**Numerical Pricing of Contingent
Claims on Multiple Assets and/or
Factors – A Low-Discrepancy
Markov Chain Approach**

J.-C. Duan, G. Gauthier
J.-G. Simonaro

Numerical Pricing of Contingent Claims on Multiple Assets and/or Factors – A Low-Discrepancy Markov Chain Approach

Jin-Chuan Duan

Rotman School of Management, University of Toronto
and Department of Finance
Hong Kong University of Science & Technology
jcduan@rotman.utoronto.ca

Geneviève Gauthier

GERAD and Méthodes quantitatives de gestion
HEC Montréal
Montréal (Québec) Canada, H3T 2A7
genevieve.gauthier@hec.ca

Jean-Guy Simonato

Service de l'enseignement de la finance
HEC Montréal
Montréal (Québec) Canada, H3T 2A7
jean-guy.simonato@hec.ca

November 2004

Abstract

We develop a Markov chain pricing method capable of handling several state variables. The Markov chain construction of Duan and Simonato (2000) is modified so that higher-dimensional valuation problems can be dealt with. Their design relies on a Cartesian product, which grows in a power fashion as the number of assets/factors increases. We use a multi-dimensional low-discrepancy sequence as the building block for constructing the Markov chain in order to take advantage of the high degree of uniformity inherent in such sequences. Our design contains two critical components. First, we have devised a way of computing analytically the entries of the transition probability matrix and then shown that such a Markov chain converges weakly to the target Markov process. Second, we have utilized an elliptical restriction as a way of removing non-critical components of the Markov chain to enhance the computational efficiency. Numerical examples are provided.

Résumé

Nous développons une méthode de tarification basée sur les chaînes de Markov qui peut s'accommoder de plusieurs variables d'états. La chaîne de Markov proposée par Duan et Simonato (2002) est modifiée de façon à ce que les problèmes d'évaluation de plus grande dimension puisse être traités. Leur design est basé sur un produit Cartésien qui croît de façon exponentielle lorsque le nombre de facteurs augmente. Nous utilisons une suite à discrépence faible multi-dimensionnelle afin de construire notre chaîne de Markov. Notre design contient deux composantes importantes. Premièrement, les entrées de la matrice de transition sont calculées de façon analytique et nous pouvons alors démontrer qu'une telle chaîne de Markov converge faiblement vers le processus markovien cible. Deuxièmement, nous avons utilisé une restriction elliptique afin de retirer les états de la chaîne qui ont une très faible probabilité d'être atteint. Ceci permet d'améliorer l'efficacité de la méthode. Des exemples numériques sont fournis.

1 Introduction

For the valuation of American-style derivatives on several underlying assets and/or factors, the only practical approach known is Monte Carlo simulation. However, even when considering the recent developments in this literature, (Broadie, Glasserman and Jain, 1997; Boyle, Kolkiewicz and Tan, 2000; Garcia, 2000; Longstaff and Schwartz, 2000), the Monte Carlo approach appears to be of limited applicability to derivatives on several assets and/or factors with many early exercise possibilities. Recently, lattice constructions for higher dimensional valuation problems have also been attempted in MaCarthy and Webber (1999) and Chen and Yang (2000) by applying some geometric restrictions to the multidimensional state partition. Due to their highly restrictive nature, these methods appear to have limited potential for models beyond the constant volatility geometric Brownian motion.

In this paper we develop a Markov chain pricing method capable of handling several state variables. It can be used in the traditional geometric Brownian motion context as well as for more complex pricing models such as stochastic volatility and GARCH. We use the Markov chain pricing idea of Duan and Simonato (2000) but modify the specific Markov chain construction so that higher-dimensional valuation problems can be dealt with. In a nutshell, Duan and Simonato's (2000) design relies on a Cartesian product, which grows in a power fashion as the number of assets/factors increases. For practical purposes, Duan and Simonato's (2000) specific design is workable up to two dimensions due to the memory constraint of typical computers. Our modification abandons the use of Cartesian product in handling additional dimensions. Instead, we use a multi-dimensional low-discrepancy sequence as the building block for constructing the Markov chain. In other words, we want to take advantage of the high degree of uniformity inherent in low-discrepancy sequences. There are several challenges related to this redesign. The first and foremost is to figure out a way of computing analytically the entries of the transition probability matrix and then to show that such a Markov chain converges weakly to the target Markov process underlying the pricing model. It is also beneficial to remove non-critical components of the Markov chain so that the computational speed can be improved. Indeed, we show that our low-discrepancy Markov chain has the desirable weak convergence property. We also employ an elliptical restriction and control variables to enhance computational efficiency of the low-discrepancy Markov chain.

The idea of partitioning the state space using a low-discrepancy sequence is not new. To our knowledge, Broadie (1999) was the first to introduce such an idea into a lattice design. Our method differs from Broadie (1999) by relying on the Markov chain idea of Duan and Simonato (2000) instead of using a lattice as in Broadie (1999). Moreover, the transition probabilities for the Markov chain are obtained analytically by inferring directly

from the target process as opposed to relying on an optimization procedure which is likely to be time-consuming and to confront potentially competing objective functions.

Our numerical findings reveal that the Markov chain algorithm performs well and converge to the target benchmark prices. The convergence to the benchmark prices is faster for larger time steps or shorter maturities. As the number of assets increases, the dimension of the chain needs to be increased substantially to achieve convergence. However, sparse matrix techniques allow for the use of Markov chains with a large dimension because in most cases, many elements of the transition probability matrix are effectively zero.

2 Markov chain approach

Consider the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_{t^*} : t^* \in \mathcal{T}\}, P)$ where $\mathcal{T}$ is a set of indices such as $[0, \infty)$, $[0, T^*]$, $\{0, 1, ..., T\}$, etc. Denote the time-t^* value of the American contingent claim by $v(\mathbf{S}_{t^*}, t^*)$ where $\{\mathbf{S}_{t^*} = (S_1(t^*), ..., S_d(t^*)) : t^* \in \mathcal{T}\}$ is a d-dimensional stochastic process that captures all pricing-relevant informations of the contingent claim: underlying assets, interest rates, exchange rates, convenience yield, volatility, etc. The typical derivative pricing theory implies that, under the hypothesis of no-arbitrage opportunity, there is an equivalent pricing measure Q with which the discounted values of the tradable assets are martingales. The logical consequence is that replicable contingent claims can be priced by simply taking an expectation under Q of their associated discounted payoffs.

The early exercise possibility related to Bermudan/American style derivatives creates complication in practice. In principle, one can invoke the dynamic programming technique to compute (or approximate) the value of the Bermudan (American) contingent claim recursively:

$$v(\mathbf{S}_{t^*}, t^*) = \max\left\{h(\mathbf{S}_{t^*}, t^*), e^{-r_{t^*}} \mathrm{E}_{t^*}^Q[v(\mathbf{S}_{t^*+\tau}, t^* + \tau)]\right\}, \ t^* \in \{0, \tau, 2\tau, ..., (T-1)\tau\}$$

$$v(\mathbf{S}_{T^*}, T^*) = h(\mathbf{S}_{T^*}, T^*)$$

$$(1)$$

where r_{t^*} is the one-period (from time t^* to $t^* + \tau$) risk-free interest rate (continuously compounded); $h(\mathbf{S}_{t^*}, t^*)$ is the time-t^* payoff of the contingent claim if exercised; T^* is the maturity date of the contingent claim; τ is the length of the time periods in this discrete-time setting[1]; $T = T^*/\tau$ is the number of time periods before the expiration of the contingent claim; and $\mathrm{E}_{t^*}^Q[\bullet]$ is a short-hand notation standing for the conditional expectation $\mathrm{E}^Q[\bullet|\mathcal{F}_{t^*}]$.

---

[1]In the following, the time is marked by a star while the number of time periods is represented without such superscript. This distinction is needed since we use a discrete-time setting to approximate a time-continuous model and we need to keep the liberty of choosing the length τ of the time periods to suit the caracteristic of the contingent claim.

The major difficulty in implementing (1) is to compute the conditional expectation. The approach we propose is to replace the stochastic process $\mathbf{S}$ by a discrete-state process $\mathbf{S}^{(n)}$ for which the conditional expectation is easy to compute. Since the methodology is based on a Markov chain, we need to assume that there is a time-homogeneous Markovian process $\mathbf{X}$ hidden behind $\mathbf{S}$. More formally, this approach assumes the existence of a function $g : \mathbb{R}^d \times \mathcal{T} \to \mathbb{R}^{d^*}$ such that 1) the d^*-dimensional process $\mathbf{X} = \{\mathbf{X}_{t^*} : t^* \in \mathcal{T}\}$ where $\mathbf{X}_{t^*} = (X_1(t^*), ..., X_{d^*}(t^*)) = g(\mathbf{S}_{t^*}; t^*)$ is a time-homogeneous Markovian process under Q and 2) for any fixed t^*, $g(\bullet; t^*)$ is invertible ensuring that $\mathbf{S}_{t^*} = g^{-1}(\mathbf{X}_{t^*}; t^*)$. An illustration of this transformation is given in Section 5.1.

Then, the stochastic process $\mathbf{X}$ is approximated by constructing a $n-$states Markov chain

$$\left\{ \mathbf{X}_t^{(n)} : t \in \{0, 1, 2, ..., T\} \right\}$$

so that $\mathbf{X}_t^{(n)}$ converges to $\mathbf{X}_{t\tau}$ weakly for any $t \in \{0, 1, 2, ..., T\}$ as the number n of states increases to infinity. Because the prices of the contingent claims obtained from the Markov chain have to converge to their respective theoretical price (given by the model $\mathbf{X}$), we also want that $\phi\left(\mathbf{X}_t^{(n)}, t\tau\right)$ converges to $\phi(\mathbf{X}_{t\tau}, t\tau)$ in mean (as $n \to \infty$) for a reasonable class of payoff functions ϕ.

Under the Markov chain approximation, the dynamic programming recursion can be handled with

$$\overrightarrow{v}_t^{(n)} = \max\left[h\left(g^{-1}(\mathbf{X}^{(n)}; t\tau), t\tau\right), e^{-r_t\tau} \mathbf{\Pi}^{(n)} \overrightarrow{v}_{t+1}^{(n)} \right], \ t \in \{0, 1, 2, ..., (T-1)\}$$

$$\overrightarrow{v}_T^{(n)} = h\left(g^{-1}(\mathbf{X}^{(n)}; T\tau), T\tau\right)$$

(2)

where $\mathbf{X}^{(n)}$ is an $n \times d$ matrix that contains each of the n possible states of the Markov chain (each state belongs to $\mathbb{R}^d$); $h\left(g^{-1}(\mathbf{X}^{(n)}; t), t\tau\right)$ is also a $n \times 1$ vector filled with payoffs of the contingent claim, if exercised at time $t^* = t\tau$, corresponding to n states of the chain; $\mathbf{\Pi}^{(n)}$ is the time-homogeneous $n \times n$ transition probability matrix over a time period of length τ, and the $n \times 1$ vector $\overrightarrow{v}_t^{(n)}$ contains, for each of the possible states, the time $t^* = t\tau$ values of the contingent claim. Note that $\mathrm{E}_{t^*}^Q[v(\mathbf{S}_{t^*+\tau}, t^* + \tau)]$ is being approximated by $\mathbf{\Pi}^{(n)} \overrightarrow{v}_{t+1}^{(n)}$, a simple matrix operation. After the recursion is completed, the element of $\overrightarrow{v}_0^{(n)}$ corresponding to the initial state of the Markov system is the model value of the contingent claim.

Although weak convergence of the Markov chain approximation to the target process is crucial, it is not enough to ensure price convergence. Due to the payoff function of the derivative contract and the transformation used in obtaining time homogeneity, the weak

convergence property may not be carried over to $h\left(g^{-1}(\mathbf{X}^{(n)}\left(t^{*}\right);t^{*}),t^{*}\right)$, and such a quantity may also lack uniform integrability so that the derivative's price cannot be ensured to converge to the right value. Since $g^{-1}\left(\bullet;t\right)$ and h are usually continuous functions, the weak convergence of $h\left(g^{-1}(\mathbf{X}^{(n)}\left(t\right);t)\right)$ to $h\left(g^{-1}(\mathbf{X}\left(t\tau\right);t\tau)\right)$ is easy to obtain. Convergence in mean, i.e., price convergence, requires more work. The issue of convergence will be addressed in detail later.

In constructing the weakly converging Markov chain, Duan and Simonato (2000) employed a Cartesian product construction. In other words, for the one-dimensional problem, the state space is partitioned into, say, m discrete states, and for the two-dimensional problem, the state space is partitioned into mk discrete states with k being the number of states in the second dimension. Relating to the Markov chain notation, $n = m$ for the one-dimensional problem and $n = mk$ for the two-dimensional problem. The Cartesian product construction method is appealing because it allows for a rather simple way of computing the entries of $\mathbf{\Pi}^{(n)}$. Although the Cartesian product approach is easier to construct and comprehend, it is rather impractical once the dimension of the valuation problem is increased to three or four because the number of states grows in a power fashion. It is worth noting that this problem is not unique to Duan and Simonato's (2000) approach. The traditional lattice and finite difference methods all face a similar problem.

This paper is set out to devise a Markov chain that is capable of dealing with higher-dimensional valuation problems while retains the convenience and power of the Markov chain approach. We opt for a construction that relies on points derived from a low-discrepancy sequence instead of employing a Cartesian product of partitions one dimension at a time.

3   Low-discrepancy Markov chain

Let $f(\mathbf{z}|\mathbf{x})$ be the one-period density function of the random vector $\mathbf{X}\left(t+1\right)$ conditional on $\mathbf{X}\left(t\right) = \mathbf{x}$. We assume this density function is independent of time so that a time-homogenous Markov chain can be constructed.

In the following, we will construct a sequence $\left\{\mathbf{X}^{(n)} = \left\{\mathbf{X}^{(n)}\left(t\right) : t \in \{0,1,...,T\}\right\} : n \in \mathbb{N}\right\}$ of Markov chains. For each n, the state space and the transition matrix need to be defined. First, consider the d-dimensional hyper-rectangle

$$\mathcal{R}_n = \prod_{i=1}^{d}\left[a_i^{(n)}, b_i^{(n)}\right] \subset \mathbb{R}^d \tag{3}$$

over which the n states of $\mathbf{X}^{(n)}$ will be uniformly scattered. Because we want to achieve the weak convergence of $\mathbf{X}^{(n)}\left(t\right)$ to $\mathbf{X}\left(t\right)$, we need that $\mathcal{R}_n \rightarrow \mathbb{R}^d$ as $n \rightarrow \infty$. Therefore

we assume that for each i, the sequence $\left\{a_i^{(n)} : n \in \mathbb{N}\right\}$ decreases to $-\infty$ and the sequence $\left\{b_i^{(n)} : n \in \mathbb{N}\right\}$ increases to ∞ so that the sequence $\{\mathcal{R}_n : n \in \mathbb{N}\}$ of hyper-rectangles in $\mathbb{R}^d$ satisfies

$$\mathcal{R}_1 \subset \mathcal{R}_2 \subset \cdots \subset \mathcal{R}_n \subset \cdots \text{ and } \mathcal{R}_n \uparrow \mathbb{R}^d \text{ as } n \uparrow \infty.$$

The Lebesgue measure (or volume) of the hyper-rectangle is

$$\lambda\left(\mathcal{R}_n\right) \equiv \prod_{i=1}^{d} \left(b_i^{(n)} - a_i^{(n)}\right). \tag{4}$$

If we take any particular subset A of $\mathbb{R}^d$, then we want the number of states contained in A to increase as n tends to infinity. Therefore, we need the volume of the hyper-rectangle to go up to ∞ at a rate slower than n raised to an appropriate power to reflect the dimension of the problem, as n approaches ∞. This condition is formalized as follows.

Partition Condition: (1) $\mathcal{R}_n \uparrow \mathbb{R}^d$ as $n \uparrow \infty$, and (2) $\lambda\left(\mathcal{R}_n\right) \frac{(\ln n)^d}{n} \to 0$ as $n \to \infty$.

The form of condition (2) comes from the discrepancy of the particular point-set used to construct the state space of the Markov chain. If the support of the distribution is finite, denoted by $\mathcal{S}$, then $a_i^{(n)}$ and $b_i^{(n)}$ only need to be bounded sequences as long as $\mathcal{S} \subset \mathcal{R}_n$ for some large n. It is fairly easy to find a specific construction that satisfies the partition condition. For example, if the one-period standard deviation σ_i of the ith component of $\mathbf{X}$ is state independent, then we may take $a_i^{(n)} = -2\sigma_i \ln n$ and $b_i^{(n)} = 2\sigma_i \ln n$. In that case,

$$\lambda\left(\mathcal{R}_n\right) \frac{(\ln n)^d}{n} = \left(\prod_{i=1}^{d} 4\sigma_i\right) \frac{(\ln n)^{2d}}{n} \to 0 \text{ as } n \to \infty,$$

and $\mathcal{R}_n \to \mathbb{R}^d$ as $n \to \infty$.

We consider n points (d-dimensional) uniformly scattered over the hyper-rectangle $\mathcal{R}_n$. Denote these points by $\mathbf{x}_1^{(n)}, ..., \mathbf{x}_n^{(n)}$. The low-discrepancy Markov chain $\mathbf{X}^{(n)}$ has these n points forming its state space and the transition probability matrix $\mathbf{\Pi}^{(n)}$ has its (i, j) entry equal to

$$\pi_{i,j}^{(n)} = \frac{f\left(\mathbf{x}_j^{(n)} \middle| \mathbf{x}_i^{(n)}\right)}{\sum_{k=1}^{n} f\left(\mathbf{x}_k^{(n)} \middle| \mathbf{x}_i^{(n)}\right)}. \tag{5}$$

It is clear that $\mathbf{\Pi}^{(n)}$ constitutes a proper transition probability matrix because its entries are all non-negative and each row sums to 1. The key task later is to show that under the

partition condition, $\mathbf{X}^{(n)}(t)$ converges (weakly) to the target d-dimensional process $\mathbf{X}(t)$ for $t \in \{0, 1, 2, ..., T\}$.

In order to find n uniformly distributed d-dimensional points in $\mathcal{R}_n$, we rely on low-discrepancy sequences. To facilitate the discussion, we formally define the notion of discrepancy.

Definition 1 *For a sequence* $\mathbf{u}_1, \mathbf{u}_2, \cdots$ *in* $[0, 1]^d$, *discrepancy* $D_n(\mathbf{u}_1, ..., \mathbf{u}_n)$ *is defined by*

$$D_n(\mathbf{u}_1, ..., \mathbf{u}_n) = \sup_{B \in \mathcal{I}} \left| \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_{\{\mathbf{u}_k \in B\}} - \lambda(B) \right|$$

where $\mathbf{1}_{\{\mathbf{u}_k \in B\}} = 1$ *if* $\mathbf{u}_k \in B$ *and* 0 *otherwise,* $\mathcal{I}$ *is the family of sets with the form* $\prod_{i=1}^{d} [\alpha_i, \beta_i]$ *with* $0 \le \alpha_i \le \beta_i \le 1$, *and* λ *is the Lebesgue measure.*

Low-discrepancy sequences are sequences with discrepancy in the order of $\frac{(\ln n)^d}{n}$. In other words, low-discrepancy sequences fill in a unit hyper-cube in a highly uniform manner.

To fill in our target set $\mathcal{R}_n$, some contracting/stretching in scale and shifting in location are needed. Let $\mathbf{U}^{(n)}$ denote the $n \times d$ matrix of the first n points of a low-discrepancy sequence. Define

$$\mathbf{M}^{(n)} \equiv \begin{pmatrix} b_1^{(n)} - a_1^{(n)} & 0 & 0 & \cdots & 0 \\ 0 & b_2^{(n)} - a_2^{(n)} & 0 & \cdots & 0 \\ 0 & 0 & b_3^{(n)} - a_3^{(n)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & b_d^{(n)} - a_d^{(n)} \end{pmatrix} \tag{6}$$

and

$$\mathbf{A}_{n \times d}^{(n)} \equiv \begin{pmatrix} a_1^{(n)} & a_2^{(n)} & ... & a_d^{(n)} \\ \vdots & \vdots & & \vdots \\ a_1^{(n)} & a_2^{(n)} & ... & a_d^{(n)} \end{pmatrix}. \tag{7}$$

The suitable vector containing n points to be used to fill in our target set $\mathcal{R}_n$ can be created by the following linear transformation:

$$\mathbf{X}_{n \times d}^{(n)} \equiv \mathbf{U}_{n \times d}^{(n)} \mathbf{M}_{d \times d}^{(n)} + \mathbf{A}_{n \times d}^{(n)}. \tag{8}$$

Using all points contained in $\mathbf{X}^{(n)}$ turns out to be an inefficient way of constructing the low-discrepancy Markov chain. The reason is that the commonly used multi-dimensional normal density function concentrates its density on an elliptical region, i.e., the densities for points in the corners of the hyper-rectangle $\mathcal{R}_n$ are negligible. It is therefore natural to

discard the points lying outside of the elliptical region. We now discuss how an appropriate elliptical region is identified.

For this purpose, we first go back to the d-dimensional unit hyper-cube $\mathcal{C}$. Note that $\mathcal{C}$ is centered at the point $\frac{1}{2} = \left(\frac{1}{2}, ..., \frac{1}{2}\right)'$. We consider the ellipse $\mathcal{E}$ enclosed in $\mathcal{C}$:

$$\mathcal{E} \equiv \left\{ \mathbf{y} \in \mathcal{C} : \left(\mathbf{y} - \frac{1}{2}\right)' \Sigma^{-1} \left(\mathbf{y} - \frac{1}{2}\right) \leq \frac{1}{4} \right\} \tag{9}$$

where $\boldsymbol{\Sigma}$ is the correlation matrix corresponding to one period conditional distribution of the stochastic process $\mathbf{X}$. Hidden here is the assumption that the one-period conditional correlation matrix is state independent. If it is not the case, the particular feature of a given model may suggest another restriction set. It is worth noting that the ellipsoid shape of the set is not essential in the convergence proofs. All one needs is the partition condition. It is therefore possible to choose a different shape of set restriction suitable for a given model.

We now combine this elliptical restriction with scaling and location shifting to obtain the final set of points defining the state space for our n-point low-discrepancy Markov chain:

$$\mathbf{X}^{*(n)} \equiv \mathbf{M}^{(n)} \mathbf{U}^{*(n)} + \mathbf{A}^{(n)} \tag{10}$$

where $\mathbf{U}^{*(n)}$ contains the first n points that are from a low-discrepancy sequence and contained in $\mathcal{E}$. We choose $\mathbf{M}^{(n)}$ and $\mathbf{A}^{(n)}$ such that $\mathbf{M}^{(n)} \left(\frac{1}{2}, ..., \frac{1}{2}\right)' + \left(a_1^{(n)}, ..., a_d^{(n)}\right)'$, the state corresponding to the center of the ellipse, is identical to the initial state of the target Markov system.

The following convergence result justifies the use of $\boldsymbol{\Pi}^{(n)} \overrightarrow{v}^{(n)}(t+1)$ for $\mathrm{E}_t^Q [v(\mathbf{S}(t+1), t+1)]$. Moreover, it also justifies the use of the low-discrepancy Markov chain to compute the prices for derivative contracts.

Definition 2 *Let $\mathbf{x}_0 \in \mathbb{R}^d$ be arbitrarily chosen. We define $\mathbf{x}_0^{(n)}$ as a point in the state space $\left\{\mathbf{x}_1^{(n)}, ..., \mathbf{x}_n^{(n)}\right\}$ which minimizes the distance to $\mathbf{x}_0$, that is,*

$$\mathbf{x}_0^{(n)} \equiv \left\{ \mathbf{x}_i^{(n)} \in \left\{\mathbf{x}_1^{(n)}, ..., \mathbf{x}_n^{(n)}\right\} : \left\|\mathbf{x}_i^{(n)} - \mathbf{x}_0\right\| = \min_{j \in \{1,2,...,n\}} \left\|\mathbf{x}_j^{(n)} - \mathbf{x}_0\right\| \right\}.$$

This point may not be unique. In that case, we choose one of the points that minimizes the distance to $\mathbf{x}_0$.

Theorem 1 *For any $t \in \{1, 2, ..., T\}$, let $f_t(\mathbf{x}|\mathbf{z})$ denotes the t-period conditional density function of $\mathbf{X}(t)$ given $\mathbf{X}(0) = \mathbf{x}$. Assume that $f_t(\mathbf{x}|\mathbf{z})$ is bounded and continuous in*

$(\mathbf{x}, \mathbf{z}) \in \mathbb{R}^d \times \mathbb{R}^d$ *and has bounded variation* $V(f_t)$ *on* $\mathbb{R}^d \times \mathbb{R}^d$ *in the sense of Hardy and Krause. For any bounded function* $\phi : \mathbb{R}^d \to \mathbb{R}$ *such that*

$$\Phi(\mathbf{x}, \mathbf{z}) \equiv \phi(\mathbf{x}) f(\mathbf{x} | \mathbf{z}) \tag{11}$$

$$\Psi_t(\mathbf{x}, \mathbf{z}) \equiv \left(\int_{\mathbb{R}^d} \phi(\mathbf{y}) f_t(\mathbf{y} | \mathbf{x}) \, d\mathbf{y} \right) f(\mathbf{x} | \mathbf{z}), \ t \in \{1, 2, ..., T\} \tag{12}$$

have bounded variations $V(\Phi_t)$ *and* $V(\Psi_t(\mathbf{x}, \mathbf{z}))$, *we have*

$$\lim_{n \to \infty} \left| \mathrm{E}\left[\phi(\mathbf{X}_{t+1}) | \mathbf{X}_t = \mathbf{x}_0 \right] - \mathrm{E}\left[\phi\left(\mathbf{X}_{t+1}^{(n)} \right) \Big| \mathbf{X}_t^{(n)} = \mathbf{x}_0^{(n)} \right] \right| = 0 \tag{13}$$

and, for $t \in \{1, 2, ..., T\}$,

$$\lim_{n \to \infty} \left| \mathrm{E}\left[\phi(\mathbf{X}_t) | \mathbf{X}_0 = \mathbf{x}_0 \right] - \mathrm{E}\left[\phi\left(\mathbf{X}_t^{(n)} \right) \Big| \mathbf{X}_0^{(n)} = \mathbf{x}_0^{(n)} \right] \right| = 0. \tag{14a}$$

Proof. See Appendix Appendix A.

Remark 1 *(1)* $f_1(\mathbf{x} | \mathbf{z}) = f(\mathbf{x} | \mathbf{z})$. *(2) Assumption (11) is needed in the proof of the one-period case (Equation (13)). Assumption (12) is used in the induction argument that conduct to the multi-period result. They are sufficient but not necessary conditions. (3) If the function* ϕ *is not bounded but satisfies the integrability condition*

$$\mathrm{E}\left[|\phi(\mathbf{X}_t)| \, | \mathbf{X}_0 = \mathbf{x}_0 \right] < \infty,$$

then for any positive ε, *there is a bounded function* $\phi_{\varepsilon, \mathbf{x}_0}$ *such that*

$$\left| \mathrm{E}\left[|\phi(\mathbf{X}_t)| \, | \mathbf{X}_0 = \mathbf{x}_0 \right] - \mathrm{E}\left[|\phi_{\varepsilon, \mathbf{x}_0}(\mathbf{X}_t)| \, | \mathbf{X}_0 = \mathbf{x}_0 \right] \right| < \varepsilon.$$

Therefore, the price obtained using the Markov chain approximation may be as close as desired to the target price.

Density functions are typically "well behaved" functions that decay quickly to zero. The large majority of the payoff functions usually do not have periodic spikes. As a result, the bounded variation assumption for functions Φ and Ψ is not a particularly strong requirement for applications. Actually, this assumption is only needed for the application of a type of the Koksma-Hlawka error bound in "low-discrepancy integration". It can be replaced by some assumption on the modulus of continuity for these functions (see Niederreiter, 1992).

As a consequence of Theorem 1, the low-discrepancy Markov chain can be shown to converge weakly to the discrete-time version of the target process $\mathbf{X}$.

Corollary 1 *The low-discrepancy Markov chain at time t, $\mathbf{X}^{(n)}(t)$, with the initial state $\mathbf{x}_0^{(n)}$ and all possible states derived from a low-discrepancy sequence, subject to proper scaling, location shifting and elliptical restriction, converges weakly to $\mathbf{X}(t)$ for $t \in \{0, 1, 2, ..., T\}$ as n goes to infinity.*

Proof. We obtain the $k-$periods cumulative conditional distribution functions of $\mathbf{X}(t)$ and $\mathbf{X}^{(n)}(t)$ as follows: for any $\mathbf{x} \in \mathbb{R}^d$, set $\phi_{\mathbf{x}}(\mathbf{y}) = \mathbf{1}_{\{\mathbf{y} \leq \mathbf{x}\}}$. Then

$$F(\mathbf{x}\,|\mathbf{x}_0) \;\equiv\; \mathrm{E}\left[\phi_{\mathbf{x}}(\mathbf{X}_{t+k})\,|\mathbf{X}_t = \mathbf{x}_0\right] \;\text{ and} \tag{15}$$

$$F^{(n)}\left(\mathbf{x}\,\Big|\mathbf{x}_0^{(n)}\right) \;\equiv\; \mathrm{E}\left[\phi_{\mathbf{x}}\left(\mathbf{X}_{t+k}^{(n)}\right)\,\Big|\mathbf{X}_t^{(n)} = \mathbf{x}_0^{(n)}\right]. \tag{16}$$

In this case, $\Phi(\mathbf{y}, \mathbf{z}) = \phi_{\mathbf{x}}(\mathbf{y})\, f(\mathbf{y}\,|\mathbf{z})$. Because of the particular form of $\phi_{\mathbf{x}}$, we have $V(\Phi) \leq V(f) < \infty$. Moreover, $\Psi_t(\mathbf{y}, \mathbf{z}) \equiv \left(\int_{\mathbb{R}^d} \phi_{\mathbf{x}}(\mathbf{v})\, f_t(\mathbf{v}\,|\mathbf{y})\; d\mathbf{v}\right) f(\mathbf{y}\,|\mathbf{z}) = F_t(\mathbf{x}\,|\mathbf{y})\, f(\mathbf{y}\,|\mathbf{z})$ where $F_t(\mathbf{x}\,|\mathbf{y}) = \Pr[\mathbf{X}(t) \leq \mathbf{x}\,|\mathbf{X}(0) = \mathbf{y}]$ and Ψ has bounded variation if F_t does. The weak convergence result is therefore a consequence of Equation (14a). $\square$

4   Numerical Efficiency

The numerical performance of the Markov chain method may be improved by employing control variables. In many cases, one indeed have the knowledge about the first moment of some functions of the components of $\mathbf{S}$. The choice of the control variables depends on the contingent claim we want to price. Examples will be given later in this paper. In general, we consider k random control variables, $C_1(\mathbf{X}(t), t), ..., C_k(\mathbf{X}(t), t)$, for which 1) their respective one-period conditional expectations $\mu_1(t-1), ..., \mu_k(t-1)$ are known,

$$\mu_i(t-1) \equiv \mathrm{E}_{t-1}^Q[C_i(\mathbf{X}(t), t)], \tag{17}$$

and 2) they are correlated with the payoff of the target contingent claim. A more efficient way of implementing the low-discrepancy Markov chain method is to use the following control-variate dynamic system:

$$\begin{aligned} &\overrightarrow{v}_{cv}^{(n)}(t) = \overrightarrow{v}^{(n)}(t) - \sum_{i=1}^k \beta_{i,t}\left(\mathbf{\Pi}^{(n)}\overrightarrow{C_i}(t+1) - \overrightarrow{\mu_i}(t)\right), \; t \in \{0, 1, ..., T-1\} \\ &\overrightarrow{v}^{(n)}(t) = \max\left[h\left(g^{-1}(\mathbf{X}^{(n)}; t)\right), e^{-r(t)}\mathbf{\Pi}^{(n)}\,\overrightarrow{v}^{(n)}(t+1)\right] \\ &\overrightarrow{v}^{(n)}(T) = h\left(g^{-1}(\mathbf{X}^{(n)}(T); T)\right). \end{aligned} \tag{18}$$

where $\overrightarrow{C_i}(t+1)$ is an $n \times 1$ vector containing the value of the ith control variable for each state of the Markov chain, $\overrightarrow{\mu_i}(t)$ is an $n \times 1$ vector filled with $\mathrm{E}_t^Q\big[C_i(\mathbf{X}(t+1), t+1)\,\big|\mathbf{X}(t) = \mathbf{x}_j^{(n)}\big]$ for $j \in \{1, ..., n\}$, and $\beta_{i,t}$ is the regression coefficient from linearly regressing $\overrightarrow{v}^{(n)}(t)$ on $\mathbf{\Pi}^{(n)}\overrightarrow{C_i}(t+1)$.

5 Examples

5.1 The multivariate Black-Scholes model

In this section, we demonstrate how an n-point Markov chain can be constructed and used for pricing derivatives on several assets governed by a multivariate geometric Brownian motion. The geometric Brownian motion for stock i is written as

$$\ln S_i(t) = \ln S_i(0) + \left(r - \frac{1}{2}\sigma_i^2 \right) t + \sigma_i W_i(t) \quad \text{for } i \in \{1, ..., d\} \tag{19}$$

where r and σ_i are, respectively, the risk-free rate and diffusion coefficient under the risk neutral probability measure Q. The d-dimensional Brownian motion $\mathbf{W} = \{(W_1(t), ..., W_d(t)) : t \geq 0\}$ is such that

$$Corr\left[W_i(t), W_j(t) \right] = \rho_{ij} \text{ for any } t \geq 0. \tag{20}$$

Let

$$\mathbf{X}(t) = \begin{bmatrix} \ln S_1(t) - \ln S_1(0) - (r - \frac{\sigma_1^2}{2})t \\ \ln S_2(t) - \ln S_2(0) - (r - \frac{\sigma_2^2}{2})t \\ \vdots \\ \ln S_d(t) - \ln S_d(0) - (r - \frac{\sigma_d^2}{2})t \end{bmatrix}. \tag{21}$$

Since

$$
\begin{aligned}
X_i(t+1) &= \ln S_i(t+1) - \ln S_i(0) - \left(r - \frac{\sigma_1^2}{2} \right)(t+1) \\
&= \ln S_i(t) - \ln S_i(0) - \left(r - \frac{1}{2}\sigma_i^2 \right)t + \sigma_i\left(W_i(t+1) - W_i(t) \right) \\
&= X_i(t) + \sigma_i\left(W_i(t+1) - W_i(t) \right), \tag{22}
\end{aligned}
$$

then the stochastic process $\mathbf{X}$ is a time-homogeneous Markov process and its one period conditional joint density function is

$$f(\mathbf{x}_j \,|\mathbf{x}_i) = \left(\frac{1}{2\pi} \right)^{d/2} |\mathbf{\Lambda}|^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{x}_j - \mathbf{x}_i)' \, \mathbf{\Lambda}^{-1}(\mathbf{x}_j - \mathbf{x}_i) \right) \tag{23}$$

where $\mathbf{\Lambda}$ is the covariance matrix $(\sigma_i\sigma_j\rho_{ij})_{ij}$. Note that if σ_i's are annualized figures, they should be multiplied by $\sqrt{h}$ where h is the length of one period measured in years.

We construct an n-point Markov chain to approximate $\mathbf{X}(t)$ for $t \in \{0, 1, 2, ..., T\}$ in three steps. For illustration purposes and without loss of generality, we consider a

two-dimensional problem. Specifically, we consider a system with the two-dimensional correlation matrix:

$$\mathbf{\Sigma} = \left[\begin{array}{cc} 1.0 & 0.5 \\ 0.5 & 1.0 \end{array} \right]. \tag{24}$$

- **Step 1:** generate a sequence of low-discrepancy 2-dimensional Sobol numbers and denote them by $\{\mathbf{u}_1, \mathbf{u}_2, \cdots\}$. This task can be accomplished using a standard computer code.

- **Step 2:** remove low-likelihood points using a suitable ellipse to obtain n points. This ellipse is based upon the covariance matrix of the components of the stochastic process $\mathbf{X}$. The less correlated are the components of $\mathbf{X}$, the closer is the ellipse to the unit circle. Specifically, pick the first n points of $\{\mathbf{u}_1, \mathbf{u}_2, \cdots\}$ such that

$$\left(\mathbf{u}_i - \left(\begin{array}{c} 1/2 \\ 1/2 \end{array} \right) \right)' \mathbf{\Sigma}^{-1} \left(\mathbf{u}_i - \left(\begin{array}{c} 1/2 \\ 1/2 \end{array} \right) \right) \leq \frac{1}{4} \tag{25}$$

  and denote these points by $\{\mathbf{u}_1^*, \mathbf{u}_2^*, \cdots, \mathbf{u}_n^*\} \subset \mathcal{E}$. With our chosen $\mathbf{\Sigma}$, the first n points must satisfy

$$\left(u_1 - \frac{1}{2} \right)^2 - \left(u_1 - \frac{1}{2} \right) \left(u_2 - \frac{1}{2} \right) + \left(u_2 - \frac{1}{2} \right)^2 \leq \frac{3}{16}. \tag{26}$$

- **Step 3:** perform scaling and location shifting to create an elliptical region enclosed in $\mathcal{R}_n$ and centered at $\mathbf{X}_0$. The Markov chain's discrete state values can be generated as follows:

$$\mathbf{x}_i = \left[\begin{array}{c} \ln S_1(0) \\ \ln S_2(0) \end{array} \right] + b^{(n)} \left[\begin{array}{cc} \sigma_1 \sqrt{T} & 0 \\ 0 & \sigma_2 \sqrt{T} \end{array} \right] \mathbf{u}_i^*. \tag{27}$$

  and $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\} \subset \mathcal{E}_n$. Note that the two respective standard deviations are used to reflect the range of values for each of two variables T periods into the future. Since the construction of the Sobol sequence implies that $\mathbf{u}_1^* = \mathbf{0}$, we have

$$\mathbf{x}_1 = \left[\begin{array}{c} \ln S_1(0) \\ \ln S_2(0) \end{array} \right]. \tag{28}$$

In other words, the first element of the Markov chain state values is the initial state value of the target Markov process. We need to choose $b^{(n)}$ so that the partition conditions are met, and a good candidate is

$$b^{(n)} \simeq 2 \ln \left(\ln \left(n \right) \right). \tag{29}$$

For the Markov chain method to perform reasonably well, some care must be given to ensure that a reasonable coverage is obtained. One can, for example, set a target

probability of the elliptical region at $\alpha\%$. Our Monte Carlo calculations suggest that $b^{(n)}$ are roughly between 2 and 4 for α being 95 or 99. for various numbers of the underlying assets. Note that these results are independent of n due to the geometric Brownian motion assumption.

6 Numerical results

Numerical results are presented in Tables 1 to 4 for European put options on the minimum of several assets in the context of the Black Scholes model. In each table, we consider three maturities (1, 3 and 9 months) and three strike-to-asset price ratios (1.1, 1.0 and 0.9). The current asset prices are set at 50 for all assets. The annualized risk-free rate (continuously compounded) is 5% and the assets' standard deviations are 20% for all assets. The correlation is set to 0.5 between all pairs of assets. The first row of this table presents the benchmark prices obtained using Monte Carlo simulations except for the one-asset case where the Black-Scholes formula is applicable. In all tables, the first panel presents prices computed with a weekly time step, whereas the second panel presents the prices obtained with a monthly time step. We use the individual Black-Scholes prices as control variables. In other words, we assume several individual put options using the same strike price. Of course, for the one-asset case the control variate adjustment is not made.

There are several observations about the results in these tables:

- The method produces prices that suitably converge to the correct benchmark prices, consistent with the theoretical predication.

- The convergence to the benchmark prices is faster for shorter maturities.

- The convergence to the benchmark prices is faster when the time steps are larger relative to the maturity. Indeed, the prices obtained with a weekly time step require a Markov chain with a larger dimension to achieve the penny accuracy as compared to the cases using a monthly time step.

- As the number of assets increases, the dimension of the chain needs to be increased substantially to achieve reasonable accuracy. However, sparse matrix techniques allow for the use of high-dimensional Markov chains because in most cases, many entries of the transition probability matrix are effectively zero. For example, for three assets with a maturity of 12 weeks, 97% of the entries in the transition probability matrix are zero.

- As the number of assets increases, more points are removed due to the elliptical restriction. In these tables, n_i denotes the would-be dimension of the Markov chain if no elliptical restriction were applied. The effective dimension with the elliptical restriction is denoted by n_e.

Table 1: European put options on one asset

	Maturity = 4 weeks			Maturity = 12 weeks			Maturity = 36 weeks		
K	55.00	50.00	45.00	55.00	50.00	45.00	55.00	50.00	45.00
Benchmark prices									
	4.8470	1.0290	0.0271	4.9049	1.6583	0.2601	5.2221	2.5116	0.8945
Markov Chain prices									
$n_i = 16, \ n_e = 16$	4.9056	0.8565	0.0000	4.9915	1.3622	0.1307	5.1989	1.9490	0.5768
$n_i = 32, \ n_e = 32$	4.9403	1.0901	0.0121	5.0319	1.7627	0.2347	5.3957	2.6568	0.9479
$n_i = 64, \ n_e = 64$	4.9514	1.0677	0.0209	5.0982	1.7437	0.2493	5.4688	2.6264	0.8612
$n_i = 128, \ n_e = 128$	4.8877	1.0573	0.0250	4.9662	1.7072	0.2584	5.3127	2.5870	0.9236
$n_i = 256, \ n_e = 256$	4.8474	1.0107	0.0256	4.8975	1.6202	0.2511	5.1831	2.4430	0.8634
$n_i = 512, \ n_e = 512$	4.8498	1.0317	0.0268	4.9091	1.6624	0.2629	5.2281	2.5175	0.9000
$n_i = 1024, \ n_e = 1024$	4.8545	1.0340	0.0270	4.9167	1.6671	0.2601	5.2402	2.5259	0.8989
$n_i = 2048, \ n_e = 2048$	4.8485	1.0304	0.0271	4.9073	1.6605	0.2611	5.2255	2.5149	0.8972

Parameters: 1 week = 1/50, Δt =0.0200, $r = 0.05$ (annualized), s_1= 50.00, σ_1= 0.20 (annualized). n_i and n_e are, respectively, the initial and effective size of the Markov chain. Average sparcity for each maturity: 0.63, 0.43, 0.26. Stock price partition computed with sobol numbers and $2 \times \log(\log(n_e))$. Markov chain on adjusted stock price. Benchmarks prices computed with Black-Scholes formula.

	Maturity = 1 month			Maturity = 4 months			Maturity = 9 months		
K	55.00	50.00	45.00	55.00	50.00	45.00	55.00	50.00	45.00
Benchmark prices									
	4.8470	1.0290	0.0271	4.9658	1.8631	0.3851	5.2221	2.5116	0.8945
Markov Chain prices									
$n_i = 16, \ n_e = 16$	4.7533	0.7711	0.0000	4.9113	1.5335	0.2051	5.1142	2.0228	0.5861
$n_i = 32, \ n_e = 32$	4.9032	1.0330	0.0062	5.1283	1.9809	0.3753	5.4318	2.6794	0.9537
$n_i = 64, \ n_e = 64$	4.8813	1.0239	0.0177	5.1402	1.9320	0.3698	5.4884	2.6398	0.8824
$n_i = 128, \ n_e = 128$	4.8748	1.0418	0.0233	5.0404	1.9165	0.3867	5.3198	2.5892	0.9233
$n_i = 256, \ n_e = 256$	4.8458	1.0156	0.0254	4.9544	1.8297	0.3747	5.1875	2.4558	0.8686
$n_i = 512, \ n_e = 512$	4.8516	1.0320	0.0265	4.9711	1.8682	0.3882	5.2288	2.5181	0.9002
$n_i = 1024, \ n_e = 1024$	4.8517	1.0314	0.0268	4.9794	1.8724	0.3853	5.2404	2.5253	0.8986
$n_i = 2048, \ n_e = 2048$	4.8489	1.0303	0.0270	4.9687	1.8658	0.3863	5.2259	2.5151	0.8971

Parameters: 1 month = 1/12.5, Δt =0.0800, $r = 0.05$ (annualized), s_1= 50.00, σ_1= 0.20 (annualized). n_i and n_e are, respectively, the initial and effective size of the Markov chain. Average sparcity for each maturity: 0.75, 0.63, 0.48. Stock price partition computed with Sobol numbers and $2 \times \log(\log(n_e))$. Markov chain on adjusted stock price.Benchmarks prices computed with Black-Scholes formula.

Table 2: European put options on the minimum of 2 assets

	Maturity = 4 weeks			Maturity = 12 weeks			Maturity = 36 weeks		
K	55.00	50.00	45.00	55.00	50.00	45.00	55.00	50.00	45.00
				Benchmark prices					
	5.9164	1.5458	0.0501	6.4639	2.4979	0.4496	7.2916	3.8015	1.4735
				Markov Chain prices					
$n_i = 128,\ n_e = 90$	5.8044	1.5218	0.0508	5.7547	2.2926	0.4133	5.8375	3.2302	1.2584
$n_i = 256,\ n_e = 178$	5.8752	1.5068	0.0510	6.1943	2.3573	0.4356	6.1600	3.2805	1.3191
$n_i = 512,\ n_e = 352$	5.9130	1.5418	0.0503	6.3631	2.4500	0.4485	6.4669	3.3807	1.3462
$n_i = 1024,\ n_e = 701$	5.9228	1.5496	0.0502	6.4535	2.5031	0.4498	7.0478	3.7146	1.4188
$n_i = 2048,\ n_e = 1392$	5.9036	1.5431	0.0501	6.4281	2.4887	0.4476	7.1908	3.7613	1.4587
$n_i = 4096,\ n_e = 2785$	5.9153	1.5507	0.0503	6.4586	2.5072	0.4504	7.2660	3.8040	1.4710

Parameters: 1 week = 1/50, Δt =0.0200, $r = 0.05$ (annualized), s_1= 50.00, σ_1= 0.20 (annualized). s_2= 50.00, σ_2= 0.20 (annualized).Correlation for all pair of assets =0.50. n_i and n_e are, respectively, the initial and effective size of the Markov chain. Average sparcity for each maturity: 0.34, 0.14, 0.05. Stock price partition computed with Sobol numbers and $2 \times \log(\log(n_e))$. Benchmarks are Monte Carlo price computed with 500000 sample paths. Markov chain on adjusted stock price. Markov chain dimension reduced according to an elliptical pattern.

	Maturity = 1 month			Maturity = 3 months			Maturity = 9 months		
K	55.00	50.00	45.00	55.00	50.00	45.00	55.00	50.00	45.00
				Benchmark prices					
	5.9164	1.5458	0.0501	6.4639	2.4979	0.4496	7.2916	3.8015	1.4735
				Markov Chain prices					
$n_i = 128,\ n_e = 90$	5.8813	1.5347	0.0509	6.3193	2.4740	0.4448	6.6424	3.6011	1.3610
$n_i = 256,\ n_e = 178$	5.9106	1.5288	0.0512	6.4110	2.4472	0.4503	7.0155	3.6372	1.4174
$n_i = 512,\ n_e = 352$	5.9166	1.5444	0.0501	6.4585	2.4947	0.4534	7.1969	3.7614	1.4726
$n_i = 1024,\ n_e = 701$	5.9209	1.5484	0.0501	6.4701	2.5038	0.4497	7.2862	3.8145	1.4746
$n_i = 2048,\ n_e = 1392$	5.9144	1.5468	0.0502	6.4487	2.4950	0.4482	7.2497	3.7897	1.4707
$n_i = 4096,\ n_e = 2785$	5.9182	1.5497	0.0503	6.4644	2.5053	0.4503	7.2929	3.8156	1.4775

Parameters: 1 month = 1/12.5, Δt =0.0800, $r = 0.05$ (annualized), s_1= 50.00, σ_1= 0.20 (annualized). s_2= 50.00, σ_2= 0.20 (annualized).Correlation for all pair of assets =0.50. n_i and n_e are, respectively, the initial and effective size of the Markov chain. Average sparsity for each maturity: 0.71, 0.42, 0.18. Stock price partition computed with Sobol numbers and $2 \times \log(\log(n_e))$. Benchmarks are Monte Carlo price computed with 500000 sample paths. Markov chain on adjusted stock price. Markov chain dimension reduced according to an elliptical pattern.

Table 3: European put options on the minimum of 3 assets

		Maturity = 4 weeks			Maturity = 12 weeks			Maturity = 36 weeks		
K		55.00	50.00	45.00	55.00	50.00	45.00	55.00	50.00	45.00
					Benchmark prices					
		6.4673	1.8879	0.0707	7.3271	3.0552	0.6027	8.5221	4.6628	1.9091
					Markov Chain prices					
$n_i = 128$, $n_e = 48$		5.7944	1.7223	0.0722	5.0069	2.3393	0.5021	5.5226	3.4961	1.4865
$n_i = 256$, $n_e = 96$		5.7561	1.7158	0.0733	5.2845	2.5755	0.4951	5.5569	3.5065	1.4592
$n_i = 512$, $n_e = 187$		5.8946	1.6425	0.0689	5.3358	2.2166	0.4849	5.5428	3.4276	1.4118
$n_i = 1024$, $n_e = 378$		6.1934	1.8016	0.0679	5.8235	2.5022	0.4760	6.3478	3.9241	1.3847
$n_i = 2048$, $n_e = 760$		6.3368	1.8490	0.0698	6.3229	2.7054	0.4992	6.5395	4.1585	1.3928
$n_i = 4096$, $n_e = 1515$		6.3834	1.8669	0.0693	6.7133	2.8706	0.5193	6.4534	3.8450	1.3907
$n_i = 8192$, $n_e = 3049$		6.4266	1.8780	0.0702	6.9555	2.9261	0.5474	7.5997	4.1282	1.4477
$n_i = 16384$, $n_e = 6080$		6.4631	1.8860	0.0707	7.2416	3.0198	0.5932	7.6590	4.1925	1.6025
$n_i = 32768$, $n_e = 12152$		6.4685	1.8883	0.0708	7.3067	3.0434	0.5971	8.1753	4.4452	1.7681
$n_i = 65536$, $n_e = 24294$		6.4688	1.8882	0.0707	7.3241	3.0536	0.6013	8.3566	4.5512	1.8368
$n_i = 131072$, $n_e = 48595$		–	–	–	7.3231	3.0532	0.6024	8.4290	4.6016	1.8729
$n_i = 262144$, $n_e = 97187$		–	–	–	7.3227	3.0525	0.6019	8.4682	4.6248	1.8843
$n_i = 524288$, $n_e = 194226$		–	–	–	–	–	–	8.5161	4.6576	1.9057

Parameters: 1 week = 1/50, Δt =0.0200, $r = 0.05$ (annualized), s_1= 50.00, $\sigma_1 = 0.20$ (annualized). s_2= 50.00, $\sigma_2 = 0.20$ (annualized).s_3= 50.00, $\sigma_3 = 0.20$ (annualized). Correlation for all pair of assets =0.50. n_i and n_e are, respectively, the initial and effective size of the Markov chain. Average sparsity for each maturity: 0.14, 0.03, 0.01. Stock price partition computed with Sobol numbers and $2 \times \log(\log(n_e))$. Benchmarks are Monte Carlo price computed with 500000 sample paths. Markov chain on adjusted stock price. Markov chain dimension reduced according to an elliptical patern. Black-Scholes control variates with a put option on each individual asset.

		Maturity = 1 month			Maturity = 3 months			Maturity = 9 months		
K		55.00	50.00	45.00	55.00	50.00	45.00	55.00	50.00	45.00
					Benchmark prices					
		6.4673	1.8879	0.0707	7.3271	3.0552	0.6027	8.5221	4.6628	1.9091
					Markov Chain prices					
$n_i = 128$, $n_e = 48$		6.3210	1.8429	0.0740	6.6532	2.8859	0.5732	5.8656	3.6653	1.5611
$n_i = 256$, $n_e = 96$		6.4404	1.8846	0.0731	6.7083	2.8783	0.5746	6.1549	3.9264	1.5222
$n_i = 512$, $n_e = 187$		6.3512	1.8251	0.0703	6.7282	2.7795	0.5819	6.0608	3.4531	1.5338
$n_i = 1024$, $n_e = 378$		6.4046	1.8802	0.0702	7.0656	2.9784	0.5885	7.0422	3.9247	1.5845
$n_i = 2048$, $n_e = 760$		6.4789	1.9109	0.0712	7.2057	3.0380	0.6055	7.5965	4.2442	1.6870
$n_i = 4096$, $n_e = 1515$		6.4508	1.8882	0.0709	7.2504	3.0366	0.5943	8.0486	4.4550	1.7531
$n_i = 8192$, $n_e = 3049$		6.4707	1.8943	0.0711	7.2924	3.0519	0.6022	8.2396	4.5272	1.8299
$n_i = 16384$, $n_e = 6080$		6.4658	1.8879	0.0711	7.3228	3.0533	0.6030	8.4773	4.6415	1.9024
$n_i = 32768$, $n_e = 12152$		6.4667	1.8875	0.0711	7.3288	3.0559	0.6026	8.5136	4.6567	1.9039
$n_i = 65536$, $n_e = 24294$		–	–	–	7.3292	3.0553	0.6028	8.5250	4.6633	1.9088
$n_i = 131072$, $n_e = 48595$		–	–	–	–	–	–	8.5205	4.6612	1.9084

Parameters: 1 month = 1/12.5, Δt =0.0800, $r = 0.05$ (annualized), s_1= 50.00, $\sigma_1 = 0.20$ (annualized). s_2= 50.00, $\sigma_2 = 0.20$ (annualized). s_3= 50.00, $\sigma_3 = 0.20$ (annualized). Correlation for all pair of assets =0.50. n_i and n_e are, respectively, the initial and effective size of the Markov chain. Average sparsity for each maturity: 0.46, 0.14, 0.04. Stock price partition computed with Sobol numbers and $2 \times \log(\log(n_e))$. Benchmarks are Monte Carlo price computed with 500000 sample paths. Markov chain on adjusted stock price. Markov chain dimension reduced according to an elliptical pattern. Black-Scholes control variates with a put option on each individual asset.

Table 4: European put options on the minimum of 4 assets

		Maturity = 4 weeks			Maturity = 12 weeks			Maturity = 36 weeks		
K		55.00	50.00	45.00	55.00	50.00	45.00	55.00	50.00	45.00
Benchmark prices										
		6.8196	2.1350	0.0894	7.8936	3.4580	0.7300	9.3557	5.2867	2.2529
Markov Chain prices										
$n_i = 8192,\ n_e = 1416$		6.3112	2.0160	0.0751	6.4034	2.9785	0.5106	6.5789	4.0457	1.4738
$n_i = 16384,\ n_e = 2802$		6.5260	2.0319	0.0812	6.4759	2.9802	0.5417	5.5873	3.5669	1.4806
$n_i = 32768,\ n_e = 5640$		6.6802	2.0863	0.0849	6.9281	3.0592	0.5589	6.4765	3.8854	1.4760
$n_i = 65536,\ n_e = 11275$		6.7576	2.1124	0.0867	7.2665	3.1769	0.5905	6.8278	3.8775	1.4687
$n_i = 131072,\ n_e = 22549$		6.7697	2.1155	0.0869	7.4834	3.2605	0.6256	7.3054	4.1398	1.4936
$n_i = 262144,\ n_e = 45132$		6.7772	2.1211	0.0873	7.6300	3.3455	0.6577	8.0059	4.4465	1.5986
$n_i = 524288,\ n_e = 90285$		6.8165	2.1378	0.0889	7.8214	3.4348	0.7195	8.2798	4.6479	1.8066
$n_i = 1048576,\ n_e = 180766$		–	–	–	7.8743	3.4495	0.7233	8.7852	4.8885	1.9550
$n_i = 2097152,\ n_e = 361506$		–	–	–	–	–	–	9.0762	5.0819	2.1036
$n_i = 4194304,\ n_e = 722992$		–	–	–	–	–	–	9.2151	5.1787	2.1729

Parameters: 1 week = 1/50, Δt =0.0200, r = 0.05 (annualized), s_1= 50.00, σ_1= 0.20 (annualized). s_2= 50.00, σ_2= 0.20 (annualized).s_3= 50.00, σ_3= 0.20 (annualized). s_4= 50.00, σ_4= 0.20 (annualized). Correlation for all pair of assets =0.50. n_i and n_e are, respectively, the initial and effective size of the Markov chain. Average sparsity for each maturity: 0.08, 0.01, 0.00. Stock price partition computed with Sobol numbers and $2 \times \log(\log(n_e))$. Benchmarks are Monte Carlo price computed with 500000 sample paths. Markov chain on adjusted stock price. Markov chain dimension reduced according to an elliptical pattern. Black-Scholes control variates with a put option on each individual asset.

		Maturity = 1 months			Maturity = 3 months			Maturity = 9 months		
K		55.00	50.00	45.00	55.00	50.00	45.00	55.00	50.00	45.00
Benchmark prices										
		6.8196	2.1350	0.0894	7.8936	3.4580	0.7300	9.3557	5.2867	2.2529
Markov Chain prices										
$n_i = 8192,\ n_e = 1416$		6.7091	2.1183	0.0895	7.3965	3.3157	0.6657	7.9952	4.6454	1.7441
$n_i = 16384,\ n_e = 2802$		6.7646	2.1107	0.0889	7.6318	3.3404	0.6877	8.1021	4.6772	1.8726
$n_i = 32768,\ n_e = 5640$		6.7984	2.1288	0.0896	7.7843	3.4097	0.7095	8.6226	4.8463	1.9685
$n_i = 65536,\ n_e = 11275$		6.8238	2.1398	0.0902	7.8627	3.4443	0.7232	8.8881	4.9897	2.0564
$n_i = 131072,\ n_e = 22549$		6.8233	2.1393	0.0903	7.8685	3.4444	0.7238	9.0640	5.0755	2.1071
$n_i = 262144,\ n_e = 45132$		–	–	–	7.8746	3.4507	0.7252	9.1784	5.1597	2.1555
$n_i = 524288,\ n_e = 90285$		–	–	–	–	–	–	9.3382	5.2833	2.2500

Parameters: 1 month = 1/12.5, Δt =0.0800, r = 0.05 (annualized), s_1= 50.00, σ_1= 0.20 (annualized). s_2= 50.00, σ_2= 0.20 (annualized).s_3= 50.00, σ_3= 0.20 (annualized). s_4= 50.00, σ_4= 0.20 (annualized). Correlation for all pair of assets =0.50. n_i and n_e are, respectively, the initial and effective size of the Markov chain. Average sparsity for each maturity: 0.40, 0.13, 0.02. Stock price partition computed with Sobol numbers and $2 \times \log(\log(n_e))$. Benchmarks are Monte Carlo price computed with 500000 sample paths. Markov chain on adjusted stock price. Markov chain dimension reduced according to an elliptical pattern. Black-Scholes control variates with a put option on each individual asset.

Appendix A   Proof of Theorem 1

Let $\mathbf{u}_1^*, ..., \mathbf{u}_n^*$ denotes the first n points of a low-discrepancy sequence on the $d-$dimensional ellipse $\mathcal{E}$ enclosed in the unit hypercube $\mathcal{C}$. To obtain a sample on the $d-$dimensional ellipse $\mathcal{E}_n$ contained in the hyper-rectangle $\mathcal{R}_n$, we rescale these points :

$$\mathbf{x}_i^{(n)} \equiv \mathbf{M}^{(n)}\mathbf{u}_i^* + \mathbf{A}^{(n)}.$$

Note that the ith state $\mathbf{x}_i^{(n)}$ of the $n-$states low-discrepancy Markov chain will go away from the origin as n increases. Let $\mathbf{x}_0 \in \mathbb{R}^d$ be the initial state of the Markov process $\mathbf{X}$. For any $n \in \mathbb{N}$, we choose the state $\mathbf{x}_0^{(n)}$ of the Markov chain approximation which is the closest to $\mathbf{x}_0$, that is,

$$\mathbf{x}_0^{(n)} \equiv \left\{ \mathbf{x}_i^{(n)} \in \left\{ \mathbf{x}_1^{(n)}, ..., \mathbf{x}_n^{(n)} \right\} : \left\| \mathbf{x}_i^{(n)} - \mathbf{x}_0 \right\| = \min_{j \in \{1,2,...,n\}} \left\| \mathbf{x}_j^{(n)} - \mathbf{x}_0 \right\| \right\}. \tag{30}$$

This point may not be unique. In that case, we choose one of the points that minimizes the distance to $\mathbf{x}_0$. Because we are working with a low-discrepancy sequence and since the volume $\lambda(\mathcal{R}_n)$ increases slower that the number n of points, then

$$\left\| \mathbf{x}_0^{(n)} - \mathbf{x}_0 \right\| \to 0 \text{ as } n \to \infty. \tag{31}$$

Appendix A.1   The one-period case

We will first consider the one-period case. Recall that $\phi : \mathbb{R}^d \to \mathbb{R}$ is a bounded function, that is

$$\|\phi\|_\infty \equiv \sup_{\mathbf{x} \in \mathbb{R}^d} |\phi(\mathbf{x})| < \infty.$$

Consider the one-period conditional expectation of $\phi(\mathbf{X}_{t+1})$ for the original process and Markov chain respectively:

$$\mathrm{E}\left[\phi(\mathbf{X}_{t+1}) | \mathbf{X}_t = \mathbf{x}\right] = \int_{\mathbb{R}^d} \phi(\mathbf{y}) f(\mathbf{y}|\mathbf{x}) \, d\mathbf{y};$$

$$\mathrm{E}\left[\phi\left(\mathbf{X}_{t+1}^{(n)}\right) \middle| \mathbf{X}_t^{(n)} = \mathbf{x}_i^{(n)}\right] = \sum_{j=1}^{n} \phi\left(\mathbf{x}_j^{(n)}\right) \pi_{ij}^{(n)}$$

$$= \frac{\frac{\lambda(\mathcal{R}_n)}{n} \sum_{j=1}^{n} \phi\left(\mathbf{x}_j^{(n)}\right) f\left(\mathbf{x}_j^{(n)} \middle| \mathbf{x}_i^{(n)}\right)}{\frac{\lambda(\mathcal{R}_n)}{n} \sum_{j=1}^{n} f\left(\mathbf{x}_j^{(n)} \middle| \mathbf{x}_i^{(n)}\right)}. \tag{32}$$

We want to prove that

$$\lim_{n \to \infty} \left| \mathrm{E}\left[\phi(\mathbf{X}_{t+1}) | \mathbf{X}_t = \mathbf{x}_0\right] - \mathrm{E}\left[\phi\left(\mathbf{X}_{t+1}^{(n)}\right) \middle| \mathbf{X}_t^{(n)} = \mathbf{x}_0^{(n)}\right] \right| = 0. \tag{33}$$

The goal of the next few steps is to reduce this problem to a simpler one. As proved in Appendix Appendix B.1, Equation (33) is a consequence of

$$\lim_{n\to\infty} \left| \mathrm{E}\left[\phi\left(\mathbf{X}_{t+1}\right)|\mathbf{X}_t = \mathbf{x}_0\right] - \frac{\lambda\left(\mathcal{R}_n\right)}{n} \sum_{j=1}^{n} \phi\left(\mathbf{x}_j^{(n)}\right) f\left(\mathbf{x}_j^{(n)} \middle| \mathbf{x}_0^{(n)}\right) \right| = 0. \qquad (34)$$

Because $\mathcal{E}_n \uparrow \mathbb{R}^d$ as $n \uparrow \infty$ and since $\int_{\mathbb{R}^d} |\phi\left(\mathbf{y}\right)| f\left(\mathbf{y}|\mathbf{x}_0\right) d\mathbf{y} < \infty$, then

$$\lim_{n\to\infty} \left| \mathrm{E}\left[\phi\left(\mathbf{X}_{t+1}\right)|\mathbf{X}_t = \mathbf{x}_0\right] - \int_{\mathcal{E}_n} \phi\left(\mathbf{y}\right) f\left(\mathbf{y}|\mathbf{x}_0\right) \, d\mathbf{y} \right| = 0. \qquad (35)$$

Therefore, using the triangle inequality, Equation (34) reduces to

$$\lim_{n\to\infty} \left| \int_{\mathcal{E}_n} \phi\left(\mathbf{y}\right) f\left(\mathbf{y}|\mathbf{x}_0\right) \, d\mathbf{y} - \frac{\lambda\left(\mathcal{R}_n\right)}{n} \sum_{j=1}^{n} \phi\left(\mathbf{x}_j^{(n)}\right) f\left(\mathbf{x}_j^{(n)} \middle| \mathbf{x}_0^{(n)}\right) \right| = 0. \qquad (36)$$

In the last expression, the two terms has not the same initial point. However, as proved in Appendix Appendix B.2, the continuity of the one-period conditional density function $f\left(\mathbf{x}|\mathbf{x}_0\right)$ in $\mathbb{R}^d \times \mathbb{R}^d$ is used to claim that

$$\lim_{n\to\infty} \left| \int_{\mathcal{E}_n} \phi\left(\mathbf{y}\right) f\left(\mathbf{y}|\mathbf{x}_0\right) \, d\mathbf{y} - \int_{\mathcal{E}_n} \phi\left(\mathbf{y}\right) f\left(\mathbf{y}\middle|\mathbf{x}_0^{(n)}\right) \, d\mathbf{y} \right| = 0. \qquad (37)$$

Consequently, Equation (36) is valid if

$$\lim_{n\to\infty} \left| \int_{\mathcal{E}_n} \phi\left(\mathbf{y}\right) f\left(\mathbf{y}\middle|\mathbf{x}_0^{(n)}\right) \, d\mathbf{y} - \frac{\lambda\left(\mathcal{R}_n\right)}{n} \sum_{j=1}^{n} \phi\left(\mathbf{x}_j^{(n)}\right) f\left(\mathbf{x}_j^{(n)} \middle| \mathbf{x}_0^{(n)}\right) \right| = 0. \qquad (38)$$

We now rescale the integral $\int_{\mathcal{E}_n} \phi\left(\mathbf{y}\right) f\left(\mathbf{y}\middle|\mathbf{x}_0^{(n)}\right) \, d\mathbf{y}$ on the $d-$dimensional ellipse $\mathcal{E}$:

$$\int_{\mathcal{E}_n} \phi\left(\mathbf{y}\right) f\left(\mathbf{y}\middle|\mathbf{x}_0^{(n)}\right) \, d\mathbf{y} = \lambda\left(\mathcal{R}_n\right) \int_{\mathcal{E}} \phi\left(\mathbf{M}^{(n)}\mathbf{u} + \mathbf{A}^{(n)}\right) f\left(\mathbf{M}^{(n)}\mathbf{u} + \mathbf{a}^{(n)} \middle| \mathbf{x}_0^{(n)}\right) \, d\mathbf{u}.$$

where $\mathbf{a}^{(n)} = \left(a_1^{(n)}, ..., a_d^{(n)}\right)'$. Equation (38) is therefore equivalent to

$$\lim_{n\to\infty} \left| \begin{array}{l} \lambda\left(\mathcal{R}_n\right) \int_{\mathcal{E}} \phi\left(\mathbf{M}^{(n)}\mathbf{u} + \mathbf{a}^{(n)}\right) f\left(\mathbf{M}^{(n)}\mathbf{u} + \mathbf{a}^{(n)} \middle| \mathbf{x}_0^{(n)}\right) \, d\mathbf{u} \\ -\frac{\lambda\left(\mathcal{R}_n\right)}{n} \sum_{j=1}^{n} \phi\left(\mathbf{M}^{(n)}\mathbf{u}_j^* + \mathbf{a}^{(n)}\right) f\left(\mathbf{M}^{(n)}\mathbf{u}_j^* + \mathbf{a}^{(n)} \middle| \mathbf{x}_0^{(n)}\right) \end{array} \right| = 0. \qquad (39)$$

Finally, we can prove that Equation (39) is satisfied using some classical error bound theorem (often called the Koksma-Hlawka inequality). Indeed,

$$
\begin{aligned}
&\lambda\left(\mathcal{R}_n\right)\left|\begin{array}{c} \int_{\mathcal{E}} \phi\left(\mathbf{M}^{(n)}\mathbf{u}+\mathbf{a}^{(n)}\right) f\left(\mathbf{M}^{(n)}\mathbf{u}+\mathbf{a}^{(n)}\left|\mathbf{x}_0^{(n)}\right.\right)\, d\mathbf{u} \\ -\frac{1}{n}\sum_{j=1}^n \phi\left(\mathbf{M}^{(n)}\mathbf{u}_j^*+\mathbf{a}^{(n)}\right) f\left(\mathbf{M}^{(n)}\mathbf{u}_j^*+\mathbf{a}^{(n)}\left|\mathbf{x}_0^{(n)}\right.\right) \end{array}\right| \\
=\ &\lambda\left(\mathcal{R}_n\right)\left|\begin{array}{c} \int_{\mathcal{C}} \phi\left(\mathbf{M}^{(n)}\mathbf{u}+\mathbf{a}^{(n)}\right) f\left(\mathbf{M}^{(n)}\mathbf{u}+\mathbf{a}^{(n)}\left|\mathbf{x}_0^{(n)}\right.\right)\mathbf{1}_{\mathbf{u}\in\mathcal{E}}\, d\mathbf{u} \\ -\frac{1}{n}\sum_{j=1}^n \phi\left(\mathbf{M}^{(n)}\mathbf{u}_j^*+\mathbf{a}^{(n)}\right) f\left(\mathbf{M}^{(n)}\mathbf{u}_j^*+\mathbf{a}^{(n)}\left|\mathbf{x}_0^{(n)}\right.\right)\mathbf{1}_{\mathbf{u}_j^*\in\mathcal{E}} \end{array}\right| \\
\leq\ &\lambda\left(\mathcal{R}_n\right) V\left(\phi\left(\bullet\right) f\left(\bullet\left|\mathbf{x}_0^{(n)}\right.\right)\mathbf{1}_{\bullet\in\mathcal{E}_n}\right) D_n^*\left(\mathbf{u}_1^*,...,.\mathbf{u}_n^*\right) \\
\leq\ &\kappa\lambda\left(\mathcal{R}_n\right) D_n^*\left(\mathbf{u}_1^*,...,.\mathbf{u}_n^*\right)
\end{aligned}
$$

where κ is some positive constant, $V\left(g\left(\bullet\right)\right)$ is the variation of the fuction g in the sense of Hardy and Krause and $D_n^*\left(\mathbf{u}_1^*,...,.\mathbf{u}_n^*\right)$ is the star discrepancy[2] of the point-set $\mathbf{u}_1^*,...,.\mathbf{u}_n^*$. It suffices to use a sequence of $d-$dimensional rectangles with volume $\lambda\left(\mathcal{R}_n\right)$ that increases at a smaller rate than the star-discrepancy $D_n^*\left(\mathbf{u}_1^*,...,.\mathbf{u}_n^*\right)$ of our low-discrepancy sequence. $\square$

Appendix A.2   The multi-periods proof

We want to prove that

$$
\lim_{n\to\infty}\left|\mathrm{E}\left[\phi\left(\mathbf{X}_t\right)|\mathbf{X}_0=\mathbf{x}_0\right]-\mathrm{E}\left[\phi\left(\mathbf{X}_t^{(n)}\right)\left|\mathbf{X}_0^{(n)}=\mathbf{x}_0^{(n)}\right.\right]\right|=0. \tag{40}
$$

The proof is based on an induction argument on t. The one-period case has been proved in Section Appendix A.1. Assume that Equation (40) is satisfied for $(t-1)$ periods and for any bounded function ϕ that satisfies the conditions of the theorem. We will prove that the result holds for t periods.

Using embedded conditional expectation, we can write

$$
\mathrm{E}\left[\phi\left(\mathbf{X}_t\right)|\mathbf{X}_0=\mathbf{x}_0\right]\ =\ \int_{\mathbb{R}^d}\mathrm{E}\left[\phi\left(\mathbf{X}_t\right)|\mathbf{X}_1=x_1\right] f\left(\mathbf{x}_1|\mathbf{x}_0\right)\, d\mathbf{x}_1
$$

$$
\mathrm{E}\left[\phi\left(\mathbf{X}_t^{(n)}\right)\left|\mathbf{X}_0^{(n)}=\mathbf{x}_0^{(n)}\right.\right]\ =\ \sum_{j=1}^n\mathrm{E}\left[\phi\left(\mathbf{X}_t^{(n)}\right)\left|\mathbf{X}_1^{(n)}=\mathbf{x}_j^{(n)}\right.\right]\frac{f\left(\mathbf{x}_j^{(n)}\left|\mathbf{x}_0^{(n)}\right.\right)}{\sum_{k=1}^n f\left(\mathbf{x}_k^{(n)}\left|\mathbf{x}_0^{(n)}\right.\right)}.
$$

---

[2]

Definition 3 *For a sequence* $\mathbf{u}_1,\mathbf{u}_2,\cdots$ *in* $[0,1]^d$, *the star-discrepancy* $D_n^*\left(\mathbf{u}_1,...,\mathbf{u}_n\right)$ *is defined by*

$$
D_n^*\left(\mathbf{u}_1,...,\mathbf{u}_n\right)=\sup_{B\in\mathcal{J}}\left|\frac{1}{n}\sum_{k=1}^n\mathbf{1}_{\{\mathbf{u}_k\in B\}}-\lambda\left(B\right)\right|
$$

where $\mathbf{1}_{\{\mathbf{u}_k\in B\}}=1$ *if* $\mathbf{u}_k\in B$ *and* 0 *otherwise,* $\mathcal{J}$ *is the family of sets with the form* $\prod_{i=1}^d[0,\beta_i]$ *with* $0\leq\beta_i\leq 1$, *and* λ *is the Lebesgue measure.*

For each $\mathbf{x} \in \mathbb{R}^d$, we define

$$\psi(\mathbf{x}) \equiv \mathrm{E}\left[\phi(\mathbf{X}_{t-1}) | \mathbf{X}_0 = \mathbf{x}\right]$$

and

$$\zeta_n(\mathbf{x}) \equiv \mathrm{E}\left[\phi(\mathbf{X}_{t-1}) | \mathbf{X}_0 = \mathbf{x}\right] - \mathrm{E}\left[\phi\left(\mathbf{X}_{t-1}^{(n)}\right) \Big| \mathbf{X}_0^{(n)} = \mathbf{x}^{(n)}\right]$$

where $\mathbf{x}^{(n)}$ denotes one of the closiest state of $\mathbf{x}$ (just as is Equation (30)).

Starting with the left hand side of Equation (40),

$$\left| \mathrm{E}\left[\phi(\mathbf{X}_t) | \mathbf{X}_0 = \mathbf{x}_0\right] - \mathrm{E}\left[\phi\left(\mathbf{X}_t^{(n)}\right) \Big| \mathbf{X}_0^{(n)} = \mathbf{x}_0^{(n)}\right]\right|$$

$$\leq \left| \int_{\mathbb{R}^d} \mathrm{E}\left[\phi(\mathbf{X}_t) | \mathbf{X}_1 = \mathbf{x}_1\right] f(\mathbf{x}_1 | \mathbf{x}_0) \; d\mathbf{x}_1 - \sum_{j=1}^n \mathrm{E}\left[\phi(\mathbf{X}_t) \Big| \mathbf{X}_1 = \mathbf{x}_j^{(n)}\right] \frac{f\left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)}\right)}{\sum_{k=1}^n f\left(\mathbf{x}_k^{(n)} \Big| \mathbf{x}_0^{(n)}\right)}\right|$$

$$+ \left| \sum_{j=1}^n \left(\mathrm{E}\left[\phi(\mathbf{X}_t) \Big| \mathbf{X}_1 = \mathbf{x}_j^{(n)}\right] - \mathrm{E}\left[\phi\left(\mathbf{X}_t^{(n)}\right) \Big| \mathbf{X}_1^{(n)} = \mathbf{x}_j^{(n)}\right]\right) \frac{f\left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)}\right)}{\sum_{k=1}^n f\left(\mathbf{x}_k^{(n)} \Big| \mathbf{x}_0^{(n)}\right)}\right|$$

$$= \left| \mathrm{E}\left[\psi(\mathbf{X}_1) | \mathbf{X}_0 = \mathbf{x}_0\right] - \mathrm{E}\left[\psi\left(\mathbf{X}_1^{(n)}\right) \Big| \mathbf{X}_0^{(n)} = \mathbf{x}_0^{(n)}\right]\right| + \left| \sum_{j=1}^n \zeta_n\left(\mathbf{x}_j^{(n)}\right) \frac{f\left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)}\right)}{\sum_{k=1}^n f\left(\mathbf{x}_k^{(n)} \Big| \mathbf{x}_0^{(n)}\right)}\right|$$

$$= \left| \mathrm{E}\left[\psi(\mathbf{X}_1) | \mathbf{X}_0 = \mathbf{x}_0\right] - \mathrm{E}\left[\psi\left(\mathbf{X}_1^{(n)}\right) \Big| \mathbf{X}_0^{(n)} = \mathbf{x}_0^{(n)}\right]\right| + \left| \mathrm{E}\left[\zeta_n\left(\mathbf{X}_1^{(n)}\right) \Big| \mathbf{X}_0^{(n)} = \mathbf{x}_0^{(n)}\right]\right|$$

Note that because ϕ is bounded, so is ψ. Moreover, the function $\Psi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ defined as

$$\Psi(\mathbf{x}, \mathbf{z}) = \psi(\mathbf{x}) f(\mathbf{x} | \mathbf{z}) = \left(\int_{\mathbb{R}^d} \phi(\mathbf{y}) f_{t-1}(\mathbf{y} | \mathbf{x}) \; d\mathbf{y} \right) f(\mathbf{x} | \mathbf{z})$$

has bounded variation $V(\Psi)$ on $\mathbb{R}^d \times \mathbb{R}^d$ (it is one of the assumptions of the theorem). Therefore, the first term tends to zero as $n \to \infty$ from the one-period result.

What about the second term? Applying the induction hypothesis, we have for any $\mathbf{x} \in \mathbb{R}^d$.

$$\lim_{n \to \infty} \zeta_n(\mathbf{x}) = 0.$$

Hence, $\zeta_n(\mathbf{X}_1) \to 0$ Q−almost-surely. Moreover, the boundedness of ϕ implies that we can apply the dominated convergence theorem:

$$\lim_{n \to \infty} \mathrm{E}\left[\zeta_n(\mathbf{X}_1) | \mathbf{X}_0 = \mathbf{x}_0\right] = \mathrm{E}\left[\lim_{n \to \infty} \zeta_n(\mathbf{X}_1) | \mathbf{X}_0 = \mathbf{x}_0\right] = 0. \tag{41}$$

Since,

$$\left| \mathrm{E}\left[\zeta_n\left(\mathbf{X}_1^{(n)}\right) \Big| \mathbf{X}_0^{(n)} = \mathbf{x}_0^{(n)}\right]\right| \leq \left| \mathrm{E}\left[\zeta_n\left(\mathbf{X}_1^{(n)}\right) \Big| \mathbf{X}_0^{(n)} = \mathbf{x}_0^{(n)}\right] - \mathrm{E}\left[\zeta_n(\mathbf{X}_1) | \mathbf{X}_0 = \mathbf{x}_0\right]\right|$$

$$+ \left| \mathrm{E} \left[\zeta_n \left(\mathbf{X}_1 \right) | \mathbf{X}_0 = \mathbf{x}_0 \right] \right|,$$

then, to conclude the proof, it suffices to show that

$$\lim_{n \to \infty} \left| \mathrm{E} \left[\zeta_n \left(\mathbf{X}_1^{(n)} \right) \Big| \mathbf{X}_0^{(n)} = \mathbf{x}_0^{(n)} \right] - \mathrm{E} \left[\zeta_n \left(\mathbf{X}_1 \right) | \mathbf{X}_0 = \mathbf{x}_0 \right] \right| = 0. \tag{42}$$

This result may be easily established by mimicing the one-period proof. $\square$

Appendix B   Proofs of intermediate results

Appendix B.1   Proof of Equation (34)

In this section, we prove that Equation (34) implies Equation (33). Indeed, assume that Equation (34) is satisfied which gives

$$\lim_{n \to \infty} \left| \mathrm{E} \left[\phi \left(\mathbf{X}_{t+1} \right) | \mathbf{X}_t = \mathbf{x}_0 \right] - \frac{\lambda \left(\mathcal{R}_n \right)}{n} \sum_{j=1}^{n} \phi \left(\mathbf{x}_j^{(n)} \right) f \left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)} \right) \right| = 0$$

and, as a particular case, if we set $\phi \left(\mathbf{x} \right) \equiv 1$,

$$\lim_{n \to \infty} \left| 1 - \frac{\lambda \left(\mathcal{R}_n \right)}{n} \sum_{j=1}^{n} f \left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)} \right) \right| = 0.$$

We deduce that

$$\lim_{n \to \infty} \left| \frac{\lambda \left(\mathcal{R}_n \right)}{n} \sum_{j=1}^{n} \phi \left(\mathbf{x}_j^{(n)} \right) f \left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)} \right) \right| = \left| \mathrm{E} \left[\phi \left(\mathbf{X}_{t+1} \right) | \mathbf{X}_t = \mathbf{x}_0 \right] \right| \leq \| \phi \|_\infty < \infty$$

and

$$\lim_{n \to \infty} \left| 1 - \frac{1}{\frac{\lambda (\mathcal{R}_n)}{n} \sum_{j=1}^{n} f \left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)} \right)} \right| = 0.$$

Therefore,

$$\left| \mathrm{E} \left[\phi \left(\mathbf{X}_{t+1} \right) | \mathbf{X}_t = \mathbf{x}_0 \right] - \mathrm{E} \left[\phi \left(\mathbf{X}_{t+1}^{(n)} \right) \Big| \mathbf{X}_t^{(n)} = \mathbf{x}_0^{(n)} \right] \right|$$

$$= \left| \mathrm{E} \left[\phi \left(\mathbf{X}_{t+1} \right) | \mathbf{X}_t = \mathbf{x}_0 \right] - \frac{\frac{\lambda (\mathcal{R}_n)}{n} \sum_{j=1}^{n} \phi \left(\mathbf{x}_j^{(n)} \right) f \left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)} \right)}{\frac{\lambda (\mathcal{R}_n)}{n} \sum_{j=1}^{n} f \left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)} \right)} \right| \quad \text{(from Equation (32))}$$

$$\leq \left| \mathrm{E} \left[\phi \left(\mathbf{X}_{t+1} \right) | \mathbf{X}_t = \mathbf{x}_0 \right] - \frac{\lambda \left(\mathcal{R}_n \right)}{n} \sum_{j=1}^{n} \phi \left(\mathbf{x}_j^{(n)} \right) f \left(\mathbf{x}_j^{(n)} \Big| \mathbf{x}_0^{(n)} \right) \right|$$

$$+ \left| \frac{\lambda\left(\mathcal{R}_n\right)}{n} \sum_{j=1}^{n} \phi\left(\mathbf{x}_j^{(n)}\right) f\left(\mathbf{x}_j^{(n)} \middle| \mathbf{x}_0^{(n)}\right) \right| \left| 1 - \frac{1}{\frac{\lambda(\mathcal{R}_n)}{n} \sum_{j=1}^{n} f\left(\mathbf{x}_j^{(n)} \middle| \mathbf{x}_0^{(n)}\right)} \right|$$

$$\rightarrow \quad 0 \text{ as } n \rightarrow \infty. \ \square$$

Appendix B.2   Proof of Equation (37)

Recall that $\lim_{n\to\infty} \left\| \mathbf{x}_0 - \mathbf{x}_0^{(n)} \right\| = 0$. The assumption that the one-period conditional density function $f\left(\mathbf{x} \middle| \mathbf{x}_0\right)$ is continuous in $\mathbb{R}^d \times \mathbb{R}^d$ implies that for any $\mathbf{x} \in \mathbb{R}^d$, $\lim_{n\to\infty} f\left(\mathbf{x} \middle| \mathbf{x}_0^{(n)}\right) = f\left(\mathbf{x} \middle| \mathbf{x}_0\right)$. We will need this last result in the application of Scheffé's Theorem to conclude that $\int_{\mathbb{R}^d} \left| f\left(\mathbf{y} \middle| \mathbf{x}_0\right) - f\left(\mathbf{y} \middle| \mathbf{x}_0^{(n)}\right) \right| d\mathbf{y} \rightarrow 0$ as $n \rightarrow \infty$. To complete the proof, consider

$$\left| \int_{\mathcal{E}_n} \phi\left(\mathbf{y}\right) f\left(\mathbf{y} \middle| \mathbf{x}_0\right) \ d\mathbf{y} - \int_{\mathcal{E}_n} \phi\left(\mathbf{y}\right) f\left(\mathbf{y} \middle| \mathbf{x}_0^{(n)}\right) \ d\mathbf{y} \right|$$

$$\leq \quad \int_{\mathbb{R}^d} \left| \phi\left(\mathbf{y}\right) f\left(\mathbf{y} \middle| \mathbf{x}_0\right) - \phi\left(\mathbf{y}\right) f\left(\mathbf{y} \middle| \mathbf{x}_0^{(n)}\right) \right| \ d\mathbf{y}$$

$$\leq \quad \|\phi\|_\infty \int_{\mathbb{R}^d} \left| f\left(\mathbf{y} \middle| \mathbf{x}_0\right) - f\left(\mathbf{y} \middle| \mathbf{x}_0^{(n)}\right) \right| \ d\mathbf{y} \rightarrow 0 \text{ as } n \rightarrow \infty. \ \square$$

References

[1] Billingsley, P. 1986, *Probability and Measure* second edition, Wiley.

[2] Boyle, P., A. Kolkiewicz and K. Tan, 2000, *Pricing American Style Options Using Low Discrepancy Mesh Methods.*

[3] Broadie, M., 1999, *Low Discrepancy Lattices for Pricing Multi-Dimensional American Options*, unpublished manuscript, Columbia University.

[4] Broadie, M., Glasserman, P., and G. Jain, 1997, Enhanced Monte Carlo Estimates of American Option Prices, *Journal of Derivatives* , (Fall) 25-44.

[5] Chen, R.-R. and T. Yang, 2000, *An Arbitrage Free Model in a Multi-Asset Economy*, unpublished manuscript, Rutgers University and Freddie Mac.

[6] Duan, J.C. and J.G. Simonato, 2000, American Option Pricing under GARCH by a Markov Chain Approximation, *Journal of Economic Dynamics and Control*, forthcoming.

[7] Garcia, D., 2000, *A Monte Carlo Method for Pricing American Options*, unpublished manuscript, Dartmouth College.

[8] Longstaff F. and E. Schwartz, 2000, Valuing American Options By Simulation: A Simple Least-Squares Approach, forthcoming in *Review of Financial Studies.*

[9] Niederreiter H., 1992, *Random Number Generation and Quasi-Monte Carlo Method*, SIAM,.Philadelphia, 237 pages.