MATHEMATICS OF COMPUTATION
Volume 73, Number 248, Pages 1953–1966
S 0025-5718(04)01627-8
Article electronically published on January 5, 2004

A QUASI–MONTE CARLO SCHEME FOR SMOLUCHOWSKI'S COAGULATION EQUATION

CHRISTIAN LÉCOT AND WOLFGANG WAGNER

ABSTRACT. This paper analyzes a Monte Carlo algorithm for solving Smoluchowski's coagulation equation. A finite number of particles approximates the initial mass distribution. Time is discretized and random numbers are used to move the particles according to the coagulation dynamics. Convergence is proved when quasi-random numbers are utilized and if the particles are relabeled according to mass in every time step. The results of some numerical experiments show that the error of the new algorithm is smaller than the error of a standard Monte Carlo algorithm using pseudo-random numbers without reordering the particles.

INTRODUCTION

Models of coalescence (i.e., coagulation, gelation, aggregation, agglomeration, accretion, etc.) mainly stem from the work of Smoluchowski on coagulation processes in colloids [15, 16]. Smoluchowski proposed the following infinite system of differential equations for the evolution of the number $N_0 c(i,t)$ of clusters of mass i for $i = 1, 2, 3 \ldots$:

$$\frac{\partial c}{\partial t}(i,t) = \frac{1}{2} \sum_{1 \le j < i} K(i-j,j) c(i-j,t) c(j,t) - \sum_{j \ge 1} K(i,j) c(i,t) c(j,t). \tag{1}$$

Here N_0 is the total number of clusters at time $t = 0$, so that $\sum_{i \ge 1} c(i,0) = 1$, and $K(i,j)$ is the coagulation kernel. Numerical solution of the Smoluchowski's coagulation equation is a difficult task for deterministic methods, so several stochastic algorithms have been proposed [8, 3, 17, 7, 11, 14, 2, 4]. The Monte Carlo (MC) schemes take a system of test particles which interact and form clusters according to the dynamics described in (1). Random numbers are used to find out which clusters interact and to determine the size of the new clusters. Despite the versatility of MC methods, a drawback is their slow convergence. An approach to acceleration is to change the choice of random numbers used. Quasi–Monte Carlo (QMC) methods use quasi-random numbers instead of pseudo-random numbers and can achieve better convergence in certain cases [5].

The efficiency of a QMC method depends on the quality of the quasi-random points that are used. These points should form a low-discrepancy point set. We recall from [13] some basic notations and concepts. If $s \ge 1$ is a fixed dimension,

Received by the editor November 11, 2002 and, in revised form, March 14, 2003.
2000 *Mathematics Subject Classification.* Primary 65C05; Secondary 70-08, 82C80.
Computation was supported by the Centre Grenoblois de Calcul Vectoriel du Commissariat à l'Énergie Atomique, France.

then $I^s := [0,1)^s$ is the s-dimensional half-open unit cube and λ_s denotes the s-dimensional Lebesgue measure. For a point set X consisting of $\mathbf{x}_0, \ldots, \mathbf{x}_{N-1} \in I^s$ and for a Lebesgue-measurable subset Q of I^s we define the *local discrepancy* by

$$D_N(Q,X) := \frac{1}{N} \sum_{0 \le p < N} c_Q(\mathbf{x}_p) - \lambda_s(Q),$$

where c_Q is the characteristic function of Q. The *discrepancy* of the point set X is defined by

$$D_N(X) := \sup_Q |D_N(Q,X)|,$$

the supremum being taken over all subintervals of I^s. The *star discrepancy* of X is

$$D_N^\star(X) := \sup_{Q^\star} |D_N(Q^\star,X)|,$$

where $Q^\star$ runs through all subintervals of I^s with one vertex at the origin. The idea of (t,m,s)-nets is to consider point sets X for which $D_N(Q,X) = 0$ for a large family of intervals Q. Such point sets should have a small discrepancy. For an integer $b \ge 2$, an interval of the form

$$\prod_{r=1}^{s} \left[\frac{a_r}{b^{d_r}}, \frac{a_r+1}{b^{d_r}} \right),$$

with integers $d_r \ge 0$ and $0 \le a_r < b^{d_r}$ for $1 \le r \le s$, is called an *elementary interval in base b*. If $0 \le t \le m$ are integers, a *(t,m,s)-net in base b* is a point set X consisting of b^m points in I^s such that $D_N(Q,X) = 0$ for every elementary interval Q in base b with measure $\lambda_s(Q) = b^{t-m}$. The sequence analog of this concept is as follows. If $b \ge 2$ and $t \ge 0$ are integers, a sequence $\mathbf{x}_0, \mathbf{x}_1, \ldots$ of points in I^s is a *(t,s)-sequence in base b* if, for all integers $n \ge 0$ and $m > t$, the points $\mathbf{x}_p$ with $nb^m \le p < (n+1)b^m$ form a (t,m,s)-net in base b. The following result is shown in [12].

Lemma 1. *Let X be a (t,m,s)-net in base b. For any elementary interval $Q' \subset I^{s-1}$ in base b and for any $x_s \in \bar{I}$,*

$$|D_{b^m}(Q' \times [0,x_s), X)| \le b^{t-m}.$$

The effectiveness of QMC methods has limitations. First, while they are valid for integration problems, they may not be directly applicable to simulations, due to the correlations between the points of a quasi-random sequence. This problem can be overcome by writing the desired result as an integral. This leads to a second limitation: the improved accuracy of QMC methods may be lost for problems in which the integrand is not smooth. It is necessary to take special measures to make optimal use of the greater uniformity associated with quasi-random sequences. This is achieved here through the additional step of reordering the particles at each time step. The aim of the paper is to construct and investigate a QMC method for Smoluchowski's coagulation equation. In Section 1 we present a particle scheme using quasi-random numbers for the solution of the equation. In Section 2 we prove the convergence of the method, as the number of simulated particles increases. In Section 3 we carry out numerical experiments based on a comparison of the method with a standard MC scheme.

1. The algorithm

We assume that the coagulation kernel $K(i,j)$ is nonnegative and symmetric

$$K(i,j) \geq 0 \quad \text{and} \quad K(i,j) = K(j,i).$$

Multiplying (1) by i and summing over all i, indicates that mass is conserved

$$\frac{d}{dt}\sum_{i\geq 1} ic(i,t) = 0, \tag{2}$$

provided an interchange of summation order on the right is valid. We refer to [9] for a study of existence, unicity, and conservation of mass of solutions. Rather than approximating the density of clusters $c(i,t)$, one can approximate the *mass density* $f(i,t) := ic(i,t)$, which satisfies the following equation for $i = 1, 2, 3\ldots$:

$$\frac{\partial f}{\partial t}(i,t) = \sum_{1\leq j<i} \widetilde{K}(i-j,j)f(i-j,t)f(j,t) - \sum_{j\geq 1} \widetilde{K}(i,j)f(i,t)f(j,t), \tag{3}$$

where

$$\widetilde{K}(i,j) := \frac{K(i,j)}{j}.$$

Equation (3) has been used in [2] for constructing a stochastic algorithm for Smoluchowski's coagulation equation. If $E \subset \mathbb{N}^\star := \{1,2,3,\ldots\}$, let σ_E denote the sequence

$$\sigma_E(i) = \begin{cases} 1 & \text{if } i \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Equation (3) can be given the form

$$\frac{d}{dt}\sum_{i\geq 1} f(i,t)\sigma_E(i) = \sum_{i,j\geq 1} \widetilde{K}(i,j)f(i,t)f(j,t)\left(\sigma_E(i+j) - \sigma_E(i)\right) \tag{4}$$

for any $E \subset \mathbb{N}^\star$. We denote by f_0 the initial data. We may assume

$$\sum_{i\geq 1} f_0(i) = 1. \tag{5}$$

We choose integers $b \geq 2, m \geq 1$ and we put $N := b^m$. We use a low-discrepancy sequence $X = \{\mathbf{x}_0, \mathbf{x}_1, \ldots\} \subset I^3$ for QMC approximation. We assume that X is a $(t,3)$-sequence in base b for some $t \geq 0$. If $X^n := \{\mathbf{x}_p : nN \leq p < (n+1)N\}$ and if π' denotes the projection defined by $\pi'(x_1,x_2,x_3) = (x_1,x_2)$, we assume that $\pi' X^n$ is a $(0,m,2)$-net in base b. We write δ_j for the unit mass at j

$$\delta_j(i) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases}$$

A sample J^0 of N particles $j_{0,0}, \ldots, j_{N-1,0}$ is chosen such that

$$f^0 := \frac{1}{N}\sum_{0\leq k<N} \delta_{j_{k,0}} \approx f_0.$$

It means that the point set J^0 has a small star f_0-discrepancy (see below); e.g., if we assume a monodisperse initial condition

$$f_0(1) = 1,\ f_0(2) = f_0(3) = \cdots = 0,$$

we set

$$j_{0,0} = \cdots = j_{N-1,0} = 1.$$

We assume that the kernel $\widetilde{K}(i,j)$ is bounded and we put

$$\widetilde{K}^* := \sup_{i,j\geq 1} \widetilde{K}(i,j).$$

We choose a time step Δt such that $\Delta t \widetilde{K}^* < 1$. Computations are still possible for unbounded kernels; see Section 3. We set $t_n := n\Delta t$ and $f_n(i) := f(i,t_n)$. If we assume that we have a point set J^n of N particles $j_{0,n}, \ldots, j_{N-1,n}$ such that

$$f^n := \frac{1}{N} \sum_{0\leq k<N} \delta_{j_{k,n}} \approx f_n, \tag{6}$$

we compute f^{n+1} in three steps.

- Relabeling the particles.

  $$j_{0,n} \leq j_{1,n} \leq \cdots \leq j_{N-1,n}.$$

  This ensures convergence of the scheme; see Lemma 5.
- Coagulation. We define g^{n+1} by

$$\frac{1}{\Delta t}\Big(\sum_{i\geq 1} g^{n+1}(i)\sigma_E(i) - \sum_{i\geq 1} f^n(i)\sigma_E(i)\Big) = \sum_{i,j\geq 1} \widetilde{K}(i,j) f^n(i) f^n(j) \left(\sigma_E(i+j) - \sigma_E(i)\right), \quad \text{for any } E \subset \mathbb{N}^\star,$$

  and so

$$\begin{aligned} \sum_{i\geq 1} g^{n+1}(i)\sigma_E(i) = {} & \frac{1}{N} \sum_{0\leq k<N} \Big(1 - \frac{\Delta t}{N} \sum_{0\leq \ell<N} \widetilde{K}(j_{k,n}, j_{\ell,n})\Big)\sigma_E(j_{k,n}) \\ & + \frac{\Delta t}{N^2} \sum_{0\leq k,\ell<N} \widetilde{K}(j_{k,n}, j_{\ell,n})\sigma_E(j_{k,n} + j_{\ell,n}). \end{aligned} \tag{7}$$

- QMC integration. Let $c_{k,\ell}$ be the characteristic function of

  $$R_{k,\ell} := \left[\frac{k}{N}, \frac{k+1}{N}\right) \times \left[\frac{\ell}{N}, \frac{\ell+1}{N}\right)$$

  and $\chi^n_{k,\ell}$ denote the characteristic function of $I^n_{k,\ell} := \big[0, \Delta t \widetilde{K}(j_{k,n}, j_{\ell,n})\big)$. For any $E \subset \mathbb{N}^\star$, define

$$\begin{aligned} C_E^{n+1}(\mathbf{x}) := \sum_{0\leq k,\ell<N} c_{k,\ell}(x_1,x_2)\Big(\big(1 - \chi^n_{k,\ell}(x_3)\big)\,\sigma_E(j_{k,n}) \\ + \chi^n_{k,\ell}(x_3)\sigma_E(j_{k,n} + j_{\ell,n})\Big), \quad \mathbf{x} \in I^3, \end{aligned} \tag{8}$$

  then

$$\sum_{i\geq 1} g^{n+1}(i)\sigma_E(i) = \int_{I^3} C_E^{n+1}(\mathbf{x})d\mathbf{x}. \tag{9}$$

  We obtain f^{n+1} by

$$\forall E \subset \mathbb{N}^\star \quad \sum_{i\geq 1} f^{n+1}(i)\sigma_E(i) = \frac{1}{N} \sum_{nN\leq p<(n+1)N} C_E^{n+1}(\mathbf{x}_p).$$

The last steps of the algorithm may be summarized as follows. Let us denote

$$k(x) := \lfloor Nx \rfloor, \quad x \in I.$$

Then for $nN \le p < (n+1)N$,

$$j_{k(x_{p,1}),n+1} = \begin{cases} j_{k(x_{p,1}),n} + j_{k(x_{p,2}),n} & \text{if } x_{p,3} < \Delta t \widetilde{K}\left(j_{k(x_{p,1}),n}, j_{k(x_{p,2}),n}\right), \\ j_{k(x_{p,1}),n} & \text{otherwise.} \end{cases}$$

2. Convergence analysis

We now establish a convergence result for the QMC algorithm. As is usual with particle method error estimates, convergence is shown in a weak sense. First we need to adapt the basic concepts of QMC methods to the present study. Let $s \ge 1$ be a fixed dimension. A *sequence* consists of an s-dimensional array of terms $u(\mathbf{i}) \in \mathbb{R}$, where $\mathbf{i} \in \mathbb{N}^{\star s}$. Let g be a sequence of nonnegative terms such that

$$\sum_{\mathbf{i} \in \mathbb{N}^{\star s}} g(\mathbf{i}) = 1. \tag{10}$$

For a point set J consisting of $\mathbf{j}_0, \dots, \mathbf{j}_{N-1} \in \mathbb{N}^{\star s}$ and for an arbitrary set $E \subset \mathbb{N}^{\star s}$ we define the *local g-discrepancy* by

$$D_N(E, J; g) := \frac{1}{N} \sum_{0 \le k < N} \sigma_E(\mathbf{j}_k) - \sum_{\mathbf{i} \in E} g(\mathbf{i}),$$

where σ_E denotes the sequence

$$\sigma_E(\mathbf{i}) = \begin{cases} 1 & \text{if } \mathbf{i} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The *star g-discrepancy* of the point set J is defined by

$$D_N^\star(J; g) := \sup_{\mathbf{h} \in \mathbb{N}^{\star s}} |D_N(E_{\mathbf{h}}, J; g)|,$$

where

$$E_{\mathbf{h}} := \begin{cases} \emptyset & \text{if } h_r = 1 \text{ for some } r, 1 \le r \le s, \\ \prod_{r=1}^{s} \{1, 2, \dots, h_r - 1\} & \text{otherwise.} \end{cases}$$

For a sequence u and for $\mathbf{j}, \mathbf{j}' \in \mathbb{N}^{\star s}$, let $T_{\mathbf{j}}^r u$ be the array defined by

$$T_{\mathbf{j}}^r u(\mathbf{i}) := u(i_1, \dots, i_{r-1}, j_r, i_{r+1}, \dots, i_s),$$

and $\Delta_{\mathbf{j},\mathbf{j}'}^r u := T_{\mathbf{j}'}^r u - T_{\mathbf{j}}^r u$. If $R = \{q_1, \dots, q_r\} \subset S := \{1, \dots s\}$, we set

$$T_{\mathbf{j}}^R u := T_{\mathbf{j}}^{q_1} \cdots T_{\mathbf{j}}^{q_r} u \quad \text{and} \quad \Delta_{\mathbf{j},\mathbf{j}'}^R u := \Delta_{\mathbf{j},\mathbf{j}'}^{q_1} \cdots \Delta_{\mathbf{j},\mathbf{j}'}^{q_r} u.$$

We put $T_{\mathbf{j}} u := T_{\mathbf{j}}^S u$ and $\Delta_{\mathbf{j},\mathbf{j}'} u := \Delta_{\mathbf{j},\mathbf{j}'}^S u$. For $\mathbf{j} = (j_1, \dots, j_s)$, we denote $\mathbf{j}+ := (j_1 + 1, \dots, j_s + 1)$ and $\Delta_{\mathbf{j}} u := \Delta_{\mathbf{j},\mathbf{j}+} u$. The *variation* of u is defined by

$$V^s(u) := \sum_{\mathbf{j} \in \mathbb{N}^{\star s}} |\Delta_{\mathbf{j}} u|,$$

and the *lower variation* of u is the sum

$$V_*(u) := \sum_{r=1}^{s} \sum_{\substack{R \subset S \\ \#R = r}} V^r(T_{\mathbf{1}}^{R^c} u),$$

where R^c denotes the complement of R in S. If u may be extended to infinity, we define the *upper variation* of u as

$$V^*(u) := \sum_{r=1}^{s} \sum_{\substack{R\subset S \\ \#R=r}} V^r(T_\infty^{R^c} u).$$

One can prove that if u has a bounded lower variation, then u may be extended to infinity and has a bounded upper variation. The next lemma is a version of the classical Koksma-Hlawka inequality. The proof follows the general outline of the proof of the Koksma-Hlawka inequality given by Zaremba [18].

Lemma 2. *Suppose g is a nonnegative sequence such that $\sum_{\mathbf{i}\in\mathbb{N}^{\star s}} g(\mathbf{i}) = 1$. If u is a sequence of bounded lower variation and if J is a point set consisting of $\mathbf{j}_0, \dots, \mathbf{j}_{N-1} \in \mathbb{N}^{\star s}$, then*

$$\left| \frac{1}{N} \sum_{0\le k<N} u(\mathbf{j}_k) - \sum_{\mathbf{i}\in\mathbb{N}^{\star s}} u(\mathbf{i}) g(\mathbf{i}) \right| \le V^*(u) D_N^\star(J; g).$$

The following lemma is an analogue of a result previously given in the continuous case [10] and can be proved with similar arguments.

Lemma 3. *Let u be a sequence of bounded lower variation. For $1 \le r \le s$, let p_r and $1 \le k_{0,r} \le k_{1,r} \le \cdots \le k_{p_r,r}$ be integers. For $\mathbf{n} = (n_1, \dots, n_s)$ with integers $n_r,\ 0 \le n_r < p_r$, let*

$$F_{\mathbf{n}} := \{k_{n_1,1}, \dots, k_{n_1+1,1}\} \times \cdots \times \{k_{n_s,s}, \dots, k_{n_s+1,s}\}$$

and $\mathbf{i}_\mathbf{n}, \mathbf{j}_\mathbf{n} \in F_\mathbf{n}$. Then

$$\sum_{\mathbf{n}} |u(\mathbf{j}_\mathbf{n}) - u(\mathbf{i}_\mathbf{n})| \le V_*(u) \prod_{r=1}^{s} p_r \sum_{r=1}^{s} \frac{1}{p_r}.$$

We now go back to the convergence analysis of the QMC algorithm. We define the error at time t_n as the star f_n-discrepancy of the point set J^n. For $h \in \mathbb{N}^\star$ let σ_h denote the sequence σ_{E_h}. We introduce the *truncation error*

$$\varepsilon_h^n := \frac{1}{\Delta t} \sum_{i\ge1} \left(f_{n+1}(i) - f_n(i)\right) \sigma_h(i) - \sum_{i,j\ge1} \widetilde{K}(i,j) f_n(i) f_n(j) \left(\sigma_h(i+j) - \sigma_h(i)\right),$$

the *additional error*

$$\begin{aligned} e_h^n := &\sum_{i,j\ge1} \widetilde{K}(i,j) f^n(i) f^n(j) \left(\sigma_h(i+j) - \sigma_h(i)\right) \\ &- \sum_{i,j\ge1} \widetilde{K}(i,j) f_n(i) f_n(j) \left(\sigma_h(i+j) - \sigma_h(i)\right), \end{aligned}$$

and the *integration error*

$$d_h^n := \frac{1}{N} \sum_{nN\le p<(n+1)N} C_h^{n+1}(\mathbf{x}_p) - \int_{I^3} C_h^{n+1}(\mathbf{x}) d\mathbf{x},$$

where $C_h^{n+1} := C_{E_h}^{n+1}$; see (8). We have the recurrence formula

$$D_N(E_h, J^{n+1}; f_{n+1}) = D_N(E_h, J^n; f_n) - \Delta t \varepsilon_h^n + \Delta t e_h^n + d_h^n. \tag{11}$$

The truncation error is bounded as follows:

$$|\varepsilon_h^n| \le \sum_{i\ge 1} \int_{t_n}^{t_{n+1}} \left| \frac{\partial^2 f}{\partial t^2}(i,t) \right| dt. \tag{12}$$

Next we have an upper bound for the additional error.

Lemma 4. *If for every $i, j \ge 1$ the sequences $\widetilde{K}(i,\cdot)$ and $\widetilde{K}(\cdot,j)$ are of bounded variation, then*

$$|e_h^n| \le \Big(\sup_{i\ge 1} V(\widetilde{K}(i,\cdot)) + \sup_{j\ge 1} V(\widetilde{K}(\cdot,j)) + 3\widetilde{K}^* \Big) D_N^\star(J^n; f_n).$$

Proof. Let u_h denote the array

$$u_h(i,j) := \widetilde{K}(i,j)\left(\sigma_h(i+j) - \sigma_h(i)\right), \quad i,j \in \mathbb{N}^\star.$$

Then

$$\begin{aligned} e_h^n = \frac{1}{N} \sum_{0\le k<N} \left(\frac{1}{N} \sum_{0\le \ell<N} u_h(j_{k,n}, j_{\ell,n}) - \sum_{j\ge 1} u_h(j_{k,n}, j) f_n(j) \right) \\ + \sum_{j\ge 1} \left(\frac{1}{N} \sum_{0\le k<N} u_h(j_{k,n}, j) - \sum_{i\ge 1} u_h(i,j) f_n(i) \right) f_n(j), \end{aligned}$$

and the result follows from Lemma 2. □

Finally, we need to bound the integration error.

Lemma 5. *If the sequence $\widetilde{K}$ is of bounded lower variation, then*

$$|d_h^n| \le \Big(2 + \Delta t \big(4V_*(\widetilde{K}) + 3\widetilde{K}^* \big) \Big) \frac{1}{b^{\lfloor (m-t)/3 \rfloor}}.$$

Proof. The integration error may be written as follows.

$$d_h^n = D_N(Q_{h,0}^n, X^n) - D_N(Q_{h,1}^n, X^n) + D_N(Q_{h,2}^n, X^n),$$

where

$$Q_{h,0}^n := \bigcup_{\substack{0\le k,\ell<N \\ j_{k,n}<h}} R_{k,\ell} \times I,$$

$$Q_{h,1}^n := \bigcup_{\substack{0\le k,\ell<N \\ j_{k,n}<h}} R_{k,\ell} \times I_{k,\ell}^n \quad \text{and} \quad Q_{h,2}^n := \bigcup_{\substack{0\le k,\ell<N \\ j_{k,n}+j_{\ell,n}<h}} R_{k,\ell} \times I_{k,\ell}^n.$$

We have

$$D_N(Q_{h,0}^n, X^n) = D_N(\pi' Q_{h,0}^n, \pi' X^n) = 0,$$

since $\pi' Q_{h,0}^n$ is a disjoint union of elementary intervals in base b, with measure b^{-m} and $\pi' X^n$ is a $(0,m,2)$-net in base b. Let $\kappa_{h,\alpha}^n$ for $\alpha = 1, 2$ denote the functions

$$\kappa_{h,1}^n(x_1,x_2) := \sum_{0\le k,\ell<N} c_{k,\ell}(x_1,x_2) \widetilde{K}(j_{k,n}, j_{\ell,n}) \sigma_h(j_{k,n}),$$

$$\kappa_{h,2}^n(x_1,x_2) := \sum_{0\le k,\ell<N} c_{k,\ell}(x_1,x_2) \widetilde{K}(j_{k,n}, j_{\ell,n}) \sigma_h(j_{k,n} + j_{\ell,n}),$$

for $(x_1, x_2) \in I^2$. Then,

$$Q^n_{h,\alpha} = \left\{ \mathbf{x} \in I^3 \ : \ x_3 < \Delta t \kappa^n_{h,\alpha}(x_1, x_2) \right\}, \quad \text{for } \alpha = 1, 2.$$

Let d_1, d_2 be integers such that $d_1 + d_2 \leq m - t$. For $\mathbf{a} = (a_1, a_2)$ with integers $0 \leq a_1 < b^{d_1}, 0 \leq a_2 < b^{d_2}$ we put

$$R_{\mathbf{a}} := \left[\frac{a_1}{b^{d_1}}, \frac{a_1 + 1}{b^{d_1}} \right) \times \left[\frac{a_2}{b^{d_2}}, \frac{a_2 + 1}{b^{d_2}} \right),$$

and, for $\alpha = 1, 2$,

$$\underline{Q}^n_{h,\alpha} := \bigcup_{\mathbf{a}} R_{\mathbf{a}} \times \Big[0, \Delta t \inf_{R_{\mathbf{a}}} \kappa^n_{h,\alpha} \Big), \quad \overline{Q}^n_{h,\alpha} := \bigcup_{\mathbf{a}} R_{\mathbf{a}} \times \Big[0, \Delta t \sup_{R_{\mathbf{a}}} \kappa^n_{h,\alpha} \Big),$$

$$\partial Q^n_{h,\alpha} = \bigcup_{\mathbf{a}} R_{\mathbf{a}} \times \Big[\Delta t \inf_{R_{\mathbf{a}}} \kappa^n_{h,\alpha}, \Delta t \sup_{R_{\mathbf{a}}} \kappa^n_{h,\alpha} \Big],$$

where the unions are over all $\mathbf{a} = (a_1, a_2)$ with $0 \leq a_1 < b^{d_1}, 0 \leq a_2 < b^{d_2}$. We have

$$\underline{Q}^n_{h,\alpha} \subset Q^n_{h,\alpha} \subset \overline{Q}^n_{h,\alpha} \quad \text{and} \quad \overline{Q}^n_{h,\alpha} \setminus \underline{Q}^n_{h,\alpha} \subset \partial Q^n_{h,\alpha},$$

hence

$$D_N(\underline{Q}^n_{h,\alpha}, X^n) - \lambda_3(\partial Q^n_{h,\alpha}) \leq D_N(Q^n_{h,\alpha}, X^n) \leq D_N(\overline{Q}^n_{h,\alpha}, X^n) + \lambda_3(\partial Q^n_{h,\alpha}).$$

By Lemma 1 it follows that

$$\left| D_N(\underline{Q}^n_{h,\alpha}, X^n) \right| \leq b^{d_1+d_2+t-m} \quad \text{and} \quad \left| D_N(\overline{Q}^n_{h,\alpha}, X^n) \right| \leq b^{d_1+d_2+t-m}.$$

Besides,

$$\lambda_3(\partial Q^n_{h,\alpha}) = \frac{\Delta t}{b^{d_1+d_2}} \sum_{\mathbf{a}} \Big(\sup_{R_{\mathbf{a}}} \kappa^n_{h,\alpha} - \inf_{R_{\mathbf{a}}} \kappa^n_{h,\alpha} \Big).$$

Let $u_{h,\alpha}$ for $\alpha = 1, 2$ denote the arrays

$$u_{h,1}(i,j) := \widetilde{K}(i,j)\sigma_h(i), \quad u_{h,2}(i,j) := \widetilde{K}(i,j)\sigma_h(i+j), \quad i, j \in \mathbb{N}^\star,$$

so that, for $\alpha = 1, 2$,

$$\kappa^n_{h,\alpha}(x_1, x_2) = u_{h,\alpha}(j_{k(x_1),n}, j_{k(x_2),n}).$$

Let $F^n_{\mathbf{a}}$ be the point set

$$F^n_{\mathbf{a}} := \{ j_{a_1 b^{m-d_1},n}, \ldots, j_{(a_1+1)b^{m-d_1}-1,n} \} \times \{ j_{a_2 b^{m-d_2},n}, \ldots, j_{(a_2+1)b^{m-d_2}-1,n} \}.$$

Since the particles are labeled according to mass,

$$(x_1, x_2) \in R_{\mathbf{a}} \Rightarrow (j_{k(x_1),n}, j_{k(x_2),n}) \in F^n_{\mathbf{a}},$$

and consequently

$$\sup_{R_{\mathbf{a}}} \kappa^n_{h,\alpha} - \inf_{R_{\mathbf{a}}} \kappa^n_{h,\alpha} \leq \max_{F^n_{\mathbf{a}}} u_{h,\alpha} - \min_{F^n_{\mathbf{a}}} u_{h,\alpha}.$$

Applications of Lemma 3 yield

$$\sum_{\mathbf{a}} \Big(\max_{F^n_{\mathbf{a}}} u_{h,1} - \min_{F^n_{\mathbf{a}}} u_{h,1} \Big) \leq V_*(\widetilde{K}) \left(b^{d_1} + b^{d_2} \right) + \widetilde{K}^* b^{d_2},$$

$$\sum_{\mathbf{a}} \Big(\max_{F^n_{\mathbf{a}}} u_{h,2} - \min_{F^n_{\mathbf{a}}} u_{h,2} \Big) \leq V_*(\widetilde{K}) \left(b^{d_1} + b^{d_2} \right) + \widetilde{K}^* \left(b^{d_1} + b^{d_2} - 1 \right).$$

By choosing $d_1 = d_2 = \lfloor (m-t)/3 \rfloor$, we obtain the desired result. □

We can combine the previous bounds to derive an upper bound for the error at time t_n.

Proposition. *If the sequence $\widetilde{K}$ is of bounded lower variation, then*

$$D_N^\star(J^n; f_n) \le \mathrm{e}^{ct_n} D_N^\star(J^0; f_0) + \Delta t \sum_{i\ge 1} \int_0^{t_n} \mathrm{e}^{c(t_n-t)} \left| \frac{\partial^2 f}{\partial t^2}(i,t) \right| dt + \Big(4V_*(\widetilde{K}) + 3\widetilde{K}^* + \frac{2}{\Delta t} \Big) \frac{\mathrm{e}^{ct_n}}{cb^{\lfloor (m-t)/3 \rfloor}},$$

where

$$c := \sup_{i\ge 1} V(\widetilde{K}(i,\cdot)) + \sup_{j\ge 1} V(\widetilde{K}(\cdot,j)) + 3\widetilde{K}^*.$$

Proof. The result follows from recurrence formula (11), used in conjunction with inequality (12) and the bounds in Lemmas 4 and 5. □

Remark. The proposition indicates a convergence order of $\mathcal{O}(1/N^{1/3})$, which is worse than the $\mathcal{O}(1/N^{1/2})$ of MC schemes. The examples below show a QMC method's convergence rate of $\mathcal{O}(1/N^{2/3})$ which is better than the random convergence. In addition the upper bound in the proposition grows linearly with the number of time steps, but this growth is not observed in computational experiments.

3. Numerical examples

The purpose of this section is to numerically validate the QMC algorithm described above. In our experiments, the number base b is taken to be 3 and the low-discrepancy sequence X is a $(0,3)$-sequence in base 3 constructed by Faure [6]. It has long been recognized that three particular kernels $K(i,j)$ are mathematically tractable [1]: for a monodisperse initial configuration $f_0 = \delta_1$, explicit solutions of Smoluchowski's coagulation equation are available. In the following we restrict our consideration to the kernels $K(i,j) = 1$ and $K(i,j) = i + j$ (note that the latter does not satisfy the hypothesis of the proposition). In both cases we can compute the error $D_N^\star(J^n; f_n)$ of the algorithm and we can compare it with the error given by the Monte Carlo scheme proposed in [2]. Assuming that the methods produce approximately $\mathcal{O}(1/N^q) + \mathcal{O}((\Delta t)^r)$ errors, one can estimate the exponents q and r from plots of the error versus N or Δt. If

$$C(t) := \sum_{i\ge 1} c(i,t) = \sum_{i\ge 1} \frac{f(i,t)}{i},$$

then $N_0 C(t)$ is the total number of clusters at time t. At time t_n, $C(t_n)$ is approximated according to (6) as

$$C(t_n) \approx \frac{1}{N} \sum_{0\le k<N} \frac{1}{j_{k,n}}.$$

We can study the accuracy of this approximation.

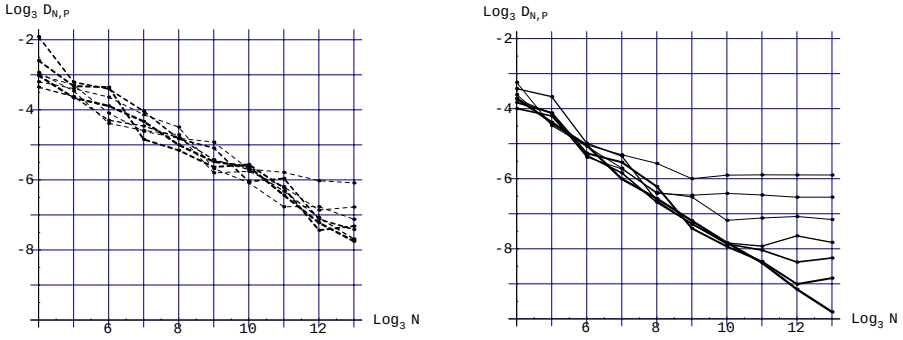

FIGURE 1. $K(i,j) = 1$. MC (left) vs. QMC simulations (right) for $P = 100$, 200, 400, 800, 1,600, 3,200, and 6,400 time steps. Thick lines correspond to small time steps.

3.1. $K(i,j) = 1$. In this case, the exact solution of (3) with a monodisperse initial condition is

$$f(i,t) = \frac{4i}{(t+2)^2}\left(\frac{t}{t+2}\right)^{i-1}. \tag{13}$$

We compute the solution up to time $t = 1.0$ with N particles and P time steps. The numerical parameter N varies between 3^4 and 3^{13}, and P varies between 1×100 and $2^6 \times 100$. In order to reduce scatter, the error is averaged as

$$D_{N,P} := \frac{1}{100}\sum_{m=1}^{100} D_N^\star(J^{mp}; f_{mp}),$$

where $p = P/100$. Figure 1 is a plot of the error $D_{N,P}$ as a function of N on a log-log scale. We see that the QMC method achieves a better rate of convergence

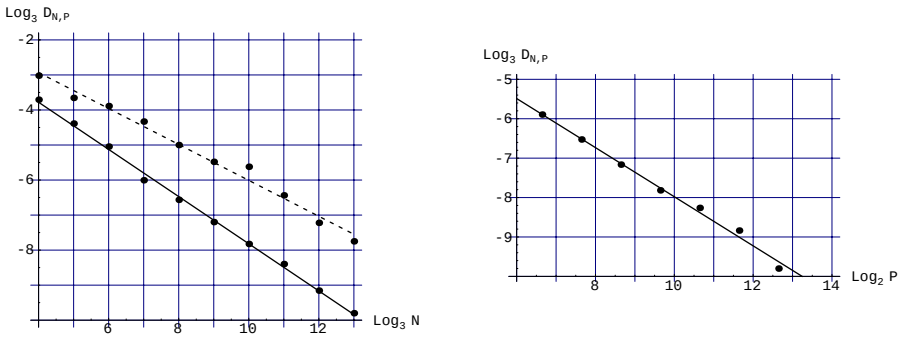

FIGURE 2. $K(i,j) = 1$. Left: linear fits to the error as a function of N for $P = 6,400$ time steps. Comparison of MC (dashed) and QMC (solid) simulations. Right: linear fit to the error as a function of P for QMC simulations with $N = 3^{13}$ particles.

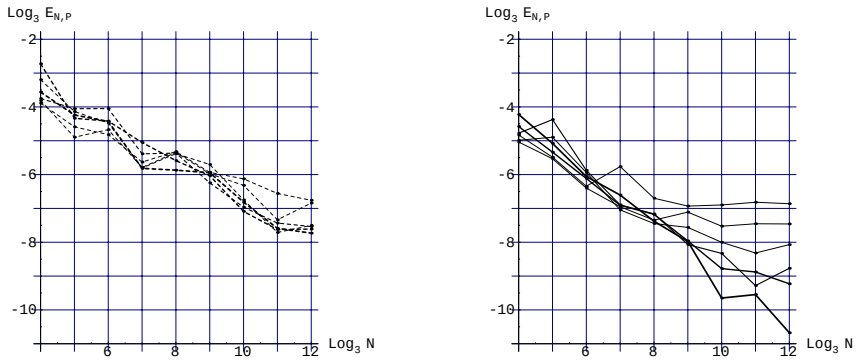

FIGURE 3. $K(i,j) = 1$. MC (left) vs. QMC simulations (right) for $P = 1,000$, $2,000$, $4,000$, $8,000$, $16,000$, and $32,000$ time steps. Thick lines correspond to small time steps.

of the error to zero with increase of N; this more rapid convergence is only tangible for small Δt.

Next we choose a time step small enough so that time discretization error is insignificant relative to the Monte Carlo error. The results are displayed in Figure 2-left. The best (in the sense of least squares) straight line fit to the log-log plot of the data gives for $P = 6,400$:

$$D_{N,P}(\text{MC}) \approx \frac{0.38}{N^{0.51}} \quad \text{and} \quad D_{N,P}(\text{QMC}) \approx \frac{0.30}{N^{0.67}}. \tag{14}$$

If we choose a large number of particles, the quasi–Monte Carlo error is negligible relative to time discretization error. So we can estimate the exponent r by a linear fit to plots of $\log D_{N,P}$ versus $\log P$. The results are shown in Figure 2-right: for $N = 3^{13}$ one obtains

$$D_{N,P}(\text{QMC}) \approx 0.15 \cdot (\Delta t)^{0.99}. \tag{15}$$

In this case one has

$$C(t) = \frac{2}{t+2}.$$

We compute the solution up to time $t = 10.0$ with N particles and P time steps. Let $C_{N,P}(t_n)$ denote the approximation of $C(t_n)$ for $0 < n \le P$. Here N varies between 3^4 and 3^{12}, and P varies between $1 \times 1,000$ and $2^5 \times 1,000$. Figure 3 shows the curves for the discrete L^1 norm

$$E_{N,P} := \frac{1}{100} \sum_{m=1}^{100} |C(t_{mp}) - C_{N,P}(t_{mp})|.$$

We see that QMC outperforms standard MC for this example.

3.2. $K(i,j) = i + j$. Here the exact solution of (3) with a monodisperse initial condition is

$$f(i,t) = \frac{i^i}{i!}(1 - \mathrm{e}^{-t})^{i-1}\mathrm{e}^{-i(1-\mathrm{e}^{-t})-t}. \tag{16}$$

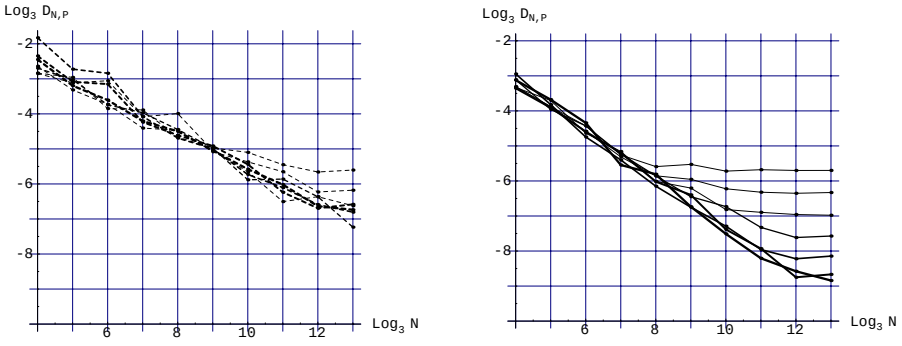

FIGURE 4. $K(i,j) = i+j$. MC (left) vs. QMC simulations (right) with $P = 200,\ 400,\ 800,\ 1,600,\ 3,200,\ 6,400$, and $12,800$ time steps. Thick lines correspond to small time steps.

The solution is computed up to time $t = 1.0$ with N particles and P time steps. The number N varies between 3^4 and 3^{13}, and P varies between 2×100 and $2^7 \times 100$. The error is averaged as in the previous model problem. The results are similar to those for the first example: we see in Figure 4 that the QMC method is superior to the MC scheme.

The gain in the rate of convergence q is illustrated in Figure 5-left. As in the previous example, if the time step size is chosen small enough, the log-log scale allows the plot of an error of the form c/N^q to appear as a straight line with slope $-q$: for $P = 12,800$, one has

$$D_{N,P}(\mathrm{MC}) \approx \frac{0.46}{N^{0.48}} \quad \text{and} \quad D_{N,P}(\mathrm{QMC}) \approx \frac{0.60}{N^{0.68}}. \tag{17}$$

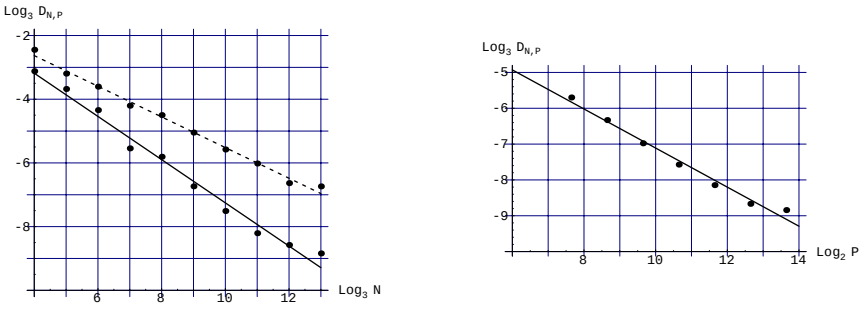

FIGURE 5. $K(i,j) = i + j$. Left: linear fits to the error as a function of N for $P = 12,800$ time steps. Comparison of MC (dashed) and QMC (solid) simulations. Right: linear fit to the error as a function of P for QMC simulations with $N = 3^{13}$ particles.

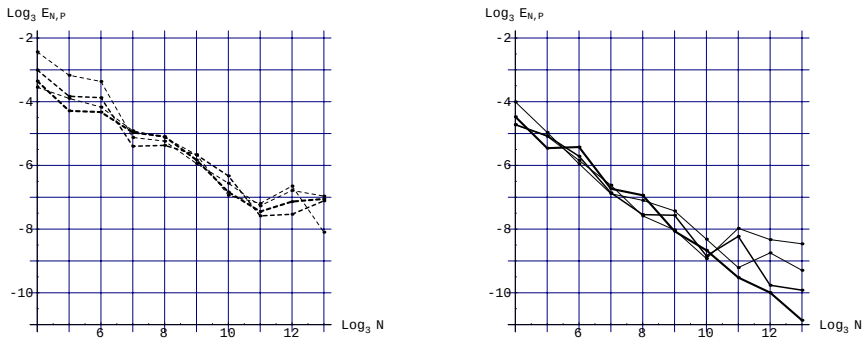

FIGURE 6. $K(i,j) = i + j$. MC (left) vs. QMC simulations (right) with $P = 3,200$, $6,400$, $12,800$, and $25,600$ time steps. Thick lines correspond to small time steps.

For a large number of particles ($N = 3^{13}$), the least-squares fit convergence rate of the QMC method is estimated as

$$D_{N,P}(\text{QMC}) \approx 0.16 \cdot (\Delta t)^{0.86}, \tag{18}$$

as shown in Figure 5-right.

In this case one obtains

$$C(t) = \mathrm{e}^{-t}.$$

We compute the solution up to time $t = 2.0$ with N particles and P time steps: N varies between 3^4 and 3^{13}, and P varies between $2^5 \times 100$ and $2^8 \times 100$. In Figure 6 we graph the $E_{N,P}$ obtained with MC and QMC strategies: once again QMC produces more accurate results than MC.

CONCLUSION

We have analyzed a procedure for solving Smoluchowski's coagulation equation. The approach is to use the Monte Carlo method to simulate the aggregation of clusters. A sample of test particles is chosen and it is assumed that their behavior is an indicator of the behavior of the medium as a whole. Time is discretized and since we approximate the mass density, the scheme works with a fixed particle number N. The standard Monte Carlo method can be quite slow, because its convergence rate is only $\mathcal{O}(1/N^{1/2})$. We have considered an improvement to this method by using quasi-random numbers in the implementation of the algorithm. To make optimal use of the greater uniformity associated with quasi-random sequences we reorder the particles at each time step. Convergence of the simulation as the number N increases has been proved. We test our analysis by comparing the QMC results with two known analytic solutions to the Smoluchowski equation. In both comparisons, the QMC results have been found to reproduce the expected distributions. Moreover, the numerical experiments show that the error in the QMC simulations is significantly less than the corresponding error for a standard MC simulation.

References

1. D.J. Aldous, *Deterministic and stochastic models for coalescence (aggregation and coagulation): a review of the mean-field theory for probabilists*, Bernoulli **5** (1999), 3–48. MR **2001c:**60153
2. H. Babovsky, *On a Monte Carlo scheme for Smoluchowski's coagulation equation*, Monte Carlo Methods Appl. **5** (1999), 1–18. MR **99m:**65010
3. Y.R. Domilovskiy, A.A. Lushnikov, and V.N. Piskurov, *Monte Carlo simulation of coagulation processes*, Izv. Atmos. Ocean. Phys. **15** (1979), 129–134. MR **81e:**86016
4. A. Eibeck and W. Wagner, *An efficient stochastic algorithm for studying coagulation dynamics and gelation phenomena*, SIAM J. Sci. Comput. **22** (2000), 802–821. MR **2001e:**82037
5. K.-T. Fang, F.J. Hickernell, and H. Niederreiter (Eds.), *Monte Carlo and Quasi–Monte Carlo Methods* 2000, Springer, Berlin, 2002.
6. H. Faure, *Discrépance de suites associées à un système de numération (en dimension s)*, Acta Arith. **41** (1982), 337–351. MR **84m:**10050
7. A.L. Garcia, C. Van den Broeck, M. Aertsens, and R. Serneels, *A Monte Carlo simulation of coagulation*, Phys. A **143** (1987), 535–546.
8. D.N. Gillespie, *An exact method for numerically simulating the stochastic coalescence process in a cloud*, J. Atmospheric Sci. **32** (1975), 1977–1989.
9. O.J. Heilmann, *Analytical solutions of Smoluschowski's coagulation equation*, J. Phys. A **25** (1992), 3763–3771. MR **93c:**82068
10. C. Lécot, *Error bounds for quasi–Monte Carlo integration with nets*, Math. Comp. **65** (1996), 179–187. MR **96d:**65012
11. K. Liffman, *A direct simulation Monte Carlo method for cluster coagulation*, J. Comput. Phys. **100** (1992), 116–127.
12. H. Niederreiter, *Point sets and sequences with small discrepancy*, Monatsh. Math. **104** (1987), 273–337. MR **89c:**11120
13. H. Niederreiter, *Random Number Generation and Quasi–Monte Carlo Methods*, SIAM, Philadelphia, 1992. MR **93h:**65008
14. K.K. Sabelfeld, S.V. Rogasinsky, A.A. Kolodko, and A.I. Levykin, *Stochastic algorithms for solving Smolouchovsky coagulation equation and applications to aerosol growth simulation*, Monte Carlo Methods Appl. **2** (1996), 41–87. MR **97c:**82042
15. M. v. Smoluchowski, *Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen*, Z. Phys. Chem. **92** (1916), 129–168.
16. M. v. Smoluchowski, *Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen*, Phys. Z. **17** (1916), 557–571, 585–599.
17. J.L. Spouge, *Monte Carlo results for random coagulation*, J. Colloid Interface Sci. **107** (1985), 38–43.
18. S.K. Zaremba, *Some applications of multidimensional integration by parts*, Ann. Polon. Math. **21** (1968), 85-96. MR **38:**4034

Laboratoire de Mathématiques, Université de Savoie, Campus scientifique, 73376 Le Bourget-du-Lac cedex, France

E-mail address: Christian.Lecot@univ-savoie.fr

Weierstrass-Institut für Angewandte Analysis und Stochastik, Mohrenstrasse 39, 10117 Berlin, Germany

E-mail address: wagner@wias-berlin.de