We have found that for simulation of diffusion we can use quasi-random sequences by making use of a technique involving renumbering the particles at each time step [CL97, Lec91, LC98, MC93, Mos97]. With quasi-random points and renumbering, Monte Carlo results can be significantly improved in terms of error versus number of particles.

In Section 2 we introduce the quasi-Monte Carlo method. In Section 3 we state and prove a convergence theorem for this method. In Section 4 we present computational results illustrating our theoretical analysis.

2. THE NUMERICAL METHOD

We consider the pure initial-value problem (1)–(2) and assume that the initial data satisfy

$$c_0 \in L^1(\mathbf{R}^s), \qquad c_0 \geqslant 0, \qquad \text{and} \qquad \int_{\mathbf{R}^s} c_0(\mathbf{x})\, d\mathbf{x} = 1. \tag{3}$$

It is easy to show that

$$\forall t > 0 \qquad \int_{\mathbf{R}^s} c(\mathbf{x}, t)\, d\mathbf{x} = 1, \tag{4}$$

so that the total amount of heat is conserved. We shall use the fundamental solution for the differential operator $\partial/\partial t - \nabla \cdot (\mathbf{D}\, \nabla)$,

$$E(\mathbf{x}, t) := H(t) \prod_{i=1}^{s} \frac{e^{-x_i^2/4D_i t}}{\sqrt{4\pi\, D_i t}},$$

where H is the Heavyside function: $H(t) = 1$ for $t > 0$, $H(t) = 0$ for $t < 0$. Then the solution of (1) satisfies

$$c(\mathbf{x}, t) = \int_{\mathbf{R}^s} E(\mathbf{x} - \mathbf{w}, t - \tau)\, c(\mathbf{w}, \tau)\, d\mathbf{w}, \qquad \mathbf{x} \in \mathbf{R}^s, \quad t > \tau, \tag{5}$$

for $\tau \geqslant 0$.

Let b and $d_1, ..., d_s$ be integers and put $m = d_1 + \cdots + d_s$ and $N = b^m$. We introduce N particles at locations $\mathbf{x}_j^{(0)}$, $0 \leqslant j < N$. We assume that these locations are sampled from the initial distribution c_0. Define the maps P' and P'' by $P'\mathbf{y} := (y_1, ..., y_s)$ and $P''\mathbf{y} := (y_{s+1}, ..., y_{2s})$ for $\mathbf{y} \in I^{2s}$. We write $\mathbf{y}'$ for $P'\mathbf{y}$ and $\mathbf{y}''$ for $P''\mathbf{y}$. Let Y be a sequence of points in I^{2s} satisfying the following properties:

(1) for all integers $n \geqslant 0$, the points $P'\mathbf{y}_j$ with $nb^m \leqslant j < (n+1)\, b^m$ form a $(0, m, s)$-net in base b,

(2) for any $\mathbf{y} \in Y$, we have $P''\mathbf{y} \in (0, 1)^s$. This condition ensures the feasibility of the scheme (see (10) below).

At time $t_n = n\, \Delta t$, we have particles at locations $\mathbf{x}_j^{(n)}$, $0 \leqslant j < N$. From Eq. (5), it follows that an approximate solution to the diffusion equation at time t_{n+1} is defined by

$$\gamma^{(n+1)}(\mathbf{x}) := \frac{1}{N} \sum_{0 \leqslant j < N} E(\mathbf{x} - \mathbf{x}_j^{(n)}, \Delta t). \tag{6}$$

If $\Omega \subseteq \mathbf{R}^s$, we define $\mathscr{S}(\Omega)$ to be the collection of all measurable characteristic functions on Ω. The members of $\mathscr{S}(\Omega)$ are the *sensor functions*. They measure the contribution of the particles to the sensor area. Define the function $e(x) := (\mathrm{erf}(x) + 1)/2$, where erf is the error function

$$\mathrm{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-w^2}\, dw \qquad \text{for} \quad x \in \mathbf{R}.$$

From the definition of $\gamma^{(n+1)}$ in Eq. (6) we find that

$$\forall \sigma \in \mathscr{S}(\mathbf{R}^s) \qquad \int_{\mathbf{R}^s} \sigma(\mathbf{x})\, \gamma^{(n+1)}(\mathbf{x})\, d\mathbf{x} = \frac{1}{N} \sum_{0 \leqslant j < N} \int_{I^s} \sigma(\mathbf{x}_j^{(n)} + \mathbf{f}(\mathbf{y}))\, d\mathbf{y}, \tag{7}$$

where $\mathbf{f}(\mathbf{y}) := (f_1(y_1), ..., f_s(y_s))$ and $f_i(y_i) := \sqrt{4D_i\, \Delta t}\, e^{-1}(y_i)$. We recover $c^{(n+1)}$ from $\gamma^{(n+1)}$ by a quasi-Monte Carlo approximation.

• The particles are relabeled using a multi-index $\mathbf{a} = (a_1, ..., a_s)$ with $0 \leqslant a_i < b^{d_i}$ for $1 \leqslant i \leqslant s$, so that

$$a_1 = b_1, ..., a_{i-1} = b_{i-1}, a_i < b_i \Rightarrow x_{\mathbf{a},i}^{(n)} \leqslant x_{\mathbf{b},i}^{(n)}.$$

This is done algorithmically as follows. The magnitude of the coordinates x_1 of the particles is used to order them into b^{d_1} subsets of level 1. In each subset of level 1, the magnitude of the coordinates x_2 of the particles is used to order them into b^{d_2} subsets of level 2 and so on: see Fig. 1 in two dimensions.

• Let $\chi_{\mathbf{a}}$ be the characteristic function of $I_{\mathbf{a}} := \prod_{i=1}^{s} [a_i b^{-d_i}, (a_i + 1)\, b^{-d_i})$. To each $\sigma \in \mathscr{S}(\mathbf{R}^s)$ there corresponds $\Sigma^{(n)} \in \mathscr{S}(I^{2s})$ by the formula

$$\Sigma^{(n)}(\mathbf{y}) := \sum_{\mathbf{a}} \chi_{\mathbf{a}}(\mathbf{y}')\, \sigma(\mathbf{x}_{\mathbf{a}}^{(n)} + \mathbf{f}(\mathbf{y}'')) \qquad \text{for} \quad \mathbf{y} = (\mathbf{y}', \mathbf{y}'') \in I^s \times I^s. \tag{8}$$

Then we find that

$$\int_{\mathbf{R}^s} \sigma(\mathbf{x})\, \gamma^{(n+1)}(\mathbf{x})\, d\mathbf{x} = \int_{I^{2s}} \Sigma^{(n)}(\mathbf{y})\, d\mathbf{y}. \tag{9}$$

We define $c^{(n+1)}$ as a sum of δ-measures by

$$\forall \sigma \in \mathscr{S}(\mathbf{R}^s) \qquad \int_{\mathbf{R}^s} \sigma(\mathbf{x})\, c^{(n+1)}(\mathbf{x}) = \frac{1}{N} \sum_{0 \leqslant j < N} \Sigma^{(n)}(\mathbf{y}_{nN+j}).$$

To summarize the algorithm, define the function

$$\mathbf{a}(\mathbf{y}) := (\lfloor b^{d_1} y_1 \rfloor, ..., \lfloor b^{d_s} y_s \rfloor) \qquad \text{for} \quad \mathbf{y} \in I^s,$$

where $\lfloor z \rfloor$ denotes the greatest integer $\leqslant z$ and put $Y^{(n)} := \{\mathbf{y}_{nN+j} : 0 \leqslant j < N\}$. Since $P' Y^{(n)}$ is a $(0, m, s)$-net in base b, the map $j \to \mathbf{a}(\mathbf{y}'_{nN+j})$ is

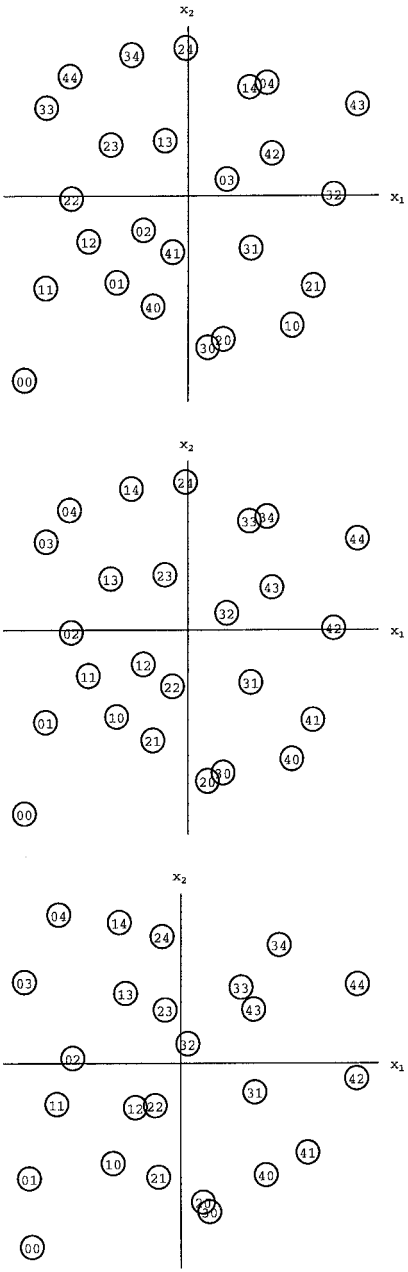

FIG. 1. Particles locations at selected times: $t = t_n$ before (top) and after relabeling (middle), and $t = t_{n+1}$ (bottom). Here $b = 5$ and $d_1 = d_2 = 1$.

one-to-one from $[0, N)$ into $[0, b^{d_1}) \times \cdots \times [0, b^{d_s})$. The location of the $\mathbf{a}(\mathbf{y}'_{nN+j})$th particle is updated according to

$$\mathbf{x}^{(n+1)}_{\mathbf{a}(\mathbf{y}'_{nN+j})} = \mathbf{x}^{(n)}_{\mathbf{a}(\mathbf{y}'_{nN+j})} + \mathbf{f}(\mathbf{y}''_{nN+j}). \tag{10}$$

Here $\mathbf{f}(\mathbf{y}''_{nN+j})$ is well defined because $\mathbf{y}''_{nN+j} \in (0, 1)^s$. Figure 1 explains the relabeling scheme and the motion of particles in two dimensions over one timestep.

3. CONVERGENCE OF THE METHOD

In this section, we will prove convergence of the quasi-Monte Carlo method. We define a weak measure of error which tells how well the particle distribution $c^{(n)}$ approximates the exact solution $c_n := c(\cdot, t_n)$. If ρ is a non-negative and Riemann-integrable function on $\mathbf{R}^s$ with $\int_{\mathbf{R}^s} \rho(\mathbf{x})\, d\mathbf{x} = 1$, the *star ρ-discrepancy* of a point set $X = \{\mathbf{x}_0, ..., \mathbf{x}_{N-1}\}$ is defined by

$$D^*_N(X; \rho) := \sup_{\mathbf{w} \in \mathbf{R}^s} \left| \frac{1}{N} \sum_{0 \leqslant j < N} \sigma_{\mathbf{w}}(\mathbf{x}_j) - \int_{\mathbf{R}^s} \sigma_{\mathbf{w}}(\mathbf{x})\, \rho(\mathbf{x})\, d\mathbf{x} \right|,$$

where $\sigma_{\mathbf{w}}$ denotes the characteristic function of the interval $\prod_{i=1}^{s} (-\infty, w_i)$ (see [HM72]). Let $X^{(n)}$ be the point set consisting of $\mathbf{x}^{(n)}_0, ..., \mathbf{x}^{(n)}_{N-1}$. For $\mathbf{w} \in \mathbf{R}^s$, we write

$$d^{(n)}_N(\mathbf{w}) := \frac{1}{N} \sum_{\mathbf{a}} \sigma_{\mathbf{w}}(\mathbf{x}^{(n)}_{\mathbf{a}}) - \int_{\mathbf{R}^s} \sigma_{\mathbf{w}}(\mathbf{x})\, c_n(\mathbf{x})\, d\mathbf{x},$$

where we sum over all $\mathbf{a} = (a_1, ..., a_s)$ with $0 \leqslant a_i < b^{d_i}$ for $1 \leqslant i \leqslant s$. The corresponding sup over all $\mathbf{w}$s is the star c_n-discrepancy of $X^{(n)}$.

It is easy to see that the quantity $d^{(n+1)}_N(\mathbf{w})$ is composed of two parts

$$d^{(n+1)}_N(\mathbf{w}) = \left(\frac{1}{N} \sum_{\mathbf{a}} e^{\Delta t}_{\mathbf{w}}(\mathbf{x}^{(n)}_{\mathbf{a}}) - \int_{\mathbf{R}^s} e^{\Delta t}_{\mathbf{w}}(\mathbf{x})\, c_n(\mathbf{x})\, d\mathbf{x} \right) + \delta^{(n)}_N(\mathbf{w}). \tag{11}$$

Here

$$e^{\Delta t}_{\mathbf{w}}(\mathbf{x}) := \int_{\mathbf{R}^s} E(\mathbf{v} - \mathbf{x}, \Delta t)\, \sigma_{\mathbf{w}}(\mathbf{v})\, d\mathbf{v} = \prod_{i=1}^{s} e\left(\frac{w_i - x_i}{\sqrt{4 D_i\, \Delta t}} \right) \qquad \text{for} \quad \mathbf{x} \in \mathbf{R}^s,$$

and $\delta_N^{(n)}(\mathbf{w})$ is the *error of the quasi-Monte Carlo approximation*

$$\delta_N^{(n)}(\mathbf{w}) := \frac{1}{N} \sum_{0 \leqslant j < N} \Sigma_{\mathbf{w}}^{(n)}(\mathbf{y}_{nN+j}) - \int_{I^{2s}} \Sigma_{\mathbf{w}}^{(n)}(\mathbf{y})\, d\mathbf{y},$$

where $\Sigma_{\mathbf{w}}^{(n)}$ is defined by Eq. (8) with $\sigma = \sigma_{\mathbf{w}}$.

Since $e_{\mathbf{w}}^{\Delta t}$ is a positive continuous function, we obtain the bound for the first difference in (11) if we use the following variant of the classical Koksma–Hlawka inequality.

LEMMA 1. *Suppose ρ is a Riemann-integrable function on $\mathbf{R}^s$ such that $\rho \geqslant 0$ and $\int_{\mathbf{R}^s} \rho(\mathbf{x})\, d\mathbf{x} = 1$. Let f be a function on $\mathbf{R}^s$ such that f and $|f|$ are of bounded variation in the sense of Hardy and Krause. If f or ρ is continuous and if $\mathbf{x}_0, ..., \mathbf{x}_{N-1}$ are points in $\mathbf{R}^s$, then*

$$\left| \frac{1}{N} \sum_{0 \leqslant j < N} f(\mathbf{x}_j) - \int_{\mathbf{R}^s} f(\mathbf{x})\, \rho(\mathbf{x})\, d\mathbf{x} \right| \leqslant V(f)\, D_N^*(X; \rho).$$

Furthermore, since $V(e_{\mathbf{w}}^{\Delta t}) = 1$, we have

$$\left| \frac{1}{N} \sum_{\mathbf{a}} e_{\mathbf{w}}^{\Delta t}(\mathbf{x}_{\mathbf{a}}^{(n)}) - \int_{\mathbf{R}^s} e_{\mathbf{w}}^{\Delta t}(\mathbf{x})\, c_n(\mathbf{x})\, d\mathbf{x} \right| \leqslant D_N^*(X^{(n)}; c_n). \tag{12}$$

Consider now $\delta_N^{(n)}(\mathbf{w})$. Define

$$e_{\mathbf{a},i}^{(n)}(w) := e\left(\frac{w - x_{\mathbf{a},i}^{(n)}}{\sqrt{4D_i\, \Delta t}} \right) \qquad \text{for} \quad w \in \mathbf{R},$$

$$\mathscr{E}_{\mathbf{w}}^{(n)} := \bigcup_{\mathbf{a}} I_{\mathbf{a}} \times \prod_{i=1}^{s} [0, e_{\mathbf{a},i}^{(n)}(w_i)).$$

Then $\Sigma_{\mathbf{w}}^{(n)}$ is the characteristic function of the staircase set $\mathscr{E}_{\mathbf{w}}^{(n)}$. Hence

$$\delta_N^{(n)}(\mathbf{w}) = D(\mathscr{E}_{\mathbf{w}}^{(n)}, Y^{(n)}). \tag{13}$$

LEMMA 2. *We have the upper bound*

$$|\delta_N^{(n)}(\mathbf{w})| \leqslant b^{d_1 + \cdots + d_{s-1} + \lfloor d_s/2 \rfloor} D_N(Y^{(n)}) + \frac{1}{b^{d_1}} + \cdots + \frac{1}{b^{d_{s-1}}} + \frac{1}{b^{\lfloor d_s/2 \rfloor}}.$$

Proof. Let $\delta_s \leqslant d_s$ be an integer. Define the function

$$\mathbf{e}_{\mathbf{a}}^{(n)}(\mathbf{w}) := (e_{\mathbf{a},1}^{(n)}(w_1), ..., e_{\mathbf{a},s}^{(n)}(w_s)) \qquad \text{for} \quad \mathbf{w} \in \mathbf{R}^s.$$

Because the particles are reordered so that

$$a_1 = b_1, ..., a_{i-1} = b_{i-1}, a_i < b_i \Rightarrow x_{\mathbf{a},i}^{(n)} \leqslant x_{\mathbf{b},i}^{(n)},$$

we can define partitions of $[0, 1]$,

$$0 = y_{0,1}^{(n)}(\mathbf{w}) \leqslant y_{1,1}^{(n)}(\mathbf{w}) \leqslant \cdots \leqslant y_{b^{d_1},1}^{(n)}(\mathbf{w}) = 1,$$

$$0 = y_{\alpha_1,0,2}^{(n)}(\mathbf{w}) \leqslant y_{\alpha_1,1,2}^{(n)}(\mathbf{w}) \leqslant \cdots \leqslant y_{\alpha_1,b^{d_2},2}^{(n)}(\mathbf{w}) = 1, \qquad \text{for} \quad 0 \leqslant \alpha_1 < b^{d_1},$$

$$\cdots$$

$$0 = y_{\alpha_1,...,\alpha_{s-1},0,s}^{(n)}(\mathbf{w}) \leqslant y_{\alpha_1,...,\alpha_{s-1},1,s}^{(n)}(\mathbf{w}) \leqslant \cdots \leqslant y_{\alpha_1,...,\alpha_{s-1},b^{\delta_s},s}^{(n)}(\mathbf{w}) = 1,$$
$$\text{for} \quad 0 \leqslant \alpha_1 < b^{d_1}, ..., 0 \leqslant \alpha_{s-1} < b^{d_{s-1}},$$

such that, for all integers $0 \leqslant \alpha_1 < b^{d_1}, ..., 0 \leqslant \alpha_{s-1} < b^{d_{s-1}}, 0 \leqslant \alpha_s < b^{\delta_s}$ and a_s with $\alpha_s b^{d_s - \delta_s} \leqslant a_s < (\alpha_s + 1)\, b^{d_s - \delta_s}$, we have

$$\mathbf{e}_{\alpha_1,...,\alpha_{s-1},a_s}^{(n)}(\mathbf{w}) \in [\, y_{\alpha_1,1}^{(n)}(\mathbf{w}),\, y_{\alpha_1+1,1}^{(n)}(\mathbf{w})\,] \times \cdots \times [\, y_{\alpha_1,...,\alpha_s,s}^{(n)}(\mathbf{w}),\, y_{\alpha_1,...,\alpha_s+1,s}^{(n)}(\mathbf{w})\,].$$

If we put $J_\alpha := \prod_{i=1}^{s-1} [\alpha_i b^{-d_i}, (\alpha_i + 1)\, b^{-d_i}) \times [\alpha_s b^{-\delta_s}, (\alpha_s + 1)\, b^{-\delta_s})$ and

$$\underline{\mathscr{E}}_{\mathbf{w}}^{(n)} := \bigcup_{\alpha} J_\alpha \times [0,\, y_{\alpha_1,1}^{(n)}(\mathbf{w})) \times \cdots \times [0,\, y_{\alpha_1,...,\alpha_s,s}^{(n)}(\mathbf{w})),$$

$$\overline{\mathscr{E}}_{\mathbf{w}}^{(n)} := \bigcup_{\alpha} J_\alpha \times [0,\, y_{\alpha_1+1,1}^{(n)}(\mathbf{w})) \times \cdots \times [0,\, y_{\alpha_1,...,\alpha_s+1,s}^{(n)}(\mathbf{w})),$$

$$\partial\mathscr{E}_{\mathbf{w}}^{(n)} := \bigcup_{\alpha} J_\alpha \times ([\, y_{\alpha_1,1}^{(n)}(\mathbf{w}),\, y_{\alpha_1+1,1}^{(n)}(\mathbf{w})\,] \times I^{s-1} \cup \cdots \cup [0,\, y_{\alpha_1,1}^{(n)}(\mathbf{w})) \times \cdots \times [\, y_{\alpha_1,...,\alpha_s,s}^{(n)}(\mathbf{w}),\, y_{\alpha_1,...,\alpha_s+1,s}^{(n)}(\mathbf{w})\,]),$$

then $\underline{\mathscr{E}}_{\mathbf{w}}^{(n)} \subset \mathscr{E}_{\mathbf{w}}^{(n)} \subset \overline{\mathscr{E}}_{\mathbf{w}}^{(n)}$ and $\overline{\mathscr{E}}_{\mathbf{w}}^{(n)} \backslash \underline{\mathscr{E}}_{\mathbf{w}}^{(n)} \subset \partial\mathscr{E}_{\mathbf{w}}^{(n)}$. Thus

$$D(\underline{\mathscr{E}}_{\mathbf{w}}^{(n)},\, Y^{(n)}) - \lambda_{2s}(\partial\mathscr{E}_{\mathbf{w}}^{(n)}) \leqslant D(\mathscr{E}_{\mathbf{w}}^{(n)},\, Y^{(n)}) \leqslant D(\overline{\mathscr{E}}_{\mathbf{w}}^{(n)},\, Y^{(n)}) + \lambda_{2s}(\partial\mathscr{E}_{\mathbf{w}}^{(n)}).$$

The subsets $\underline{\mathscr{E}}_{\mathbf{w}}^{(n)}$ and $\overline{\mathscr{E}}_{\mathbf{w}}^{(n)}$ are disjoint unions of $b^{d_1+\cdots+d_{s-1}+\delta_s}$ subintervals of I^{2s}, hence

$$\max(D(\underline{\mathscr{E}}_{\mathbf{w}}^{(n)}, Y^{(n)}), D(\overline{\mathscr{E}}_{\mathbf{w}}^{(n)}, Y^{(n)})) \leqslant b^{d_1+\cdots+d_{s-1}+\delta_s} D_N(Y^{(n)}).$$

On the other hand,

$$\lambda_{2s}(\partial \mathscr{E}_{\mathbf{w}}^{(n)}) \leqslant \frac{1}{b^{d_1}} + \cdots + \frac{1}{b^{d_{s-1}}} + \frac{1}{b^{\delta_s}}.$$

If we choose $\delta_s = \lfloor d_s/2 \rfloor$, we obtain the desired inequality. ∎

We can combine the previous bounds to obtain an upper bound for the error of the quasi-Monte Carlo method.

PROPOSITION 1. *The star c_n-discrepancy of $X^{(n)}$ satisfies*

$$D_N^*(X^{(n)}; c_n) \leqslant D_N^*(X^{(0)}; c_0) + b^{d_1+\cdots+d_{s-1}+\lfloor d_s/2 \rfloor} \sum_{m=0}^{n-1} D_N(Y^{(m)}) + n\left(\frac{1}{b^{d_1}} + \cdots + \frac{1}{b^{d_s-1}} + \frac{1}{b^{\lfloor d_s/2 \rfloor}}\right).$$

Remark 1. In two dimensions or higher, there are many choices for $d_1, \ldots, d_s$ with $b^{d_1+\cdots+d_s} = N$. The *optimal choice* will minimize the right-hand side of the bound for the error.

4. COMPUTATIONAL EXAMPLES

In this section, we present results of numerical experiments which illustrate the theoretical analysis. The quasi-Monte Carlo (QMC) method was applied to the initial-value problem (1)–(2) with diffusion coefficients $D_i = 1.0$ for two different sets of initial data:

- Problem 1 (smooth data)

$$c_0(\mathbf{x}) = \prod_{i=1}^{s} \frac{1}{\sqrt{\pi}} \mathrm{e}^{-x_i^2}.$$

- Problem 2 (non-smooth data)

$$c_0(\mathbf{x}) = \prod_{i=1}^{s} \chi_{[-1/2,\,+1/2)}(x_i),$$

where $\chi_{[-1/2,\,+1/2)}$ is the characteristic function of the interval $[-\frac{1}{2}, +\frac{1}{2})$.

For each of these problems, a closed-form analytic solution is available with which to compare the numerical solution.

All simulations were conducted by first sampling N particles from the initial distribution

$$c_0(\mathbf{x}) = \prod_{i=1}^{s} c_{0,\,i}(x_i).$$

We used a $(0, m, s)$-net in base b, Ξ and the sampling was done by mapping Ξ to $\mathbf{R}^s$ using the inverse function of

$$C_0(\mathbf{x}) := \left(\int_{-\infty}^{x_i} c_{0,\,i}(w)\, dw \right)_{1 \leqslant i \leqslant s}.$$

For the QMC simulations we need a low-discrepancy sequence Y of points in I^{2s}. We used the $2s$-dimensional Faure sequence in base b, where b is the least prime $\geqslant 2s$ (see [Fau82]). The Faure sequence $S^b_{[1,\,2s]} = \{\mathbf{z}_0, \mathbf{z}_1, \ldots\}$ is a $(0, 2s)$-sequence in base b. We chose

$$\Xi = \left\{ \frac{2j+1}{2N} : 0 \leqslant j < N \right\}$$

in dimension $s = 1$ and $\Xi = P'\{\mathbf{z}_j : 0 \leqslant j < N\}$ for $s > 1$. In addition we chose $Y = \{\mathbf{z}_j : N \leqslant j\}$.

In dimension $s > 1$, the computation time for the discrepancy $D^*_N(X^{(n)}; c_n)$ would be prohibitive, therefore the *error* was measured by computing

$$\tilde{D}^*_N(X^{(n)}; c_n) := \max_{1 \leqslant i \leqslant s} \sup_{w \in \mathbf{R}} \left| \frac{1}{N} \sum_{\mathbf{a}} \sigma^{(i)}_w(\mathbf{x}^{(n)}_{\mathbf{a}}) - \int_{\mathbf{R}^s} \sigma^{(i)}_w(\mathbf{x})\, c_n(\mathbf{x})\, d\mathbf{x} \right|,$$

where $\sigma^{(i)}_w$ is the characteristic function of the interval $\{\mathbf{w} \in \mathbf{R}^s : w_i < w\}$. Assuming that the error is approximately αN^{-p}, one can estimate α and p

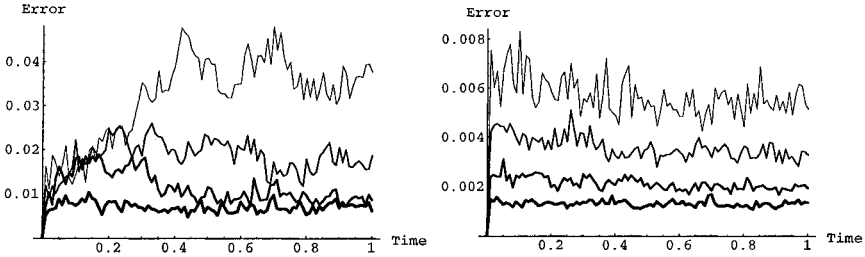

FIG. 2. Problem 1, dimension $s=1$, $N=2^{10}, 2^{11}, 2^{12}, 2^{13}$. MC (left) vs QMC (right) results. Thick lines correspond to large N.

from plots of the error versus N. This can be accomplished as follows. The simulation is run up to M time steps, and the *averaged error*

$$D_N := \frac{1}{M+1} \sum_{n=0}^{M} \tilde{D}_N^*(X^{(n)}; c_n)$$

is computed. The constant prefactor α and the exponent p are then estimated by finding the best (in the sense of least squares) straight line fit to the log-log plot of the data.

Figures 2–7 compare errors obtained in solving Problems 1 and 2 using two different methods:

1. quasi-Monte Carlo (QMC) using renumbering,
2. Monte Carlo (MC) utilizing pseudo-random numbers without renumbering.

For Problem 1 a time step of $\Delta t = 0.01$ was chosen and the particles were moved 100 steps out to time $T=1.0$. For Problem 2 we advance up to $T=0.01$ with a time step of $\Delta t = 0.0001$.

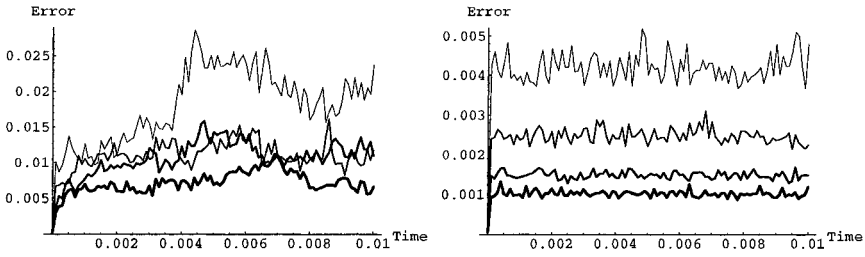

FIG. 3. Problem 2, dimension $s=1$, $N=2^{10}, 2^{11}, 2^{12}, 2^{13}$. MC (left) vs QMC (right) results. Thick lines correspond to large N.

Figure 2 shows the result of applying MC and QMC methods to Problem 1 in dimension $s=1$ with the number N of particles made progressively greater. Note the difference in vertical scales. The quasi-random strategy is much better than pseudo-random here. The same error level is attained by the QMC method with $N=2^{10}$ particles and by the MC scheme with $N=2^{13}$ particles.

The effect of using pseudo-random and quasi-random sequences is illustrated also in Fig. 3 for Problem 2 in dimension $s=1$. The results for Problem 2 are similar to those for the first problem.

The results of the convergence experiment for these problems are given in Fig. 4. One finds for Problem 1

$$D_N(\mathrm{MC}) \approx \frac{0.71}{N^{0.49}}, \qquad D_N(\mathrm{QMC}) \approx \frac{0.76}{N^{0.71}}, \tag{14}$$

and for Problem 2

$$D_N(\mathrm{MC}) \approx \frac{0.55}{N^{0.50}}, \qquad D_N(\mathrm{QMC}) \approx \frac{0.61}{N^{0.71}}. \tag{15}$$

In both cases the QMC method outperforms the MC method for sufficiently large numbers of particles.

To further investigate the performance of quasi-random sequences and renumbering for simulations, other experiments were run to solve two-dimensional problems. Figure 5 depicts results obtained in solving Problem 1 in dimension $s=2$ using MC method and QMC method. Note the change of scale. For the QMC method, only results of the optimal choice are depicted, since it was found that other choices can fail to improve on

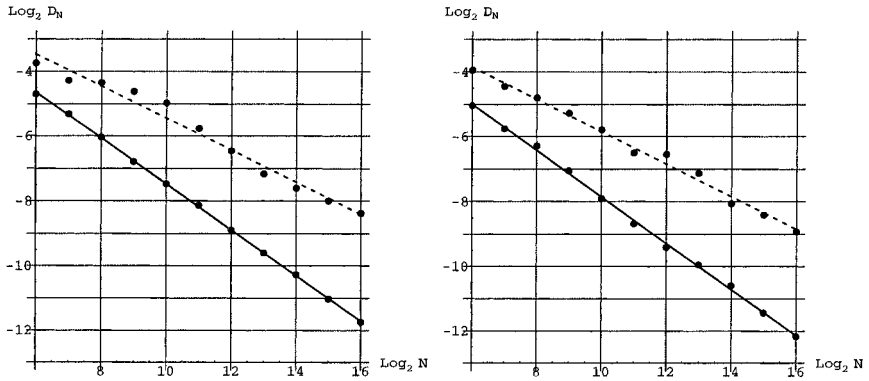

FIG. 4. Dimension $s=1$, linear fits to the averaged error for Problem 1 (left) and Problem 2 (right). Dashed lines correspond to MC results and solid lines to QMC results.

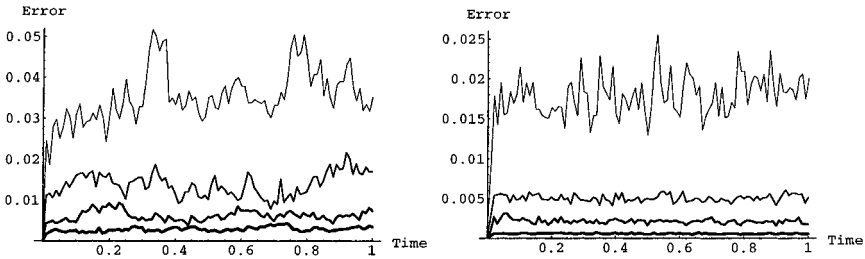

FIG. 5. Problem 1, dimension $s=2$, $N=5^{2+2}$, 5^{2+3}, 5^{3+3}, 5^{3+4}. MC (left) vs QMC (right) results. Thick lines correspond to large N.

Monte Carlo. We see that quasi-random simulation still clearly outperforms standard pseudo-random simulation, although not to as large an extent as in dimension $s=1$. The QMC error with $N=5^{3+3}$ particles is significantly less than the MC error with $N=5^7$ particles.

A comparison of pseudo-random and quasi-random strategies for Problem 2 in dimension $s=2$ is shown in Fig. 6. Once again, the superiority of the QMC algorithm is clearly demonstrated.

Figure 7 shows plots of the averaged error as a function of N on a log-log scale. One finds for Problem 1

$$D_N(\mathrm{MC}) \approx \frac{0.97}{N^{0.52}}, \qquad D_N(\mathrm{QMC}) \approx \frac{1.08}{N^{0.66}}, \tag{16}$$

and for Problem 2

$$D_N(\mathrm{MC}) \approx \frac{0.65}{N^{0.51}}, \qquad D_N(\mathrm{QMC}) \approx \frac{1.01}{N^{0.66}}, \tag{17}$$

As in dimension $s=1$, the QMC method produces lower errors than the MC errors and faster convergence.

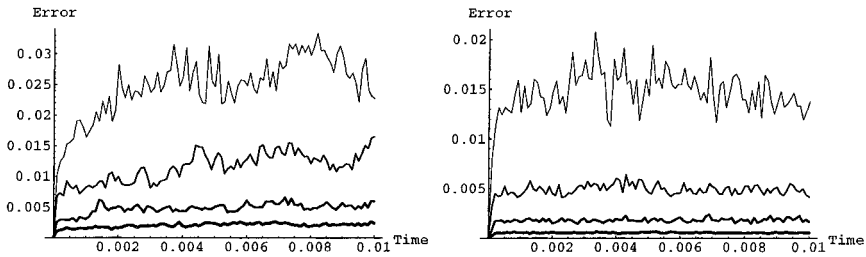

FIG. 6. Problem 2, dimension $s=2$, $N=5^{2+2}$, 5^{2+3}, 5^{3+3}, 5^{3+4}. MC (left) vs QMC (right) results. Thick lines correspond to large N.

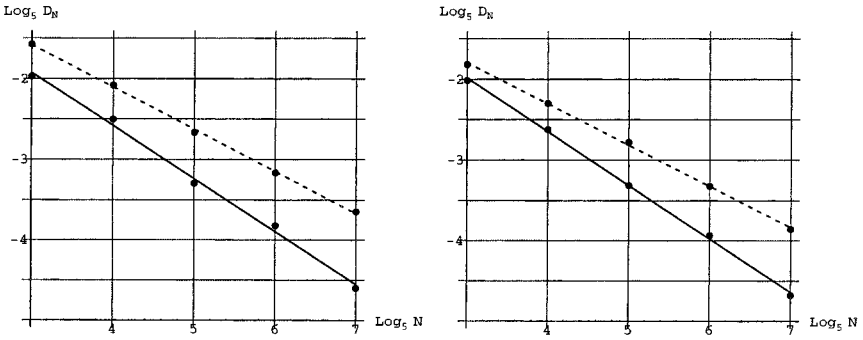

FIG. 7. Dimension $s = 2$, linear fits to the averaged error for Problem 1 (left) and Problem 2 (right). Dashed lines correspond to MC results and solid lines to QMC results.

In dimension $s = 2$, a more conventional measure of the error is to form an approximation to the particle solution and to compute the L^2-norm of the difference between this empirical distribution and an approximation to the exact solution, constructed in the same way. We introduce a *cut-off* function

$$\zeta(\mathbf{x}) := \frac{1}{\pi} \mathrm{e}^{-(x_1^2 + x_2^2)}, \qquad \text{for} \quad \mathbf{x} = (x_1, x_2).$$

Then we set for all $\varepsilon > 0$

$$\zeta_\varepsilon(\mathbf{x}) := \frac{1}{\varepsilon^2} \zeta\left(\frac{\mathbf{x}}{\varepsilon}\right).$$

Next we define

$$c_\varepsilon^{(n)} := c^{(n)} * \zeta_\varepsilon \qquad \text{and} \quad c_\varepsilon(\cdot, t) := c_\varepsilon(\cdot, t) * \zeta_\varepsilon.$$

The L^2-norm of the difference $c_\varepsilon^{(n)} - c_\varepsilon(\cdot, t_n)$ was computed. In Fig. 8 we display the result of solving Problem 1 with the choice $\varepsilon = 0.1$. One can see that the MC results are again worst.

The QMC method requires more computational effort, since the particles have to be renumbered after each time step. Table I lists the CPU times in seconds on a Cray C90 computer for Problems 1 and 2 in dimension $s = 1$. The CPU times in dimension $s = 2$ are listed in Table II. During each time step, each particle takes a step chosen from a Gaussian distribution. Gaussian numbers are generated from uniform numbers by inverting the error function. This technique is time-consuming. The renumberings do not add a significant amount of CPU time to the overall method.

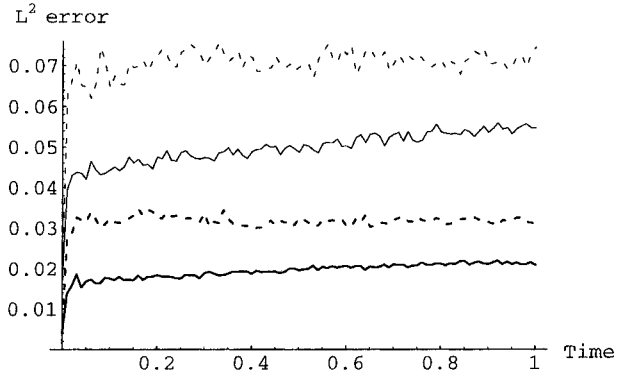

FIG. 8. Problem 1, dimension $s=2$, L^2 errors of MC (dashed) and QMC (solid) methods, $N=5^{2+3}$, 5^{3+3}. Thick lines correspond to large N.

TABLE I

CPU Times in Seconds of One-Dimensional Simulations

	N	2^{10}	2^{11}	2^{12}	2^{13}
Problem 1	MC	14.34	28.69	57.39	114.71
	QMC	14.54	29.16	58.43	117.07
Problem 2	MC	14.54	29.01	58.12	116.30
	QMC	14.75	29.56	59.35	118.89

TABLE II

CPU Times in Seconds of Two-Dimensional Simulations

	N	5^{2+2}	5^{2+3}	5^{3+3}	5^{3+4}
Problem 1	MC	17.92	89.53	447.81	2240.32
	QMC	18.28	91.11	460.50	2295.01
Problem 2	MC	18.05	90.26	451.09	2243.10
	QMC	18.08	91.68	455.57	2276.48

Remark 2. The application of quasi-random sequences to random walk simulations cannot be implemented in a straightforward fashion, because of correlations. Without some kind of scrambling, this will cause the particles to just repeat the same pattern over and over, and thus the method will not converge. For Problem 1 in dimension $s = 1$, it was found that the error of the method using a quasi-random sequence without renumbering was larger than 0.10 after 7 time steps. There was little difference in the errors obtained with $N = 2^{10}$, $N = 2^{11}$, or $N = 2^{12}$ particles.

5. CONCLUSION

In this paper we have presented a particle method which replaces the pseudo-random samples in the random walk method by quasi-random points. In every time step the number order of the particles is scrambled according to their positions before assigning a new quasi-random point to each particle. We have shown that the algorithm converges as the number of particles increases. For some model problems, the error was found to be significantly less when quasi-random points were used than when a standard random walk calculation was performed. Quasi-random points have also been applied to simple nonlinear problems. We refer to [Lec91, CL97] for 1-D kinetic equations. The results show that quasi-Monte Carlo simulation is superior to standard Monte Carlo in magnitude of error and in convergence rate. The improvement degrades in two dimensions or higher. This will probably be true for the random vortex method for viscous flow [Cho73], but further investigation will be required to fully answer this question.

ACKNOWLEDGMENTS

The numerical experiments were performed on the Cray C94 computer at the Centre Grenoblois de Calcul Vectoriel du Commissariat à l'Énergie Atomique. We thank this institution for providing us access to the machine. We also thank the referees for their constructive criticism on the first version of this paper and for several valuable suggestions.

REFERENCES

[Cho73] A. J. Chorin, Numerical study of slightly viscous flow, *J. Fluid Mech.* **57** (1973), 785–796.

[CL97] I. Coulibaly and C. Lécot, Particle simulations of the Kac model of the Boltzmann equation, *J. Comput. Appl. Math.* **87** (1997), 169–193.

[DT97] M. Drmota and R. F. Tichy, "Sequences, Discrepancies and Applications," Springer-Verlag, Berlin, 1997.

[Fau82] H. Faure, Discrépance de suites associées à un système de numération (en dimension s), *Acta Arith.* **41** (1982), 337–351.

[HM72] E. Hlawka and R. Mück, A transformation of equidistributed sequences, *in* "Applications of Number Theory to Numerical Analysis" (S. K. Zaremba, Ed.), pp. 371–388, 1972.

[HM97] N. Hofmann and P. Mathé, On quasi-Monte Carlo simulation of stochastic differential equations, *Math. Comp.* **66** (1997), 573–589.

[Lec91] C. Lécot, A quasi-Monte Carlo method for the Boltzmann equation, *Math. Comp.* **56** (1991), 621–644.

[LC98] C. Lécot and I. Coulibaly, A quasi-Monte Carlo scheme using nets for a linear Boltzmann equation, *SIAM J. Numer. Anal.* **35** (1998), 51–70.

[MC93] W. J. Morokoff and R. E. Caflisch, A quasi-Monte Carlo approach to particle simulation of the heat equation, *SIAM J. Numer. Anal.* **30** (1993), 1558–1573.

[Mos97] B. S. Moskowitz, Neutron transport calculations using quasi-Monte Carlo methods, *in* "Proc. Internat. Conference on Mathematical Methods and Supercomputing for Nuclear Applications, Saratoga Springs," pp. 531–540, 1997.

[Nie87] H. Niederreiter, Point sets and sequences with small discrepancy, *Monatsh. Math.* **104** (1987), 273–337.

[Nie88] H. Niederreiter, Low-discrepancy and low-dispersion sequences, *J. Number Theory* **30** (1988), 51–70.

[Nie92] H. Niederreiter, "Random Number Generation and Quasi-Monte Carlo Methods," Soc. for Industr. Appl. Math., Philadelphia, 1992.

[NS95] H. Niederreiter and P. J.-S. Shiue (Eds.), "Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing," Springer-Verlag, New York, 1995.

[NHLZ98] H. Niederreiter, P. Hellekalek, G. Larcher, and P. Zinterhof (Eds.), "Monte Carlo and Quasi-Monte Carlo Methods 1996," Springer-Verlag, New York, 1998.