Les Cahiers du GERAD ISSN: 0711–2440

Simple stratified sampling for simulating multi-dimensional Markov chains

R. El Haddad, C. Lécot, P. L'Ecuyer

G–2023–68

December 2023

La publication de ces rapports de recherche est rendue possible grâce au soutien de HEC Montréal, Polytechnique Montréal, Université McGill, Université du Québec à Montréal, ainsi que du Fonds de recherche du Québec – Nature et technologies.

Dépôt légal – Bibliothèque et Archives nationales du Québec, 2023
– Bibliothèque et Archives Canada, 2023

The publication of these research reports is made possible thanks to the support of HEC Montréal, Polytechnique Montréal, McGill University, Université du Québec à Montréal, as well as the Fonds de recherche du Québec – Nature et technologies.

Legal deposit – Bibliothèque et Archives nationales du Québec, 2023
– Library and Archives Canada, 2023

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Simple stratified sampling for simulating multi-dimensional Markov chains

Rami El Haddad [a]

Christian Lécot [b]

Pierre L'Ecuyer [c, d]

[a] *Faculté des Sciences, Université Saint-Joseph, Mar Mikhaël Beyrouth 1104 2020, Liban*

[b] *Université Savoie Mont Blanc, LAMA, UMR 5127 CNRS, 73376 Le Bourget-du-Lac, France*

[c] *DIRO, Université de Montréal, Montréal (Qc), Canada, H3C 3J7*

[d] *GERAD, Montréal (Qc), Canada, H3T 1J4*

rami.haddad@usj.edu.lb
christian.lecot@univ-smb.fr
lecuyer@iro.umontreal.ca

December 2023
Les Cahiers du GERAD
G–2023–68

Abstract : Monte Carlo (MC) is widely used for the simulation of discrete time Markov chains. We consider the case of a d-dimensional continuous state space and we restrict ourselves to chains where the d components are advanced independently from each other, with d random numbers used at each step. We simulate N copies of the chain in parallel, and we replace pseudorandom numbers on $I^d := (0,1)^d$ with stratified random points over I^{2d}: for each point, the first d components are used to select a state and the last d components are used to advance the chain by one step. We use a simple stratification technique: let p be an integer, then for $N = p^{2d}$ samples, the unit hypercube is dissected into N hypercubes of measure $1/N$ and there is one sample in each of them. The strategy outperforms classical MC if a well-chosen multivariate sort of the states is employed to order the chains at each step. We prove that the variance of the stratified sampling estimator is bounded by $\mathcal{O}(N^{-(1+1/(2d))})$, while it is $\mathcal{O}(N^{-1})$ for MC. In numerical experiments, we observe empirical rates that satisfy the bounds. We also compare with the Array-RQMC method.

1 Introduction

Markov chains with a large state space can be used to model a variety of real life systems in domains such as particles physics, telecommunications, queueing theory, mathematical finance, etc. In many situations, neither analytic solutions are available nor deterministic numerical methods are practicable because the state space is too large. So Monte Carlo (MC) simulation becomes the standard way to solve the model.

We consider a discrete time Markov chain, with state space $\mathcal{X} := \prod_{i=1}^{d} \mathcal{X}_i$, for some integer $d > 0$, where $\mathcal{X}_i = (a_i, b_i)$, with $-\infty \leq a_i < b_i \leq +\infty$, for $1 \leq i \leq d$. We assume that d-dimensional random variates are used to advance the chain by one step. So X_0 is a d-dimensional random variable (it can be degenerate over a single state) and the chain evolves according to the recurrence:

$$X_{n+1} = \varphi_{n+1}(X_n, U_{n+1}), \quad n \geq 0. \tag{1}$$

Here $U_1, U_2, \ldots$ are i.i.d. uniform random variables over I^d, where $I := (0, 1)$, and $\varphi_1, \varphi_2, \ldots$ are measurable functions $\mathcal{X} \times I^d \to \mathcal{X}$. Let P_n denote the law of X_n. We want to estimate $\int_{\mathcal{X}} c(x) dP_n(x)$, for some cost function c, by a mean $(\sum_j c(X_j^n))/N$, where $\{X_j^n : 1 \leq j \leq N\}$ are realizations of the chain at step n (to be defined). The mean must be unbiased and the convergence of the method is evaluated by $\mathrm{Var}((\sum_j c(X_j^n))/N)$. The MC approach uses pseudorandom numbers as realizations of uniform random variables over I^d. The drawback is that convergence can be slow, with respect to the number of samples, classically as $\mathcal{O}(N^{-1})$.

Stratified sampling is a technique for increasing efficiency of MC methods: see, e.g., [8, 14]. We use here the following strategy, called *simple stratified sampling (SSS)*: for a dimension δ, the unit cube I^δ is partitioned into $N = p^\delta$ intervals $J_\kappa := \prod_{i=1}^{\delta} [(k_i - 1)/p, k_i/p)$, for $\kappa = (k_1, \ldots, k_\delta)$, where each integer k_i lies between 1 and p; in each J_κ, a point is selected at random. The SSS is analyzed in [2, 3, 9] for integration of smooth functions: the authors show an improved convergence rate for the variance (compared to classical MC). SSS for simulating Markov chains has been proposed in [4]. When the state space is one-dimensional, variance reduction was theoretically analyzed in [5], for a discrete space and in [6] for a continuous state space.

This approach resembles the *randomized quasi-Monte Carlo* (Array-RQMC) method initiated in [11, 12]. It has been observed that, in many cases, Array-RQMC reduces the variance [1, 13, 15]. However, an improved convergence rate compared with MC has been proved only for very specific cases, such as a one-dimensional state space with a stratification approach, for which a rate of $\mathcal{O}(N^{-3/2})$ was established under certain conditions [11]. For both methods, the N copies of the chain are sorted at each step before being moved forward; this ensures theoretical and numerical convergence. Array-RQMC has been observed empirically to perform very well for several applications [1, 11, 12, 15], often with much better rates than $\mathcal{O}(N^{-1})$ for the variance, but we still have no formal proof for these better rates in more than one dimension (a difficulty is to find a multidimensional sort which guarantees convergence). The aim of the present paper is to offer progress in this direction.

Specifically, we prove a $\mathcal{O}(N^{-(1+1/(2d))})$ variance bound for a SSS method applied to Markov chains with a d-dimensional state space, under certain conditions. The conditions are restrictive, but they are satisfied in interesting situations (e.g. the Gaussian random walk). The SSS method examined here differs slightly from Array-RQMC combined with stratification as defined in [11]. We point out the difference at the end of Section 2.

The remainder is organized as follows. In Section 2, we recall the MC method for simulating such Markov chains and we present the approach using simple stratified (SS) samples. In Section 3, we compare the bounds for the variance of the estimator obtained by the MC method and the SSS strategy. We provide numerical illustrations in Section 4: we compute the empirical variances of the estimators and compare them with the theoretical bounds previously established. We also compare with the Array-RQMC method. Conclusions are drawn in Section 5.

2 Markov chain simulation

Consider a chain with the recurrence given by (1). We suppose that the d components of the chain are advanced independently from each other:

$$\varphi_{n+1}(x,u) = (\varphi_{n+1,1}(x_1,u_1),\ldots,\varphi_{n+1,d}(x_d,u_d)). \tag{2}$$

This is a strong limitative assumption. There are interesting situations in which this assumption is satisfied. This is the case of the random walk construction of a standard d-dimensional Brownian motion at discrete times. It is used for solving physical problems that involve a combination of convection, reaction and diffusion. Besides, a restricted convergence result may be obtained under a less limitative hypothesis (see Remark 1). A random walk is also utilized within the Black-Scholes model in financial engineering: an example is given in Section 4.

We present the usual MC method and the SSS version to simulate the chain. Denote by $\mathcal{B}_+(\mathcal{X})$ the set of nonnegative Borel measurable functions defined on $\mathcal{X}$. For every $s \in \mathcal{B}_+(\mathcal{X})$ we have

$$\int_{\mathcal{X}} s(x)dP_{n+1}(x) = \int_{\mathcal{X}\times I^d} s\circ\varphi_{n+1}(x,u)dP_n(x)du. \tag{3}$$

The following notations are used. If m,n are integers, then $[m,n] := \{m,m+1,\ldots,n\}$ (if $m \le n$) and $(m,n] := \{m+1,m+2,\ldots,n\}$ (if $m < n$). Let U be a random variable, then $U \sim \mathfrak{U}(\mathcal{E})$ means that U is uniformly distributed over the set $\mathcal{E}$. For $z \in \mathcal{X}$, we denote s_z the indicator function of $\prod_{i=1}^d (a_i, z_i)$.

2.1 Classical Monte Carlo

We present how to simulate in parallel instead of the standard MC sequential algorithm to highlight the difference with the variance reduction techniques described later. We choose an integer N, the number of independent samples. We suppose that, at step n, $\{X_j^n : j \in [1,N]\}$ are random variables (the states) such that, for a family $\mathcal{C} \subset \mathcal{B}_+(\mathcal{X})$ of *cost functions*,

$$\forall s \in \mathcal{C} \quad \frac{1}{N}\sum_j s(X_j^n) \approx \int_{\mathcal{X}} s(x)dP_n(x).$$

Here, the approximation is assumed to be unbiased and consistent as follows.

$$\mathbb{E}\left[\frac{1}{N}\sum_j s(X_j^n)\right] = \int_{\mathcal{X}} s(x)dP_n(x) \quad \text{and} \quad \mathrm{Var}\left(\frac{1}{N}\sum_j s(X_j^n)\right) = \mathcal{O}\left(\frac{1}{N}\right).$$

The analysis is first done for cost functions of the form s_z (Eqs. (8) and (9)); a result for regular cost functions is given next (Corollary 1).

We want to find new states at step $n+1$. Following (3), for every $s \in \mathcal{C}$:

$$\int_{\mathcal{X}} s(x)dP_{n+1}(x) \approx \frac{1}{N}\sum_j \int_{I^d} s\circ\varphi_{n+1}(X_j^n,u)du. \tag{4}$$

For $j \in [1,N]$, let 1_j denote the indicator function of $F_j := [(j-1)/N, j/N)$. We associate to any $s \in \mathcal{B}_+(\mathcal{X})$ the following function of $d+1$ variables:

$$A_s^n(\underline{u}) := \sum_j 1_j(u_0)s\circ\varphi_{n+1}(X_j^n,u), \quad \underline{u} = (u_0,u) \in I\times I^d.$$

Then,

$$\int_{I^{d+1}} A_s^n(\underline{u})d\underline{u} = \frac{1}{N}\sum_j \int_{I^d} s\circ\varphi_{n+1}(X_j^n,u)du. \tag{5}$$

A numerical quadrature is done for the transition from n to $n+1$. Let $\{U_j^{n+1} : j \in [1,N]\}$ be independent random variables with $U_j^{n+1} \sim \mathfrak{U}(I^d)$. We replace the left-hand side of (5) with the following MC estimate:

$$\widehat{X}_s^{n+1} := \frac{1}{N}\sum_j A_s^n\left(\frac{j-1}{N}, U_j^{n+1}\right) = \frac{1}{N}\sum_j s\circ\varphi_{n+1}(X_j^n, U_j^{n+1}).$$

We generate

$$X_j^{n+1} = \varphi_{n+1}(X_j^n, U_j^{n+1}),$$

so that

$$\widehat{X}_s^{n+1} = \frac{1}{N}\sum_j s(X_j^{n+1})$$

is chosen as an approximation of $\int_{\mathcal{X}} s(x)dP_{n+1}(x)$.

2.2 Simple Stratified Sampling

We choose $N = p^{2d}$, for some integer $p > 0$. This is a first limitation of the method: its practical use is restricted to small d's, because of the exponential growth of the sample size. We suppose that we have generated a set of random variables $\{Y_\mu^n : \mu = (m_1,\ldots,m_d) \in [1,p]^{d-1}\times[1,p^{d+1}]\}$ such that

$$\forall s \in \mathcal{C} \quad \frac{1}{N}\sum_\mu s(Y_\mu^n) \approx \int_{\mathcal{X}} s(x)dP_n(x).$$

Here, the Y_μ^n replace the X_j^n of MC; the indexing will be justified in the proof of Lemma 1. The approximation is assumed to be unbiased and consistent as for MC. For every $s \in \mathcal{C}$,

$$\int_{\mathcal{X}} s(x)dP_{n+1}(x) \approx \frac{1}{N}\sum_\mu \int_{I^d} s\circ\varphi_{n+1}(Y_\mu^n, u)du.$$

In order to find the new states at step $n+1$, a numerical quadrature is done. For $\mu \in [1,p]^{d-1}\times[1,p^{d+1}]$, let 1_μ denote the indicator function of

$$L_\mu := \prod_{i=1}^{d-1}\left[\frac{m_i-1}{p}, \frac{m_i}{p}\right) \times \left[\frac{m_d-1}{p^{d+1}}, \frac{m_d}{p^{d+1}}\right).$$

To any $s \in \mathcal{B}_+(\mathcal{X})$, we associate the function of $2d$ variables:

$$B_s^n(v) := \sum_\mu 1_\mu(\dot{v}) s\circ\varphi_{n+1}(Y_\mu^n, \ddot{v}), \quad v = (\dot{v},\ddot{v}) \in I^d \times I^d.$$

Then,

$$\int_{I^{2d}} B_s^n(v)dv = \frac{1}{N}\sum_\mu \int_{I^d} s\circ\varphi_{n+1}(Y_\mu^n, \ddot{v})d\ddot{v}. \tag{6}$$

The transition from n to $n+1$ has two steps: renumbering the chains and numerical quadrature.

(S1) The chains are relabeled so that if $\mu = (m_1,\ldots,m_d), \mu' = (m_1',\ldots,m_d')$,

$$m_1 = m_1', \ldots, m_{i-1} = m_{i-1}', m_i < m_i' \;\Rightarrow\; Y_{\mu,i}^n \le Y_{\mu',i}^n. \tag{7}$$

This lexicographic ordering (called *multivariate batch sort* in [13]) was first introduced in a QMC context and motivated in [10]: it guaranteed theoretical and numerical convergence. Figure 1 depicts the indices $\mu = (m_1, m_2)$ on the locations of the states Y_μ^n. The states are first grouped in two batches (horizontally) according to the first coordinate, then each batch is sorted according to the second coordinate.

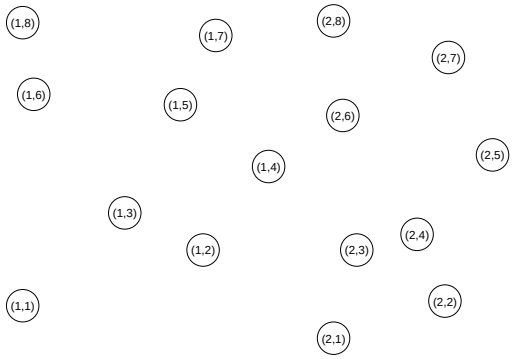

Figure 1: Lexicographic ordering of $2^{2\times 2}$ states in $\mathbb{R}^2$

This is a second limitation of the SSS method: its use requires d orderings of N numbers at each step.

(S2) For $\kappa = (k_1, \ldots, k_{2d}) \in [1,p]^{2d}$, let

$$J_\kappa := \prod_{i=1}^{2d} \left[\frac{k_i - 1}{p}, \frac{k_i}{p} \right).$$

Let $\{V_\kappa^{n+1} : \kappa \in [1,p]^{2d}\}$ be independent random variables with $V_\kappa^{n+1} \sim \mathfrak{U}(J_\kappa)$. The first d-dimensional projection $\dot{V}_\kappa^{n+1}$ of V_κ^{n+1} is used for selecting the state at step n and the second d-dimensional projection $\ddot{V}_\kappa^{n+1}$ is used for advancing the chain by one step. That is, we replace the left-hand side of (6) with the SSS estimate:

$$\widehat{Y}_s^{n+1} := \frac{1}{N} \sum_\kappa B_s^n(V_\kappa^{n+1}) = \frac{1}{N} \sum_\kappa s \circ \varphi_{n+1}(Y^n_{\mu(\dot{V}_\kappa^{n+1})}, \ddot{V}_\kappa^{n+1}),$$

where, for $u \in I^d$, the index $\mu(u) \in [1,p]^{d-1} \times [1, p^{d+1}]$ is such that $u \in L_{\mu(u)}$. We define

$$Y_\kappa^{n+1} = \varphi_{n+1}\left(Y^n_{\mu(\dot{V}_\kappa^{n+1})}, \ddot{V}_\kappa^{n+1}\right),$$

so that

$$\widehat{Y}_s^{n+1} = \frac{1}{N} \sum_\kappa s(Y_\kappa^{n+1})$$

is chosen as an approximation of $\int_{\mathcal{X}} s(x) dP_{n+1}(x)$.

It is noticeable that the correspondence $\kappa \in [1,p]^{2d} \to \mu(\dot{V}_\kappa^{n+1}) \in [1,p]^{d-1} \times [1, p^{d+1}]$ between the stratified samples V_κ^{n+1} and the states Y_μ^n is not necessarily bijective: the same state may be chosen more than once or some states may be left out. This diverges from the stratified sampling construction in the Array-RQMC algorithm [11, 13], subsequently called A-Strat.

3 Theoretical convergence

For classical MC, let $\{X_j^0 : j \in [1,N]\}$ be i.i.d. random variables with probability distribution P_0. Then, the following holds for any $n \geq 0$.

1. For any $s \in \mathcal{B}_+(\mathcal{X})$,

$$\mathbb{E}\left[\frac{1}{N}\sum_j s(X_j^n)\right] = \int_{\mathcal{X}} s(x) dP_n(x). \tag{8}$$

2. For any $z \in \mathcal{X}$,

$$\mathrm{Var}\left(\frac{1}{N}\sum_j s_z(X_j^n)\right) \le \frac{1}{4N}. \tag{9}$$

For SSS, we next show that, at every step, (1) for any $s \in \mathcal{B}_+(\mathcal{X})$, the estimator of $\int_{\mathcal{X}} s(x)dP_n(x)$ is unbiased, and (2) for any $z \in \mathcal{X}$, the variance of the estimator of $\int_{\mathcal{X}} s_z(x)dP_n(x)$ is bounded by $\mathcal{O}(N^{-(1+1/(2d))})$.

We suppose that P_0 has density f_0 and we assume the following for any $i \in [1, d]$.

A1. For any $x \in \mathcal{X}_i$, the mapping $u \in I \to \varphi_{n+1,i}(x, u) \in \mathcal{X}_i$ is strictly increasing, bijective; one defines a function $y \in \mathcal{X}_i \to \psi_{n+1,i}(x, y) \in I$ which is strictly increasing, bijective, such that $y = \varphi_{n+1,i}(x, u) \Leftrightarrow u = \psi_{n+1,i}(x, y)$.

A2. For any $x \in \mathcal{X}_i$, the mapping $y \to \psi_{n+1,i}(x, y)$ is continuously differentiable.

A3. For any $y \in \mathcal{X}_i$, the mapping $x \to \psi_{n+1,i}(x, y)$ is continuously differentiable; there exists a constant $M_{n+1,i}$ such that $\int_{\mathcal{X}_i} |\frac{\partial \psi_{n+1,i}}{\partial x}(x, y)|dx \le M_{n+1,i}$; in addition $\lim_{x\to b_i} \psi_{n+1,i}(x, y) = 0$.

We define:

$$\Pi_{n+1} := \prod_{i=1}^{d} M_{n+1,i} \quad \text{and} \quad \Sigma_{n+1} := \sum_{i=1}^{d} M_{n+1,i}.$$

For initialization, we would like to start with a better convergence rate than MC. We assume that we can define a set $\{Y_\mu^0 : \mu = (m_1, \ldots, m_d) \in [1, p]^{d-1} \times [1, p^{d+1}]\}$ of independent random variables with the following properties (which are used in proofs by induction in Proposition 1).

P1. For any $s \in \mathcal{B}_+(\mathcal{X})$,

$$\mathbb{E}\left[\frac{1}{N}\sum_\mu s(Y_\mu^0)\right] = \int_{\mathcal{X}} s(x)dP_0(x). \tag{10}$$

P2. There exists some $\beta_0 > 0$ such that, for every $z \in \mathcal{X}$,

$$\mathrm{Var}\left(\frac{1}{N}\sum_\mu s_z(Y_\mu^0)\right) \le \frac{\beta_0}{N^{1+1/2d}}. \tag{11}$$

If the density f_0 of P_0 is such that $f_0(x) = \prod_{i=1}^d f_{0,i}(x_i)$, this may be obtained by the inversion method. Let $F_{0,i}$ be the cumulative distribution function (cdf) associated with $f_{0,i}$. For $\lambda = (\ell_1, \ldots, \ell_d) \in [1, p^2]^d$, denote

$$K_\lambda := \prod_{i=1}^{d} \left[\frac{\ell_i - 1}{p^2}, \frac{\ell_i}{p^2}\right).$$

Let $\{V_\lambda : \lambda \in [1, p^2]^d\}$ be independent random variables with $V_\lambda \sim \mathfrak{U}(K_\lambda)$. We take

$$Y_\lambda^0 := \left(F_{0,1}^{-1}(V_{\lambda,1}), \ldots, F_{0,d}^{-1}(V_{\lambda,d})\right).$$

Then (10) is satisfied (after re-indexing) and there exists $\beta_0 \le d/4$ such that for any $z \in \mathcal{X}$,

$$\mathrm{Var}\left(\frac{1}{N}\sum_\lambda s_z(Y_\lambda^0)\right) \le \frac{\beta_0}{N^{1+1/d}}.$$

After, the analysis has two stages: (1) a variance bound for step n to $n+1$, conditional on the states at step n, and (2) a bound of the unconditional variance (over several steps). We first focus on one step and we assume that $\{y_\mu^n : \mu \in [1, p]^{d-1} \times [1, p^{d+1}]\}$ are given vectors with a numbering satisfying (7). Let

$$\widehat{\mathcal{Y}}_s^{n+1} := \frac{1}{N}\sum_\kappa \sum_\mu 1_\mu(\dot{V}_\kappa^{n+1}) s \circ \varphi_{n+1}(y_\mu^n, \ddot{V}_\kappa^{n+1}).$$

Lemma 1. *For the SSS method, we have:*

1. *For any* $s \in \mathcal{B}_+(\mathcal{X})$,
$$\mathbb{E}[\widehat{\mathcal{Y}}_s^{n+1}] = \frac{1}{N}\sum_{\mu}\int_{I^d} s\circ\varphi_{n+1}(y_\mu^n, u)du.$$
2. *For any* $z \in \mathcal{X}$,
$$\mathrm{Var}(\widehat{\mathcal{Y}}_{s_z}^{n+1}) \leq \frac{d(\Sigma_{n+1}+2)}{4N^{1+1/2d}}.$$

Proof.

1. Straightforward.
2. The transition from n to $n+1$ is described by a numerical quadrature of a function B_s^n; if $s = s_z$, then $B_s^n = B_{s_z}^n$ is the indicator function of the following set:
$$\mathcal{E}_z^n := \bigcup_{\mu} L_\mu \times \prod_{i=1}^{d}\left[0, \psi_{n+1,i}(y_{\mu,i}^n, z_i)\right).$$
The variable $B_{s_z}^n(V_\kappa^{n+1})$ is a Bernoulli random variable, with expectation $e_{z,\kappa}^{n+1} = N\lambda_{2d}(\mathcal{E}_z^n \cap J_\kappa)$. Hence, $\mathrm{Var}(B_{s_z}^n(V_\kappa^{n+1})) \leq 1/4$ and $\mathrm{Var}(B_{s_z}^n(V_\kappa^{n+1})) = 0$ if $J_\kappa \subset \mathcal{E}_z^n$ or if $J_\kappa \cap \mathcal{E}_z^n = \emptyset$. The only $\kappa' s$ that contribute to the variance of $\widehat{\mathcal{Y}}_{s_z}^{n+1}$ are those for which J_κ intersects the boundary of $\mathcal{E}_z^n$. Therefore,
$$\mathrm{Var}(\widehat{\mathcal{Y}}_{s_z}^{n+1}) \leq \frac{1}{4N^2}\left|\{\kappa \in [1,p]^{2d} : J_\kappa \not\subset \mathcal{E}_z^n \text{ and } J_\kappa \cap \mathcal{E}_z^n \neq \emptyset\}\right|,$$
where $|\mathcal{N}|$ denotes the cardinality of a (finite) set $\mathcal{N}$. We have:
   - if $J_\kappa \not\subset \mathcal{E}_z^n$, then
   $$\exists i \in [1,d] \quad k_{d+i} > p \min_{m_d \in ((k_d-1)p^d, k_d p^d]} \psi_{n+1,i}(y^n_{(k_1,\ldots,k_{d-1},m_d),i}, z_i),$$
   - if $J_\kappa \cap \mathcal{E}_z^n \neq \emptyset$, then
   $$\forall i \in [1,d] \quad k_{d+i} < p \max_{m_d \in ((k_d-1)p^d, k_d p^d]} \psi_{n+1,i}(y^n_{(k_1,\ldots,k_{d-1},m_d),i}, z_i) + 1.$$

   Hence,
$$\mathrm{Var}(\widehat{\mathcal{Y}}_{s_z}^{n+1}) \leq \frac{p^d}{4N^2}\left(\sum_{\dot\kappa}\sum_{i=1}^{d}\left(\max_{m_d \in ((k_d-1)p^d, k_d p^d]} \psi_{n+1,i}(y^n_{(k_1,\ldots,k_{d-1},m_d),i}, z_i)\right.\right.$$
$$\left.\left. - \min_{m_d \in ((k_d-1)p^d, k_d p^d]} \psi_{n+1,i}(y^n_{(k_1,\ldots,k_{d-1},m_d),i}, z_i)\right) + 2dp^{d-1}\right).$$

   Up to now, the proof is a generalization of the one-dimensional case done in [6]. Henceforth, we use the lexicographic ordering of the states. As a consequence of (7), there exists a covering of $\prod_{i=1}^d [a_i, b_i]$ with
$$\begin{aligned}
a_1 &= w_{0,1}^n \leq w_{1,1}^n \leq \cdots \leq w_{p,1}^n = b_1,\\
a_2 &= w_{m_1,0,2}^n \leq w_{m_1,1,2}^n \leq \cdots \leq w_{m_1,p,2}^n = b_2 \quad \text{for } m_1 \in [1,p],\\
&\cdots\\
a_d &= w_{m_1,\ldots,m_{d-1},0,d}^n \leq w_{m_1,\ldots,m_{d-1},1,d} \leq \cdots \leq w_{m_1,\ldots,m_{d-1},p,d} = b_d\\
&\quad \text{for } (m_1,\ldots,m_{d-1}) \in [1,p]^{d-1}
\end{aligned}$$

such that, if, for $\dot{\kappa} := (k_1, \ldots, k_d) \in [1,p]^d$ we define

$$R^n_{\dot{\kappa}} := [w^n_{k_1-1,1}, w^n_{k_1,1}] \times [w^n_{k_1,k_2-1,2}, w^n_{k_1,k_2,2}] \times \cdots \times [w^n_{k_1,\ldots,k_{d-1},k_d-1}, w^n_{k_1,\ldots,k_d}],$$

then we have $y^n_{(k_1,\ldots,k_{d-1},m_d)} \in R^n_{\dot{\kappa}}$ for any $m_d \in ((k_d-1)p^d, k_d p^d]$. Hence, there exist points $\check{y}^{n,i}_{\dot{\kappa}}, \hat{y}^{n,i}_{\dot{\kappa}} \in R^n_{\dot{\kappa}}$ so that

$$\begin{aligned}\mathrm{Var}(\widehat{\mathcal{Y}}^{n+1}_{s_z}) &\leq \frac{p^d}{4N^2}\left(\sum_{\dot{\kappa}}\sum_{i=1}^{d} |\psi_{n+1,i}(\hat{y}^{n,i}_{\dot{\kappa},i}, z_i) - \psi_{n+1,i}(\check{y}^{n,i}_{\dot{\kappa},i}, z_i)| + 2dp^{d-1}\right)\\ &\leq \frac{p^d}{4N^2}\left(dp^{d-1}\sum_{i=1}^{d} M_{n+1,i} + 2dp^{d-1}\right),\end{aligned}$$

where the last inequality is obtained by using a modified version of Lemma 4 in [10]. The conclusion follows. □

The proof of the next result uses techniques employed in [11].

Proposition 1. *For the SSS method, the following holds.*

1. *For any $s \in \mathcal{B}_+(\mathcal{X})$,*
$$\mathbb{E}\left[\frac{1}{N}\sum_{\mu} s(Y^n_\mu)\right] = \int_{\mathcal{X}} s(x) dP_n(x).$$
2. *For any $z \in \mathcal{X}$,*
$$\mathrm{Var}\left(\frac{1}{N}\sum_{\mu} s_z(Y^n_\mu)\right) \leq \frac{\beta_n}{N^{1+1/2d}}, \tag{12}$$
where $\beta_{n+1} = \Pi^2_{n+1}\beta_n + d(\Sigma_{n+1}+2)/4$.

Proof.

1. We prove the result by induction. It holds for $n = 0$ from (10). For $n \geq 0$,
$$D^{n+1}_s := \int_{\mathcal{X}} s(x) dP_{n+1}(x) - \frac{1}{N}\sum_{\mu} s(Y^{n+1}_\mu) = D^n_{s,1} + D^n_{s,2},$$
where
$$D^n_{s,1} := \int_{\mathcal{X}}\int_{I^d} s\circ\varphi_{n+1}(x,u)du dP_n(x) - \frac{1}{N}\sum_{\mu}\int_{I^d} s\circ\varphi_{n+1}(Y^n_\mu, u)du,$$
$$D^n_{s,2} := \frac{1}{N}\sum_{\mu}\int_{I^d} s\circ\varphi_{n+1}(Y^n_\mu, u)du - \frac{1}{N}\sum_{\kappa}\sum_{\mu} 1_\mu(\dot{V}^{n+1}_\kappa) s\circ\varphi_{n+1}(Y^n_\mu, \ddot{V}^{n+1}_\kappa).$$
By the induction hypothesis, $\mathbb{E}[D^n_{s,1}] = 0$; by Lemma 1, $\mathbb{E}[D^n_{s,2}] = 0$. The result follows.
2. We proceed by induction on n. The case $n = 0$ is given by (11). Let $n \geq 0$ be arbitrary. Since $\{Y^n_\mu : \mu \in [1,p]^{d-1} \times [1,p^{d+1}]\}$ and $\{V^{n+1}_\kappa : \kappa \in [1,p]^{2d}\}$ are independent, we have $\mathbb{E}[D^n_{s_z,1} D^n_{s_z,2}] = 0$. Consequently
$$\mathrm{Var}\left(\frac{1}{N}\sum_{\mu} s_z(Y^{n+1}_\mu)\right) = \mathbb{E}[(D^{n+1}_{s_z})^2] = \mathbb{E}[(D^n_{s_z,1})^2] + \mathbb{E}[(D^n_{s_z,2})^2]. \tag{13}$$
For the first summand, we write
$$\begin{aligned}\int_{\mathcal{X}\times I^d} s_z\circ\varphi_{n+1}(x,u)du dP_n(x) &= \int_{\mathcal{X}}\prod_{i=1}^{d}\psi_{n+1,i}(x_i,z_i)dP_n(x)\\ &= (-1)^d\int_{\mathcal{X}^2}\prod_{i=1}^{d}\frac{\partial\psi_{n+1,i}}{\partial x_i}(x_i,z_i)s_x(y)dP_n(y)dx\end{aligned}$$

and

$$\int_{I^d} s_z \circ \varphi_{n+1}(Y_\mu^n, u)du = (-1)^d \int_{\mathcal{X}} \prod_{i=1}^d \frac{\partial \psi_{n+1,i}}{\partial x_i}(x_i, z_i) s_x(Y_\mu^n)dx.$$

This gives

$$\begin{aligned} D^n_{s_z,1} &= (-1)^d \int_{\mathcal{X}} \left(\prod_{i=1}^d \frac{\partial \psi_{n+1,i}}{\partial x_i}(x_i, z_i) \right) \left(\int_{\mathcal{X}} s_x(y) dP_n(y) - \frac{1}{N} \sum_\mu s_x(Y_\mu^n) \right) dx \\ &= (-1)^d \int_{\mathcal{X}} \left(\prod_{i=1}^d \frac{\partial \psi_{n+1,i}}{\partial x_i}(x_i, z_i) \right) D^n_{s_x} dx. \end{aligned}$$

We then have

$$\begin{aligned} \mathbb{E}[(D^n_{s_z,1})^2] &= \int_{\mathcal{X}^2} \left(\prod_{i=1}^d \frac{\partial \psi_{n+1,i}}{\partial x_i}(x_i, z_i) \prod_{i=1}^d \frac{\partial \psi_{n+1,i}}{\partial x_i}(x'_i, z_i) \right) \mathbb{E}\left[D^n_{s_x} D^n_{s_{x'}} \right] dxdx' \\ &\leq \left(\prod_{i=1}^d M_{n+1,i} \right)^2 \sup_{x \in \mathcal{X}} \mathrm{Var}\left(\frac{1}{N} \sum_\mu s_x(Y_\mu^n) \right). \end{aligned}$$

For the second summand in (13), we have from Lemma 1:

$$\mathbb{E}[(D^n_{s_z,2})^2] \leq \frac{d(\Sigma_{n+1} + 2)}{4N^{1+1/2d}}.$$

With (13), the result is established by induction. □

This result is a third limitation of the SSS method: there is only a small improvement to the MC variance bound for large d's.

The variance of the MC or SSS estimator is bounded for a function of the form s_z, with $z \in \mathcal{X}$. We obtain a bound for a *regular* cost function c by the same reasoning as in Proposition 6 of [11]. Let $\mathcal{C}$ be the class of functions $c \in \mathcal{B}_+(\mathcal{X})$ which are d-times continuously differentiable and such that

1. $V^{(d)}(c) := \int_{\mathcal{X}} \left| \frac{\partial^d c}{\partial x_1 \cdots \partial x_d}(x) \right| dx < +\infty$,
2. there exists a permutation π of $[1, d]$ such that for any $i \in [1, d]$ and $x_{\pi(j)} \in (a_{\pi(j)}, b_{\pi(j)}), 1 \leq j \leq d, j \neq i$,

$$\lim_{x_{\pi(i)} \to b_{\pi(i)}} \frac{\partial^{d-i} c}{\partial x_{\pi(i+1)} \cdots \partial x_{\pi(d)}}(x) = 0.$$

Corollary 1. *If $c \in \mathcal{C}$, then*

$$\mathrm{Var}\left(\frac{1}{N} \sum_\mu c(Y_\mu^n) \right) \leq (V^{(d)}(c))^2 \frac{\beta_n}{N^{1+1/2d}}.$$

Proof. We have

$$D^n_c := \int_{\mathcal{X}} c(x) dP_n(x) - \frac{1}{N} \sum_\mu c(Y_\mu^n) = (-1)^d \int_{\mathcal{X}} \frac{\partial^d c}{\partial x_1 \cdots \partial x_d}(z) D^n_{s_z} dz.$$

Consequently,

$$\begin{aligned} &\mathbb{E}[(D^n_c)^2] \\ &\leq \int_{\mathcal{X}^2} \left| \frac{\partial^d c}{\partial x_1 \cdots \partial x_d}(z) \right| \cdot \left| \frac{\partial^d c}{\partial x_1 \cdots \partial x_d}(z') \right| \sqrt{\mathbb{E}[(D^n_{s_z})^2]} \sqrt{\mathbb{E}[(D^n_{s_{z'}})^2)]} dzdz', \end{aligned}$$

hence the conclusion. □

Remark 1. The previous results can be extended without difficulty as follows. Let d, d' be integers and put $\overline{d} := d + d'$. We consider a discrete time Markov chain $\overline{X}_n := (X_n, X'_n)$ with state space $\overline{\mathcal{X}} := \mathcal{X} \times \mathcal{X}'$, where $\mathcal{X} = \prod_{i=1}^{d} \mathcal{X}_i$, $\mathcal{X}' = \prod_{i=d+1}^{\overline{d}} \mathcal{X}_i$, with $\mathcal{X}_i = (a_i, b_i)$ and $-\infty \le a_i < b_i \le +\infty$, for $1 \le i \le \overline{d}$. Here X_0 (resp. X'_0) is a d-dimensional (resp. d'-dimensional) random variable and the chain evolves according to the recurrence:

$$X_{n+1} = \varphi_{n+1}(X_n, U_{n+1}),\ X'_{n+1} = \varphi'_{n+1}(X_n, X'_n, U_{n+1}), \quad n \ge 0. \tag{14}$$

The variables $U_1, U_2, \ldots$ are i.i.d. uniform random variables over I^d; the functions $\varphi_{n+1} : \mathcal{X} \times I^d \to \mathcal{X}$ and $\varphi'_{n+1} : \overline{\mathcal{X}} \times I^d \to \mathcal{X}'$ are measurable. We put $\overline{\varphi}_{n+1}(\overline{x}, u) := (\varphi_{n+1}(x, u), \varphi'_{n+1}(x, x', u))$, for $\overline{x} := (x, x') \in \overline{\mathcal{X}}$. We denote $\overline{P}_n$ the law of $\overline{X}_n$. We suppose that φ_{n+1} satisfies (2) and we assume (A1–A3) for any $i \in [1, d]$. We suppose that $\overline{P}_0$ has density $\overline{f}_0$. For $z \in \mathcal{X}$, we denote $\overline{s}_z$ the indicator function of $\prod_{i=1}^{d}(a_i, z_i) \times \mathcal{X}'$.

The SSS scheme is as follows. We choose $N = p^{\overline{d}+d}$, for some integer p. For $\mu = (m_1, \ldots, m_{\overline{d}}) \in [1, p]^{\overline{d}-1} \times [1, p^{d+1}]$, we set: $L_\mu := \prod_{i=1}^{\overline{d}-1}[(m_i - 1)/p, m_i/p) \times [(m_{\overline{d}} - 1)/p^{d+1}, m_{\overline{d}}/p^{d+1})$. Then, for $u \in I^{\overline{d}}$, the index $\mu(u) \in [1, p]^{\overline{d}-1} \times [1, p^{d+1}]$ is such that $u \in L_{\mu(u)}$. For initialization, we assume that we can define a set $\{\overline{Y}^0_\mu : \mu \in [1, p]^{\overline{d}-1} \times [1, p^{d+1}]\}$ of independent random variables with the following properties.

P1. For any $\overline{s} \in \mathcal{B}_+(\overline{\mathcal{X}})$,

$$\mathbb{E}\left[\frac{1}{N}\sum_\mu \overline{s}(\overline{Y}^0_\mu)\right] = \int_{\overline{\mathcal{X}}} \overline{s}(\overline{x}) d\overline{P}_0(\overline{x}).$$

P2. There exists some $\overline{\beta}_0 > 0$ such that, for every $z \in \mathcal{X}$,

$$\mathrm{Var}\left(\frac{1}{N}\sum_\mu \overline{s}_z(\overline{Y}^0_\mu)\right) \le \frac{\overline{\beta}_0}{N^{1+1/(\overline{d}+d)}}.$$

We suppose that we have generated a set of random variables $\{\overline{Y}^n_\mu : \mu \in [1, p]^{\overline{d}-1} \times [1, p^{d+1}]\}$; the transition from n to $n+1$ acts as follows. The numbering is modified so that (7) is satisfied for $1 \le i \le \overline{d}$. For $\kappa = (k_1, \ldots, k_{\overline{d}+d}) \in [1, p]^{\overline{d}+d}$, let $J_\kappa := \prod_{i=1}^{\overline{d}+d}[(k_i - 1)/p, k_i/p)$. Let $\{V^{n+1}_\kappa : \kappa \in [1, p]^{\overline{d}+d}\}$ be independent random variables with $V^{n+1}_\kappa \sim \mathfrak{U}(J_\kappa)$. For $u \in I^{\overline{d}+d}$, let $\dot{u} := (u_1, \ldots, u_{\overline{d}})$ and $\ddot{u} := (u_{\overline{d}+1}, \ldots, u_{\overline{d}+d})$. We define $\overline{Y}^{n+1}_\kappa = \overline{\varphi}_{n+1}(\overline{Y}^n_{\mu(\dot{V}^{n+1}_\kappa)}, \ddot{V}^{n+1}_\kappa)$. The following holds, with $\overline{\beta}_{n+1} = \Pi^2_{n+1}\overline{\beta}_n + (\overline{d}\Sigma_{n+1} + 2d)/4$.

1. For any $\overline{s} \in \mathcal{B}_+(\overline{\mathcal{X}})$,

$$\mathbb{E}\left[\frac{1}{N}\sum_\mu \overline{s}(\overline{Y}^n_\mu)\right] = \int_{\overline{\mathcal{X}}} \overline{s}(\overline{x}) d\overline{P}_n(\overline{x}).$$

2. For any $z \in \mathcal{X}$,

$$\mathrm{Var}\left(\frac{1}{N}\sum_\mu \overline{s}_z(\overline{Y}^n_\mu)\right) \le \frac{\overline{\beta}_n}{N^{1+1/(\overline{d}+d)}}. \tag{15}$$

4 Numerical experiments

In this section, we compare classical MC, SSS and Array-RQMC approaches for the simulation of multi-dimensional Markov chains. The Array-RQMC method gives several possibilities for the choices of QMC point sets, randomization and sorting strategy. The A-Strat version uses stratified samples, while the A-Sobol version employs Sobol′ sequences. For both schemes, a multivariate batch sort is done. In dimension d, with N states, an A-method with sorting parameters $(\alpha_1, \ldots, \alpha_d) \in (0, 1)^d$

(satisfying $\alpha_1 + \cdots + \alpha_d = 1$) firstly sorts the set of N states by their first coordinate in N^{α_1} subsets of size $N^{1-\alpha_1}$; then it sorts each subset in N^{α_2} subsets of size $N^{1-\alpha_1-\alpha_2}$ by their second coordinate, etc. Note that, for a state space of dimension d, the SSS method corresponds to the sorting parameters $\alpha_1 = \cdots = \alpha_{d-1} = 1/(2d), \alpha_d = (d+1)/(2d)$. The extension presented in Remark 1 for a state space of dimension $\overline{d}$ corresponds to the sorting parameters $\alpha_1 = \cdots = \alpha_{\overline{d}-1} = 1/(\overline{d}+d), \alpha_{\overline{d}} = (d+1)/(\overline{d}+d)$. In the following, A-Strat and A-Sobol are tested with uniform sorting parameters and also with SSS parameters.

We calculate the empirical variance of the estimators of

$$\int_{\mathbb{R}^d} s(x) dP_n(x) \quad \text{or} \quad \int_{\mathbb{R}^{\overline{d}}} \overline{s}(\overline{x}) d\overline{P}_n(\overline{x}),$$

for some s or $\overline{s}$. We plot this variance as a function of the number N of simulated chains. Assuming a model $KN^{-\gamma}$ for the variance, we can estimate the rate γ by linear regression and compare it with the theoretical bounds. Thereafter, we denote by φ and Φ the probability density function (pdf) and cdf of the standard normal distribution, respectively.

4.1 Diffusion

We consider the initial value problem for the d-dimensional diffusion equation:

$$\frac{\partial f}{\partial t}(x,t) = \nu \Delta f(x,t), \ x \in \mathbb{R}^d, \ t > 0 \quad \text{and} \quad f(x,0) = f_0(x), \ x \in \mathbb{R}^d,$$

with constant diffusivity $\nu > 0$. Here Δ is the Laplacian. We assume that the initial data satisfies $f_0 \geq 0$ and $\int_{\mathbb{R}^d} f_0(x)dx = 1$. Then, for any $t > 0$, it holds that $\int_{\mathbb{R}^d} f(x,t)dx = 1$. Let G be the fundamental solution of the heat operator:

$$G(x,t) := \frac{1}{(4\pi\nu t)^{d/2}} \mathrm{e}^{-\|x\|^2/4\nu t}, \quad x \in \mathbb{R}^d, \ t > 0,$$

where $\|x\|$ denotes the Euclidean norm of $x \in \mathbb{R}^d$. Then, for any $\tau \geq 0$, $f(x,t) = \int_{\mathbb{R}^d} G(x-w, t-\tau) f(w,\tau)dw$, for $x \in \mathbb{R}^d$ and $t > \tau$. If Δt is a time step, we set $t_n := n\Delta t$ and $f_n(x) := f(x, t_n)$. It follows that

$$f_{n+1}(x) = \frac{1}{(2\nu\Delta t)^{d/2}} \int_{\mathbb{R}^d} \prod_{i=1}^{d} \varphi\left(\frac{x_i - w_i}{\sqrt{2\nu\Delta t}}\right) f_n(w)dw.$$

Consequently, for any $s \in \mathcal{B}_+(\mathbb{R}^d)$,

$$\begin{aligned} &\int_{\mathbb{R}^d} s(x) f_{n+1}(x)dx \\ &= \int_{\mathbb{R}^d \times I^d} s(x_1 + \sqrt{2\nu\Delta t}\Phi^{-1}(u_1), \ldots, x_d + \sqrt{2\nu\Delta t}\Phi^{-1}(u_d)) f_n(x)dxdu. \end{aligned}$$

We define the Markov chain as in (1). Let X_0 have pdf f_0 and let

$$X_{n+1,i} = X_{n,i} + \sqrt{2\nu\Delta t}\Phi^{-1}(U_{n+1,i}), \quad 1 \leq i \leq d,$$

where $U_{n+1} \sim \mathfrak{U}(I^d)$. This defines a *random walk* method [7]. Here, Assumption (2) and hypotheses A1–A3 are satisfied.

As presented, the method is artificial, since we know the exact solution. This algorithm is a part of fractional step schemes when one considers problems involving a combination of convection, reaction and diffusion. In this experiment, we take $\nu = 1$ and define f_0 as the indicator function of the interval $[-1/2, 1/2]^d$. We choose $\Delta t = 0.001$ and $T = 0.01$. We compute the empirical variance (with $M = 100$ replications) of the estimate of $\int_{[0,1)^d} f(x,T)dx$. Here, s is the indicator function of $[0,1)^d$.

For $d = 2$, Figure 2 (left) shows the variance of MC and SSS as a function of N, for $N = 4^4, 5^4, \ldots, 50^4$ (in $\log_{10}$ scales). The calculations are also done with A-Strat and A-Sobol. The sorting parameters $(1/2, 1/2)$ and $(1/4, 3/4)$ are tested and the results only show a small difference between them. For both A-methods, the best choice is drawn. For $d = 3$, the variance of MC and SSS as a function of N, for $N = 3^6, 4^6, \ldots, 13^6$ (in $\log_{10}$ scales) is shown on Figure 2 (right). The calculations are also done using A-Strat and A-Sobol, with the sorting parameters $(1/3, 1/3, 1/3)$ and $(1/6, 1/6, 4/6)$, without a great difference between the results; only the best choices are drawn. The regression estimates of the convergence rate γ are given in Table 1, for $d = 2$ (second row) and $d = 3$ (third row). A-method refers to uniform sorting parameters and *A-method* (italicized) to SSS parameters. The rates of the upper bounds given in Section 3 are indicated in parenthesis. SSS and Array-RQMC produce smaller variance than classical MC (for the same N). The variances of Array-RQMC are smaller than those of SSS; nevertheless, the convergence rates of SSS are equal or slightly better than those of Array-RQMC. For SSS, the regression estimates of the rate γ somehow overtake the rates of the theoretical upper bounds.

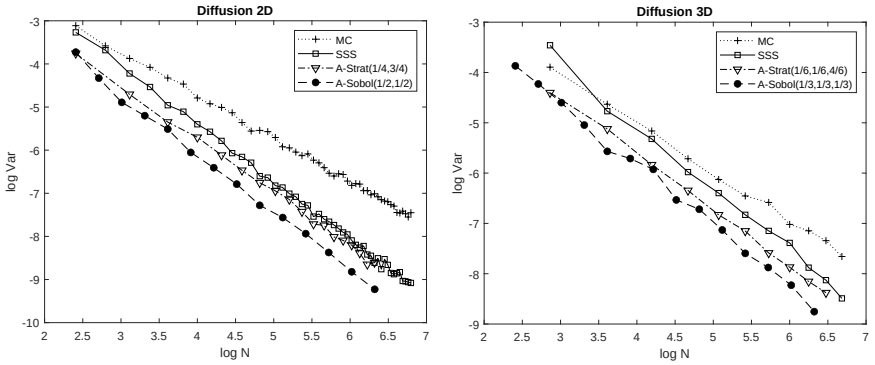

Figure 2: Random walk for diffusion: empirical variance of 100 estimations of $\int_{[0,1)^d} f(x,T)dx$ as a function of N for MC, SSS, A-Strat and A-Sobol in dimension $d = 2$ (left) and $d = 3$ (right) ($\log_{10}$-$\log_{10}$ scale)

4.2 An asian option

In the Black-Scholes model, under the risk-neutral measure, the asset price S_t at time t is given by: $S_t = S_0 \exp((r - \sigma^2/2)t + \sigma B_t)$, where r is the risk-free interest rate, σ the volatility parameter and B is a standard Brownian motion. Let T be the maturity date and $0 < t_1 < \cdots < t_q = T$ be observation times. If K is a strike price, we want to estimate the value of the call option:

$$CO = \mathrm{e}^{-rT}\mathbb{E}\left[\left(\frac{1}{q}\sum_{n=1}^{q} S_{t_n} - K\right)_+\right].$$

We refer to [8] for further details. If $\Delta t_{n+1} := t_{n+1} - t_n$, we have $S_{t_{n+1}} = S_{t_n} \exp((r - \sigma^2/2)\Delta t_{n+1} + \sigma(B_{t_{n+1}} - B_{t_n}))$, for $0 \le n < q$. By setting $X_0 := S_0$, $X_0' := 1$ and $X_n := S_{t_n}$, $X_n' := (1/n)\sum_{m=1}^{n} S_{t_m}$, for $n \ge 1$, this may be written as in (14):

$$\begin{aligned} X_{n+1} &= X_n \mathrm{e}^{(r-\sigma^2/2)\Delta t_{n+1} + \sigma\sqrt{\Delta t_{n+1}}\Phi^{-1}(U_{n+1})}, \\ X_{n+1}' &= \frac{1}{n+1}\left(X_n \mathrm{e}^{(r-\sigma^2/2)\Delta t_{n+1} + \sigma\sqrt{\Delta t_{n+1}}\Phi^{-1}(U_{n+1})} + nX_n'\right), \end{aligned}$$

where $U_{n+1} \sim \mathfrak{U}(I)$. Here, the state space is $(0, +\infty)$; $d = d' = 1$ and hypotheses A1–A3 are satisfied.

In this example, we choose $S_0 = 100$, $r = \mathrm{Ln}(1.09)$, $\sigma = 0.2$, $K = 90$, $T = 240/365$, $\Delta t = 1/365$, $q = 60$ and $t_n = T - (q - n)\Delta t$ for $n \in [1, q]$. We compare the variances of the MC, SSS, A-Strat and A-Sobol estimates of CO: we replicate the calculation independently 100 times and we compute the sample variance.

Figure 3 shows the variance of MC and SSS as a function of N (in $\log_{10}$ scales), for $N = 10^3, 15^3, \ldots, 100^3$. For the calculations done with A-Strat and A-Sobol, the sorting parameters $(1/2, 1/2)$ and $(1/3, 2/3)$ are tested. While the differences between the A-Strat versions are small, this is not the case for the A-Sobol versions. The regression estimates of the convergence rate γ are given in Table 1 (fourth row). SSS and Array-RQMC result in smaller variance than classical MC (for the same N). Here, A-Sobol with uniform sorting parameters outperforms the other schemes. With a similar convergence rate, A-Strat gives smaller variance than SSS. In Section 3 we have established for SSS a $\mathcal{O}(N^{-4/3})$ upper bound for the variance of the estimator with a cost function of the form s_z, for $z > 0$: see (15). But we did not prove a bound for the variance of a SSS estimator of CO.

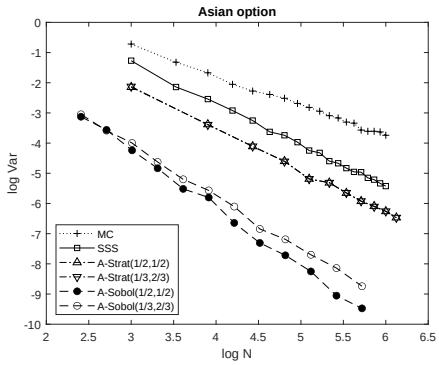

Figure 3: Asian call option: empirical variance of 100 estimates of CO as a function of N for MC, SSS, A-Strat and A-Sobol ($\log_{10}$-$\log_{10}$ scale)

Table 1: Estimation of the convergence rate γ of the sample variance: comparison of classical MC, SSS, A-Strat and A-Sobol approaches. Italicized names correspond to SSS sorting parameters. The rates of the upper bounds are indicated in parenthesis.

Experiment	MC	SSS	A-Strat	*A-Strat*	A-Sobol	*A-Sobol*
Diffusion 2D	1.00 (1.00)	1.34 (1.25)	1.19	1.24	1.36	1.28
Diffusion 3D	0.97 (1.00)	1.25 (1.17)	1.10	1.12	1.21	1.16
Asian option	0.99 (1.00)	1.37	1.38	1.39	1.95	1.72

5 Conclusion

We consider Markov chain models with a $(d + d')$-dimensional continuous state space. We assume that the first d components of the chain are advanced independently from each others and that only d random variates are used to advance by one step. We analyze two approaches for the simulation: classical MC and a method using SS samples. Upper bounds on the variance of an estimator for a cost function which only depends on the first d variables are proved. When N copies of the chain are simulated, the order is $\mathcal{O}(N^{-1})$ for MC and our upper bound is $\mathcal{O}(N^{-1+1/(2d+d')})$ for the stratified strategy. In our numerical experiments, the SSS variance decreased a bit faster than the bound. The Array-RQMC method gave a lower variance, but we unfortunately have no proof (so far) of the better convergence rate for that method. Interesting topics for further research include the extension of our analysis to more general cost functions and to other stratified approaches, and proving the better convergence rates for Array-RQMC in more than one dimension.

References

[1] Ben Abdallah, A., L'Ecuyer, P., Puchhammer, F.: Array-RQMC for option pricing under stochastic volatility models. In: Mustafee, N., Bae, K.-H.G., Lazarova-Molnar, S., Rabe, M., Szabo, C., Haas, P., Son, Y.-J. (eds.) Proceedings of the 2019 Winter Simulation Conference, pp. 440–451. IEEE Press (2019)

[2] Cheng, R.C.H., Davenport, T.: The problem of dimensionality in stratified sampling. Manage. Sci. 35, 1278–1296 (1989)

[3] El Haddad, R., Fakhreddine, R., Lécot, C.: Stratified Monte Carlo integration. In: Sabelfeld, K.K., Dimov, I. (eds.) Monte Carlo Methods and Applications, pp. 105–113. De Gruyter, Berlin (2013)

[4] El Haddad, R., Fakhereddine, R., Lécot, C., Venkiteswaran, G.: Extended Latin hypercube sampling for integration and simulation. In: Dick, J., Kuo, F.Y., Peters, G.W., Sloan, I.H. (eds.) Monte Carlo and Quasi-Monte Carlo Methods 2012, pp. 317–330. Springer-Verlag, Berlin (2013)

[5] El Haddad, R., El Maalouf, J., Lécot, C., L'Ecuyer, P.: Sudoku Latin square sampling for Markov chain simulation. In: Tuffin, B., L'Ecuyer, P. (eds.) Monte Carlo and Quasi-Monte Carlo Methods, pp. 207–230. Springer, Cham (2020)

[6] El Haddad, R., El Maalouf, J., Fakhereddine, R., Lécot, C.: Simulation of Markov chains with continuous state space by using simple stratified and Sudoku Latin square sampling. In: Botev Z., Keller, A., Lemieux, C., Tuffin, B. (eds.) Advances in Modeling and Simulation, pp. 239–260. Springer, Cham (2022)

[7] Ghoniem, A.F., Sherman, F.S.: Grid-free simulation of diffusion using random walk methods. J. Comput. Phys. 61, 1–37 (1985)

[8] Glasserman, P.: Monte Carlo Methods in Financial Engineering. Springer, New York (2004)

[9] Haber, S.: A modified Monte-Carlo quadrature. Math. Comput. 20, 361–368 (1966)

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

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

[12] L'Ecuyer, P., Lécot, C., L'Archevêque-Gaudet, A.: On array-RQMC for Markov chains: Mapping alternatives and convergence rates. In: L'Ecuyer, P., Owen, A.B. (eds.) Monte Carlo and Quasi-Monte Carlo Methods 2008, pp. 485–500. Springer-Verlag, Berlin (2009)

[13] L'Ecuyer, P., Munger, D., Lécot, C., Tuffin, B.: Sorting methods and convergence rates for Array-RQMC: Some empirical comparisons. Math. Comput. Simul. 143, 191–201 (2018)

[14] Lemieux, C.: Monte Carlo and Quasi-Monte Carlo Sampling. Springer, New York (2009)

[15] Puchhammer, F., Ben Abdellah, A., L'Ecuyer, P.: Variance reduction with Array-RQMC for tau-leaping simulation of stochastic biological and chemical reaction networks. Bull. Math. Biol. 83, 91 (2021)