Quasi-Monte Carlo methods for Markov chains with continuous multi-dimensional state space

Rami El Haddad, Christian Lécot, Pierre L'Ecuyer, Nabil Nassif

Quasi-Monte Carlo methods for Markov chains with continuous multi-dimensional state space

R. El Haddad[a,*], C. Lécot[b], P. L'Ecuyer[c], N. Nassif[d]

[a]*Département de Mathématiques, Université Saint-Joseph, BP 11-514, Riad El Solh Beyrouth 1107 2050, Liban*

[b]*LAMA, UMR 5127 CNRS & Université de Savoie, 73376 Le Bourget du Lac, France*

[c]*Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, CP 6128, Succ. Centre-Ville, Montréal, H3C 3J7, Canada*

[d]*Department of Mathematics, American University of Beirut, BP 11-0236, Riad El-Solh Beyrouth 1107 2020, Liban*

Abstract

We describe a quasi-Monte Carlo method for the simulation of discrete time Markov chains with continuous multi-dimensional state space. The method simulates copies of the chain in parallel. At each step the copies are reordered according to their successive coordinates. We prove the convergence of the method when the number of copies increases. We illustrate the method with numerical examples where the simulation accuracy is improved by large factors compared with Monte Carlo simulation.

Keywords: Markov chain, discrepancy, quasi-Monte Carlo method, simulation.

1. Introduction

Many real-life systems can be modeled using Markov chains. Fields of application are queueing theory, telecommunications, option pricing, etc. In most interesting situations, analytic formulas are not available and the state space of the chain is so large that classical numerical methods would require

[*] Corresponding author. Tel.: +961 (1) 421 372; fax: +961 (4) 532 657

Email addresses: rami.haddad@fs.usj.edu.lb (R. El Haddad), christian.lecot@univ-savoie.fr (C. Lécot), lecuyer@iro.umontreal.ca (P. L'Ecuyer), nn12@aub.edu.lb (N. Nassif)

a considerable computational time and huge memory capacity. So Monte Carlo (MC) simulation becomes the standard way of estimating performance measures for these systems. A drawback of MC methods is their slow convergence. One approach to improve the accuracy of the method is to change the random numbers used. Quasi-Monte Carlo (QMC) methods use quasi-random numbers instead of pseudo-random numbers. Pseudo-random numbers aim to simulate a sequence of independent and identically distributed (i.i.d.) random variables with a given distribution (we only consider the uniform distribution). In the example of MC integration, it is not so much the randomness of the samples that is relevant, but rather that the samples should be spread in a uniform manner over the integration domain. Quasi-random numbers are sample points for which the empirical distribution is close to the uniform distribution; unlike for random sampling, quasi-random points are not required to be independent and may be completely deterministic.

The efficiency of a QMC method depends on the quality of the quasi-random points that are used. Broadly speaking, these points should form a low-discrepancy point set. We recall from [12] some basic notations and concepts. We first denote $\mathcal{I} := [0, 1)$. Let $s \geq 1$ be a fixed dimension and denote by λ_s the s-dimensional Lebesgue measure. For a set $U = \{\mathbf{u}_0, \ldots, \mathbf{u}_{N-1}\}$ of points in the s-dimensional unit cube $\mathcal{I}^s$ and for a Borel set $B \subset \mathcal{I}^s$ we define the *local discrepancy* by

$$D(B, U) := \frac{1}{N} \sum_{0 \leq k < N} 1_B(\mathbf{u}_k) - \lambda_s(B), \tag{1}$$

where 1_B denotes the indicator function of B. The *discrepancy* of U is defined by $D(U) := \sup_Q |D(Q, U)|$, the supremum being taken over all subintervals $Q \subset \mathcal{I}^s$. The *star discrepancy* of U is $D^\star(U) := \sup_{Q^\star} |D(Q^\star, U)|$, where $Q^\star$ runs through all subintervals of $\mathcal{I}^s$ of the form $\prod_{i=1}^s [0, a_i)$. A *low-discrepancy* point set in $\mathcal{I}^s$ is a set of N points for which the discrepancy is of size $\mathcal{O}((\log N)^{s-1}/N)$, which is the minimum size possible. The most powerful current methods of constructing low-discrepancy point sets are based on the theory of (t, m, s)-nets. For an integer $b \geq 2$, an *elementary interval in base* b is an interval of the form $\prod_{i=1}^s [a_i b^{-d_i}, (a_i + 1) b^{-d_i})$, with integers $d_i \geq 0$ and $0 \leq a_i < b^{d_i}$ for $1 \leq i \leq s$. If $0 \leq t \leq m$ are integers, a (t, m, s)*-net in base* b is a point set U consisting of b^m points in $\mathcal{I}^s$ such that $D(Q, U) = 0$ for every elementary interval Q in base b with measure b^{t-m}. If $b \geq 2$ and $t \geq 0$

are integers, a sequence $\mathbf{u}_0, \mathbf{u}_1, \ldots$ of points in $\mathcal{I}^s$ is a (t, s)-*sequence in base* b if, for all integers $j \geq 0$ and $m > t$, the points $\mathbf{u}_\ell$ with $jb^m \leq \ell < (j+1)b^m$ form a (t, m, s)-net in base b.

In the example of numerical integration, the QMC method achieves a significantly higher accuracy than the MC method, with the same computational effort. It may be hoped that the improvement obtained by using quasi-random points in place of random samples can also be attained in problems of numerical analysis that can be reduced to numerical integration. QMC simulations can outperform MC simulations in some applications: we refer to the IMACS Seminars on Monte Carlo Methods [1, 2, 4, 13].

In previous communications, we first proposed QMC schemes to simulate Markov chains with a *discrete* state space, either one-dimensional [7, 8] or multi-dimensional [3]. We next applied the method to one-dimensional continuous state spaces [10, 11]. In the present work, we extend the QMC algorithm to Markov chains with *continuous multi-dimensional* state spaces.

2. The method

Our setting is an homogeneous Markov chain $\{X_j,\ j \in \mathbb{N}\}$ whose state space E is a subspace of $\mathbb{R}^s$ for some $s \in \mathbb{N}^*$. The distribution P_0 of X_0 is known, and we assume that the chain evolves according to the stochastic recurrence:

$$X_{j+1} = \varphi_{j+1}(X_j, U_{j+1}), \quad j \geq 0, \tag{2}$$

where $\{U_j,\ j \geq 1\}$ is a sequence of i.i.d. uniform random variables over $\mathcal{I}^d$ for some $d \in \mathbb{N}^*$, and $\varphi_{j+1} : E \times \mathcal{I}^d \to E$ is a measurable map for each j.

To approximate the Markov chain by ordinary MC, we proceed as follows. Given a large integer N, we draw N samples $\mathbf{x}_k^0,\ 0 \leq k < N$ from the initial distribution P_0. Then for each k, we generate a sample path of the chain via

$$\mathbf{x}_k^{j+1} = \varphi_{j+1}(\mathbf{x}_k^j, \mathbf{u}_k^{j+1}), \quad j \geq 0, \tag{3}$$

where $\mathbf{u}_k^1, \mathbf{u}_k^2, \ldots$ are pseudo-random numbers which simulate independent and uniformly distributed random variables over $\mathcal{I}^d$. In order to construct a QMC algorithm for the approximation of the Markov chain, we reduce the simulation to numerical integration.

We denote by $\mathcal{M}^+$ the set of all nonnegative measurable functions on E. If P_j denotes the distribution of X_j, then

$$\forall f \in \mathcal{M}^+ \quad \int_E f dP_{j+1} = \int_{\mathcal{I}^d} \int_E f \circ \varphi_{j+1}(\mathbf{x}, \mathbf{u}) dP_j(\mathbf{x}) d\mathbf{u}. \tag{4}$$

For $\mathbf{x} \in E$, let us write $\delta_{\mathbf{x}}$ for the unit mass at $\mathbf{x}$. We are looking for an approximation of P_j of the form

$$\widehat{P}_j := \frac{1}{N} \sum_{0 \leq k < N} \delta_{\mathbf{x}_k^j}, \tag{5}$$

for some integer N and a judiciously chosen set $X^j := \{\mathbf{x}_0^j, \ldots, \mathbf{x}_{N-1}^j\} \subset E$. Let $b \geq 2$, $d_1, \ldots, d_s$ be integers and put $N := b^m$ where $m = \sum_{i=1}^s d_i$. We shall use a low-discrepancy sequence $Y = \{\mathbf{y}_0, \mathbf{y}_1, \ldots\} \subset \mathcal{I}^{s+d}$ for QMC approximation. If Y^j is the point set $\{\mathbf{y}_\ell : jN \leq \ell < (j+1)N\}$ and if π' and π'' are the projections defined by $\pi'(u_1, \ldots, u_{s+d}) := (u_1, \ldots, u_s)$ and $\pi''(u_1, \ldots, u_{s+d}) := (u_{s+1}, \ldots, u_{s+d})$, we assume that

$$\forall j \in \mathbb{N} \quad \pi' Y^j \text{ is a } (0, m, s)\text{-net in base } b \quad \text{and} \quad \pi''(Y^j) \subset \mathring{\mathcal{I}}^d, \tag{6}$$

where $\mathring{\mathcal{I}} := (0, 1)$. For $\mathbf{u} \in \mathcal{I}^{s+d}$, we denote $\mathbf{u}' := \pi'(\mathbf{u})$ and $\mathbf{u}'' := \pi''(\mathbf{u})$. We now explain our algorithm in which N copies of the chain are simulated simultaneously.

2.1. Generating the initial states

A sample X^0 is chosen such that $\widehat{P}_0 \approx P_0$.This means that X^0 has a small star P_0-discrepancy (see section 3).

2.2. Transition

Supposing that we have calculated a set X^j of N states such that $\widehat{P}_j \approx P_j$, we compute X^{j+1} and $\widehat{P}_{j+1}$ in two steps.

2.2.1. Relabeling the states

The states are labeled $\mathbf{x}_{\mathbf{a}}^j$ using a *multi-dimensional index* in $\mathcal{A} := \{\mathbf{a} = (a_1, \ldots, a_s),\ 0 \leq a_i < b^{d_i},\ 1 \leq i \leq s\}$, such that:

if $a_1 < a'_1$ then $x_{\mathbf{a},1}^j \leq x_{\mathbf{a}',1}^j$,

if $a_1 = a'_1, a_2 < a'_2$ then $x_{\mathbf{a},2}^j \leq x_{\mathbf{a}',2}^j$,

$\cdots$

if $a_1 = a'_1, ..., a_{s-1} = a'_{s-1}, a_s < a'_s$ then $x_{\mathbf{a},s}^j \leq x_{\mathbf{a}',s}^j$.

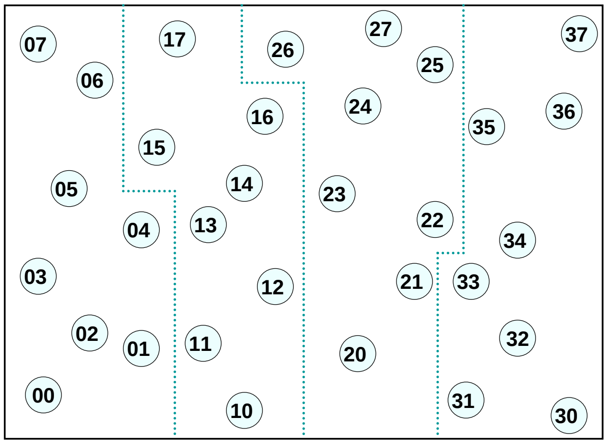

Figure 1: Relabeling the states ($b = 2,\ s = 2,\ d_1 = 2,\ d_2 = 3$)

Practically, we first sort the N states in b^{d_1} groups of size Nb^{-d_1} according to their first coordinates; then each group is sorted in b^{d_2} subgroups of size $Nb^{-d_1-d_2}$ by order of the second coordinates, and so on. The sorting is illustrated in Figure 1, with $b = 2$, $s = 2$, $d_1 = 2$ and $d_2 = 3$. Each circle represents a state and the numbers represent the pairs (a_1, a_2). This type of sorting was first introduced in [6]. It provides a good description of the distribution of the states in the state space and will guarantee theoretical convergence: since each transition can be described by a numerical integration (see Section 2.2.2 below), the sorting reverts to minimizing the amplitude of the jumps of the function to be integrated.

2.2.2. QMC integration

If we replace P_j by $\widehat{P}_j$ in the right-hand-side of (4), we define a probability measure $\widetilde{P}_{j+1}$ on E:

$$\int_E f d\widetilde{P}_{j+1} := \int_{\mathcal{I}^d} \int_E f \circ \varphi_{j+1}(\mathbf{x}, \mathbf{u}) d\widehat{P}_j(\mathbf{x}) d\mathbf{u}, \quad f \in \mathcal{M}^+. \tag{7}$$

This measure certainly approximates P_{j+1}, but it is not a sum of unit masses, like $\widehat{P}_j$. We recover this kind of approximation if we use a QMC quadrature rule. For $\mathbf{a} = (a_1, \ldots, a_s) \in \mathcal{A}$, let $\mathcal{I}_{\mathbf{a}} := \prod_{i=1}^{s} [a_i b^{-d_i}, (a_i + 1)b^{-d_i})$ and $1_{\mathbf{a}}$ be the indicator function of $\mathcal{I}_{\mathbf{a}}$. For $f \in \mathcal{M}^+$, define

$$C^j f(\mathbf{u}) := \sum_{\mathbf{a} \in \mathcal{A}} 1_{\mathbf{a}}(\mathbf{u}') f \circ \varphi_{j+1}(\mathbf{x}^j_{\mathbf{a}}, \mathbf{u}''), \quad \mathbf{u} = (\mathbf{u}', \mathbf{u}'') \in \mathcal{I}^{s+d}. \tag{8}$$

Then we have

$$\forall f \in \mathcal{M}^+ \quad \int_E f d\widetilde{P}_{j+1} = \int_{\mathcal{I}^{s+d}} C^j f(\mathbf{u}) d\mathbf{u}. \tag{9}$$

We retrieve $\widehat{P}_{j+1}$ if we perform a QMC approximation:

$$\int_E f d\widehat{P}_{j+1} := \frac{1}{N} \sum_{jN \leq \ell < (j+1)N} C^j f(\mathbf{y}_\ell), \quad f \in \mathcal{M}^+. \tag{10}$$

The last step of the algorithm may be summarized as follows. For each $\mathbf{u}' \in \mathcal{I}^s$, we associate the index $\mathbf{a}(\mathbf{u}') := (\lfloor b^{d_1} u_1 \rfloor, \ldots, \lfloor b^{d_s} u_s \rfloor)$. From (6), the mapping $k \in \{jN, jN+1, \ldots, (j+1)N-1\} \to \mathbf{a}(\mathbf{y}'_k) \in \mathcal{A}$ is one-to-one. The N states $\mathbf{x}^{j+1}_0, \ldots, \mathbf{x}^{j+1}_{N-1}$ are computed according to:

$$\mathbf{x}^{j+1}_{\mathbf{a}(\mathbf{y}'_\ell)} = \varphi_{j+1}(\mathbf{x}^j_{\mathbf{a}(\mathbf{y}'_\ell)}, \mathbf{y}''_\ell), \quad \text{for } jN \leq \ell < (j+1)N, \tag{11}$$

which must be compared with (3). This means that the projection $\pi'(\mathbf{y}_\ell)$ of each point $\mathbf{y}_\ell$ of the low discrepancy sequence is used to select the state of the chain which will advance, while the remaining components $\pi''(\mathbf{y}_\ell)$ are used to determine the next state.

3. Convergence

First we adapt the basic concepts of QMC methods to the present study. If $U = \{\mathbf{u}_0, \ldots, \mathbf{u}_{N-1}\} \subset \mathcal{I}^s$ and if $c : \mathcal{I}^s \to \mathbb{R}$ is a non-negative measurable and bounded function, we put

$$D(c, U) := \frac{1}{N} \sum_{0 \leq k < N} c(\mathbf{u}_k) - \int_{\mathcal{I}^s} c(\mathbf{u}) d\mathbf{u}. \tag{12}$$

Let now P be a probability measure on E and $X := \{\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}\} \subset E$. For a measurable subset A of E we define the *local P-discrepancy* by

$$D(A, X; P) := \frac{1}{N} \sum_{0 \leq k < N} \chi_A(\mathbf{x}_k) - P(A), \tag{13}$$

where χ_A denotes the characteristic function of A. The *star P-discrepancy* of the point set X is defined by $D^\star(X;P) := \sup_{\mathbf{z}\in E} |D(A_{\mathbf{z}}, X; P)|$, where $A_{\mathbf{z}} := \{\mathbf{x} \in E : \mathbf{x} < \mathbf{z}\}$ and $\mathbf{x} < \mathbf{z}$ means $\forall i \; x_i < z_i$. We shall also use the following notation: if $f \in \mathcal{M}^+$, then

$$D(f, X; P) := \frac{1}{N} \sum_{0\le k<N} f(\mathbf{x}_k) - \int_E f dP. \tag{14}$$

The next Lemma is a version of the classical Koksma inequality [12].

Lemma 1. *Let P be a probability measure on E, with a Riemann-integrable density function ρ. Let $f : E \to \mathbb{R}$ be a function such that f and $|f|$ are of bounded variation in the sense of Hardy and Krause. If f or ρ is continuous and if X is a point set consisting of N points in E, then*

$$|D(f, X; P)| \le V(f) D^\star(X; P). \tag{15}$$

We now go back to the convergence analysis of the QMC algorithm. We restrict ourselves to the case $s = d$ and we assume that $E = \prod_{i=1}^s E_i$ and every φ_j has the form: $\varphi_j(\mathbf{x}, \mathbf{u}'') = (\varphi_{j,1}(x_1, u_1''), \ldots, \varphi_{j,s}(x_s, u_s''))$. In addition, we assume that every P_j has a continuous density function ρ_j.

Proposition 1. *Suppose that*

(i) $\forall j \ge 1 \; \forall \mathbf{z} \in E \; \forall \mathbf{u}'' \in \mathcal{I}^s \quad V(\chi_{A_z} \circ \varphi_j(\cdot, \mathbf{u}'')) \le 1,$

and for every $j \ge 1$ and $1 \le i \le s$:

(ii) for any $x_i \in E_i$, the map $\varphi_{j,i}(x_i, \cdot) : \mathring{\mathcal{I}} \to E_i$ is strictly increasing,

(iii) for any $z_i \in E_i$, the map $x_i \to (\varphi_{j,i}(x_i, \cdot))^{-1}(z_i)$ is monotone.

Then

$$\begin{aligned} D^\star(X^J; P_J) \;\; \le \;\; & D^\star(X^0; P_0) + b^{d_1+\cdots+d_{s-1}+\lfloor d_s/2\rfloor} \sum_{j=0}^{J-1} D(Y^j) \\ & + \left(\frac{1}{b^{d_1}} + \cdots + \frac{1}{b^{d_{s-1}}} + \frac{1}{b^{\lfloor d_s/2\rfloor}} \right) J. \end{aligned} \tag{16}$$

Proof. For $j \geq 1$, $f \in \mathcal{M}^+$ and $\mathbf{x} \in E$, denote $\Psi_j f(\mathbf{x}) := \int_{\mathcal{I}^s} f \circ \varphi_j(\mathbf{x}, \mathbf{u}'') d\mathbf{u}''$. For $\mathbf{z} \in E$ we have:

$$D(A_{\mathbf{z}}, X^{j+1}; P_{j+1}) = D(\Psi_{j+1}\chi_{A_{\mathbf{z}}}, X^j; P_j) + D(C^j \chi_{A_{\mathbf{z}}}, Y^j). \tag{17}$$

By Lemma 1 and assumption (i), we get $|D(\Psi_{j+1}\chi_{A_{\mathbf{z}}}, X^j; P_j)| \leq D^\star(X^j; P_j)$. The function $C^j \chi_{A_{\mathbf{z}}}$ is the indicator function of

$$R^j_{\mathbf{z}} := \bigcup_{\mathbf{a} \in \mathcal{A}} \mathcal{I}_{\mathbf{a}} \times \{\mathbf{u}'' \in \mathcal{I}^s : \varphi_{j+1}(\mathbf{x}^j_{\mathbf{a}}, \mathbf{u}'') < \mathbf{z}\}, \tag{18}$$

hence $D(C^j \chi_{A_{\mathbf{z}}}, Y^j) = D(R^j_{\mathbf{z}}, Y^j)$. From (6) and (ii) we have $D(R^j_{\mathbf{z}}, Y^j) = D(\widetilde{R}^j_{\mathbf{z}}, Y^j)$, where

$$\widetilde{R}^j_{\mathbf{z}} := \bigcup_{\mathbf{a} \in \mathcal{A}} \mathcal{I}_{\mathbf{a}} \times \prod_{i=1}^{s} \left[0, (\varphi_{j+1,i}(x^j_{\mathbf{a},i}, \cdot))^{-1}(z_i)\right). \tag{19}$$

Let $\delta_s \leq d_s$ be an integer. Denote for $\mathbf{z} \in E$:

$$\Phi^{j+1}_{\mathbf{a}}(\mathbf{z}) = \left((\varphi_{j+1,1}(x^j_{\mathbf{a},1}, \cdot))^{-1}(z_1), \ldots, (\varphi_{j+1,s}(x^j_{\mathbf{a},s}, \cdot))^{-1}(z_s)\right). \tag{20}$$

Because the states are sorted and by (iii), there exist s partitions of $[0,1]$:

$$\begin{aligned}
&0 = w^j_{0,1}(\mathbf{z}) \leq w^j_{1,1}(\mathbf{z}) \leq \cdots \leq w^j_{b^{d_1},1}(\mathbf{z}) = 1, \\
&\cdots \\
&0 = w^j_{\alpha_1,\ldots,\alpha_{s-1},0,s}(\mathbf{z}) \leq w^j_{\alpha_1,\ldots,\alpha_{s-1},1,s}(\mathbf{z}) \leq \cdots \leq w^j_{\alpha_1,\ldots,\alpha_{s-1},b^{\delta_s},s}(\mathbf{z}) = 1, \\
&\quad \text{for } 0 \leq \alpha_1 < b^{d_1}, \ldots, \; 0 \leq \alpha_{s-1} < b^{d_{s-1}},
\end{aligned}$$

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

$$\begin{aligned}
\Phi^{j+1}_{\alpha_1,\ldots,\alpha_{s-1},a_s}(\mathbf{z}) \in \quad & [w^j_{\alpha_1,1}(\mathbf{z}), w^j_{\alpha_1+1,1}(\mathbf{z})] \times \cdots \\
& \times [w^j_{\alpha_1,\ldots,\alpha_s,s}(\mathbf{z}), w^j_{\alpha_1,\ldots,\alpha_s+1,s}(\mathbf{z})].
\end{aligned} \tag{21}$$

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

$$\underline{Q}^j_{\mathbf{z}} := \bigcup_{\alpha} \mathcal{J}_\alpha \times [0, w^j_{\alpha_1,1}(\mathbf{z})) \times \cdots \times [0, w^j_{\alpha_1,\ldots,\alpha_s,s}(\mathbf{z})), \tag{22}$$

$$\overline{Q}^j_{\mathbf{z}} := \bigcup_{\alpha} \mathcal{J}_\alpha \times [0, w^j_{\alpha_1+1,1}(\mathbf{z})) \times \cdots \times [0, w^j_{\alpha_1,\ldots,\alpha_s+1,s}(\mathbf{z})), \tag{23}$$

$$\begin{aligned}
\partial Q^j_{\mathbf{z}} := \bigcup_{\alpha} \mathcal{J}_\alpha \times \Big(& [w^j_{\alpha_1,1}(\mathbf{z}), w^j_{\alpha_1+1,1}(\mathbf{z})) \times \mathcal{I}^{s-1} \cup \cdots \\
& \cup [0, w^j_{\alpha_1,1}(\mathbf{z})) \times \cdots \times [w^j_{\alpha_1,\ldots,\alpha_s,s}(\mathbf{z}), w^j_{\alpha_1,\ldots,\alpha_s+1,s}(\mathbf{z})) \Big),
\end{aligned} \tag{24}$$

then $D(\underline{Q}^j_{\mathbf{z}}, Y^j) - \lambda_{2s}(\partial Q^j_{\mathbf{z}}) \le D(\widetilde{R}^j_{\mathbf{z}}, Y^j) \le D(\overline{Q}^j_{\mathbf{z}}, Y^j) + \lambda_{2s}(\partial Q^j_{\mathbf{z}})$. The subsets $\underline{Q}^j_{\mathbf{z}}$ and $\overline{Q}^j_{\mathbf{z}}$ are disjoint unions of $b^{d_1+\cdots+d_{s-1}+\delta_s}$ subintervals of $\mathcal{I}^{2s}$, hence $\max\left(|D(\underline{Q}^j_{\mathbf{z}}, Y^j)|, |D(\overline{Q}^j_{\mathbf{z}}, Y^j)|\right) \le b^{d_1+\cdots+d_{s-1}+\delta_s} D(Y^j)$. On the other hand, $\lambda_{2s}(\partial Q^j_{\mathbf{z}}) \le b^{-d_1} + \cdots + b^{-d_{s-1}} + b^{-\delta_s}$. By choosing $\delta_s = \lfloor d_s/2 \rfloor$, we get

$$|D(\widetilde{R}^j_{\mathbf{z}}, Y^j)| \le b^{d_1+\cdots+d_{s-1}+\lfloor d_s/2 \rfloor} D(Y^j) + \frac{1}{b^{d_1}} + \cdots + \frac{1}{b^{d_{s-1}}} + \frac{1}{b^{\lfloor d_s/2 \rfloor}}. \qquad (25)$$

The desired result follows by taking the supremum over $\mathbf{z} \in E$ and by induction on j.

4. Numerical examples

In this section, we present the results of numerical experiments which show the kind of improvement that our method can bring with respect to MC, even when the restrictive assumptions of Proposition 1 are not fulfilled. The examples we choose are artificial since the exact solutions are known and can be analytically calculated, but we use them as a benchmark to evaluate the viability of our method. The MC computations are done using the pseudorandom points generated by `MRG32k3a` of [9]. The QMC computations use Niederreiter's sequences in base $b = 2$ [12].

4.1. Asian option

We consider the pricing of an Asian option on a single asset whose value $S(t)$ obeys: $dS(t) = rS(t)dt + \sigma S(t)dB(t)$, where r is the *risk-free interest rate*, σ the *volatility parameter* and B is a standard Brownian motion (BM). Consider discrete observation times $0 = t_0 < t_1 < \cdots < t_J = T$ and write:

$$S(t_j) = S(t_{j-1}) \exp\left((r - \sigma^2/2)\delta t_j + \sigma\sqrt{\delta t_j} Z_j\right), \qquad (26)$$

where $\delta t_j := t_j - t_{j-1}$ and $\{Z_j : j \ge 1\}$ is a sequence of i.i.d. standard normal variables. The value of the call option at maturity can be written as $C_A = e^{-rT}\mathbb{E}[\max((\prod_{j=1}^J S(t_j))^{1/J} - K, 0)]$ where the constant K is the *strike price.* We want to estimate C_A by our QMC algorithm and compare the results with those given by a classical MC scheme. Thus, we define a bi-dimensional Markov chain by: $X_0 := (S(t_0), 1)$ and $X_j := (S(t_j), (\prod_{h=1}^j S(t_h))^{1/j})$, for $j \ge 1$. Here $s = 2, d = 1$ and we consider the following parameters: $S(0) = 100$, $r = 0.037$, $\sigma = 0.2$, $T = 240/365$, $K = 90$. We estimate the error for

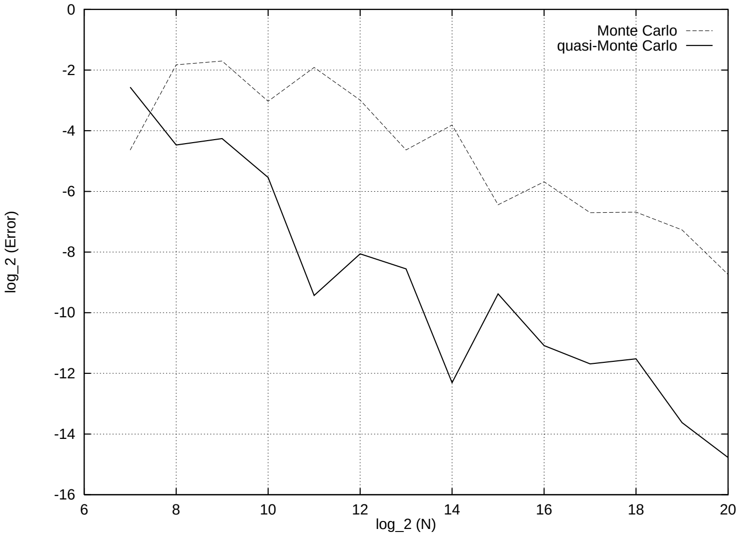

Figure 2: Asian option : the error as a function of N for MC (thin line) and QMC (thick line)

$J = 120$ as a function of N, say $\mathrm{Err}_{\mathrm{MC}}(N)$ for MC and $\mathrm{Err}_{\mathrm{QMC}}(N)$ for the QMC method. The value of N varies from 2^7 to 2^{20}. Figure 2 shows the errors, in log-log scale. A linear regression analysis estimates the empirical convergence rate of the QMC method to be of the order of $\mathcal{O}(N^{-0.84})$. Clearly, the QMC algorithm enjoys a much faster convergence than the MC scheme, whose convergence rate is known to be $\mathcal{O}(N^{-0.50})$.

4.2. European option on the maximum of two risky assets

For our second example, we consider the pricing of an European call option on the maximum of two risky assets. The model is a bivariate geometric Brownian motion $S(t) = (S_1(t), S_2(t))$ with interest rate r and volatility parameters σ_1 and σ_2. Thus, for $i = 1, 2$: $dS_i(t) = rS_i(t)dt + \sigma_i S_i(t)dB_i(t)$, where B_1 and B_2 are two standard BM with *correlation parameter* ρ. For a strike price $K > 0$, the option has discounted payoff $e^{-rT}\max(\max(S_1(T), S_2(T)) - K, 0)$ at maturity date $T > 0$. The expected value C_M of this payoff can be computed by formulas given in [5]. To estimate C_M, we discretize the problem using a set of observation times

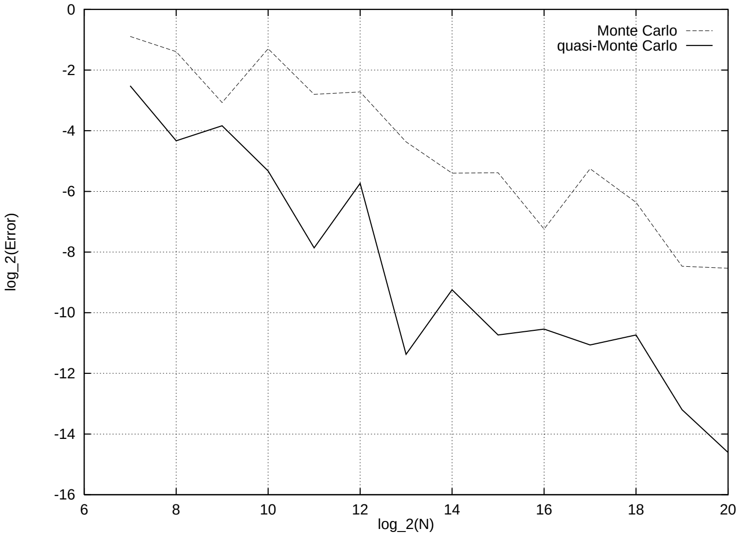

Figure 3: European option on the maximum of two risky assets : the error as a function of N for MC (thin line) and QMC (thick line)

$0 = t_0 < t_1 < \cdots < t_J$ and we simulate the assets as follows:

$$\begin{aligned} S_1(t_j) &= S_1(t_{j-1}) \exp\Big((r - \sigma_1^2/2)\delta t_j + \sqrt{\delta t_j}(\sigma_1 Z_{j,1})\Big) &(27)\\ S_2(t_j) &= S_2(t_{j-1}) \\ &\quad \cdot \exp\Big((r - \sigma_2^2/2)\delta t_j + \sqrt{\delta t_j}(\sigma_2 \rho Z_{j,1} + \sigma_2\sqrt{1-\rho^2}Z_{j,2})\Big), &(28)\end{aligned}$$

where $\{Z_j : j \geq 1\}$ are i.i.d. random variables such that $Z_j \sim \mathcal{N}(0, I_2)$ (here I_2 is the identity matrix). We define the Markov chain by $X_j = (S_1(t_j), S_2(t_j))$. Here $s = d = 2$; for numerical illustration, let $S_1(0) = S_2(0) = 40$, $r = 0.048$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $\rho = 0.5$, $T = 7/12$, $K = 35$ and $J = 100$. The number N of paths varies from 2^7 to 2^{20}. The values of the errors are shown on Figure 3. Here again, regression analysis estimates the convergence speed to be $\mathrm{Err}_{\mathrm{QMC}} = \mathcal{O}(N^{-0.84})$ for QMC, showing a strong improvement over MC.

5. Conclusion

We have presented a QMC algorithm for the simulation of Markov chains with continuous and multi-dimensional state space. The method simulates several copies of the chain in parallel and reduces the error by sorting the states used in the simulation according to their successive coordinates at each step. Under certain assumptions, we have proved a convergence result as the number of simulated paths increases. The results of some numerical examples have shown that our QMC method is clearly superior to standard MC simulation in magnitude of error and in convergence rate. In the future, we shall analyze the convergence in more general settings and we shall provide some experiments with larger and more complicated models.

[1] R. Beauwens, A. Dubus, M. Hou (Eds.), Selection of papers presented at the IMACS Seminar on Monte Carlo Methods, Math. Comput. Simulation 47 (1998) 83–518.

[2] I. Dimov (Ed.), IMACS Sponsored Special Issue: The Second IMACS Seminar on Monte Carlo Methods, Math. Comput. Simulation 55 (2001) 1–305.

[3] R. El Haddad, C. Lécot, P. L'Ecuyer, Quasi-Monte Carlo simulation of discrete-time Markov chains on multidimensional state spaces, in: S. Heinrich, A. Keller, and H. Niederreiter (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2006, Springer-Verlag, Berlin, 2008, pp. 413–429.

[4] K. Entacher, W.C. Schmid, A. Uhl (Eds.), Special Issue: 3rd IMACS Seminar on Monte Carlo Methods, Math. Comput. Simulation 62 (2003) 217–571.

[5] H. Johnson, Options on the maximum or the minimum of several assets, J. Financ. Quant. Anal. 22 (1987) 277–283.

[6] C. Lécot, A direct simulation Monte Carlo scheme and uniformly distributed sequences for solving the Boltzmann equation, Computing 41 (1989) 41–57.

[7] C. Lécot, B. Tuffin, Quasi-Monte Carlo methods for estimating transient measures of discrete time Markov chains, in: H. Niederreiter (Ed.), Monte Carlo and Quasi-Monte Carlo Methods 2002, Springer-Verlag, Berlin, 2004, pp. 329–343.

[8] C. Lécot, B. Tuffin, Comparison of quasi-Monte Carlo-based methods for simulation of Markov chains, Monte Carlo Methods Appl. 10 (2004) 377–384.

[9] P. L'Ecuyer, Good parameters and implementations for combined multiple recursive random number generators, Oper. Res. 47 (1999) 159–164.

[10] P. L'Ecuyer, C. Lécot, B. Tuffin, Randomized quasi-Monte Carlo simulation of Markov chains with an ordered state space, in: H. Niederreiter, D. Talay (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2004, Springer-Verlag, Berlin, 2006, pp. 331–342.

[11] P. L'Ecuyer, C. Lécot, B. Tuffin, A randomized quasi-Monte Carlo simulation method for Markov chains, Oper. Res. 56 (2008) 958–975.

[12] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods, SIAM, Philadelphia, 1992.

[13] K. K. Sabelfeld (Ed.), Selection of papers presented at the international conference: 4th IMACS Seminar on Monte Carlo Methods, Monte Carlo Methods Appl. 10 (2004) 183–655.