

AMERICAN OPTION PRICING WITH RANDOMIZED QUASI-MONTE CARLO SIMULATIONS

Maxime Dion
Pierre L'Ecuyer

DIRO, Université de Montreal
C.P. 6128, Succ. Centre-Ville
Montréal (Québec), H3C 3J7, CANADA

ABSTRACT

We study the pricing of American options using least-squares Monte Carlo combined with randomized quasi-Monte Carlo (RQMC), viewed as a variance reduction method. We find that RQMC reduces both the variance and the bias of the option price obtained in an out-of-sample evaluation of the retained policy, and improves the quality of the returned policy on average. Various sampling methods of the underlying stochastic processes are compared and the variance reduction is analyzed in terms of a functional ANOVA decomposition.

1 INTRODUCTION

American-style financial options are contracts that give the owner a payoff that depends on the information available (such as the prices of certain assets or commodities) at the time when he/she decides to exercise the option. Each contract specifies a (potentially infinite) set of admissible exercise times between the purchase date and the expiration date. The owner wants to find a strategy of exercise that maximizes his expected payoff, and the value (or price) of the option can be written as the expected payoff under such an optimal exercise strategy, where the expectation is taken under a risk-neutral measure. An exercise strategy can be defined as a rule which for each admissible exercise time and each possible “state” (available relevant information) at that time, specifies a decision in the set {exercise now, wait}.

Pricing an American-style option and computing an optimal exercise strategy belongs to a large class of stochastic *optimal stopping problems* for which the standard tools are stochastic control theory and dynamic programming (DP); see [Chow, Robbins, and Siegmund \(1971\)](#), [Bertsekas \(2005\)](#), [Bertsekas \(2007\)](#) and the many references given there. Equivalently, this is also a subclass of Markov decision process problems. The optimal stopping strategy and the corresponding optimal value (the option price, in our case) can be obtained in principle by solving the DP (or Bellman) recurrence equations, but this is not always practical. In some cases, the optimal stopping strategy admits an explicit formula, or has an explicit form (e.g., is defined by a threshold on a state variable), or can be computed numerically to good accuracy by solving the DP equations. But in many situations, especially when the state space is high-dimensional and/or the underlying random variables have too complicated probability distributions, the DP equations are too difficult to solve numerically.

Approximate dynamic programming techniques have been developed over the years to compute approximations to the optimal strategy and optimal value in DP problems; see [Haurie and L'Ecuyer \(1986\)](#), [L'Ecuyer \(1989\)](#), [Bertsekas and Tsitsiklis \(1996\)](#), [Bertsekas \(2007\)](#), [Chang et al. \(2007\)](#), and the references therein. In most of these methods, the value function and/or the policy, which are functions defined over the state space, are parameterized by a finite number of parameters, and these parameters are then “tuned” or “learned” to provide the best possible fit. This can be done via simple least-squares or more elaborate machine learning techniques, including neural networks. When this is done for the value function after the decision is taken, which for the optimal stopping problem means

the value function for the case where we decide not to stop (also called the *continuation value*), this is exactly equivalent to what machine-learning people call *Q-learning*. A popular approach is to select a finite set of real-valued basis functions defined over the state space, write the value function as a linear combination of those basis functions, and estimate the coefficients by least squares from the computed values at a set of evaluation points, at each step of the DP algorithm. If the dimension of the state space is not too large, the evaluation points can be carefully selected to form a grid that covers the relevant part of the state space, as done in [Haurie and L'Ecuyer \(1986\)](#), [L'Ecuyer \(1989\)](#), and [Ben Ameur, Breton, and L'Ecuyer \(2002\)](#), for example.

For high-dimensional state spaces, it becomes difficult to construct such a grid that covers the entire state space, but clever simulation-based (or Monte Carlo (MC)) methods have been designed that generate sample trajectories of the process and take the visited states at any given step as a representative set of evaluation points for the value function. This approach turns out to simplify greatly when the sample trajectories can be generated independently of the decisions taken, which is the case for optimal stopping problem where the available decisions at each step are only to continue or to stop: The simulated trajectories can then be generated simply by assuming that we never stop. [Tsitsiklis and Roy \(2001\)](#) and [Longstaff and Schwartz \(2001\)](#) proposed approximate dynamic programming algorithms based on this observation, using least-squares regression with a small set of basis functions. The variant of [Longstaff and Schwartz \(2001\)](#), called *least-squares Monte Carlo* (LSM), and using polynomials as basis functions, has become the most popular because it provides a price estimator that appears (empirically) to have less bias than that of [Tsitsiklis and Roy \(2001\)](#), which we denote by TvR. Convergence of LSM is studied theoretically by [Clément, Lamberton, and Protter \(2002\)](#), [Stentoft \(2004b\)](#) and [Zanger \(2009\)](#). [Glasserman \(2004\)](#) compares TvR and LSM on some examples.

These methods are not perfect. A good choice of basis functions can be very difficult in high dimensions, convergence is rather slow, and the approximate price is a biased estimator of the true price. The choice of basis functions is studied empirically by [Stentoft \(2004a\)](#). For a given set of basis functions, the accuracy can be improved by using variance reduction methods, such as control variates, importance sampling, and randomized quasi-Monte Carlo (RQMC), for example. RQMC and other variance reduction methods are known to be very effective, when properly applied, for pricing European-style financial options (which cannot be exercised before the expiration date); see [Glasserman \(2004\)](#), [L'Ecuyer \(2009\)](#), [Lemieux \(2009\)](#), and the references given there.

The aim of this paper is to examine the combination of TvR and LSM with RQMC, and make comparisons in terms of the quality of the retained policy. Other experimental work on the combination of LSM with QMC or RQMC has been done recently, all for options based on assets whose sample paths are geometric Brownian motions (GBM). [Chaudhary \(2005\)](#) examines empirically the convergence of LSM combined with a bridge sampling method, with the pseudorandom numbers replaced by (deterministic) Sobol' or Niederreiter point sets, for pricing three types of American options. He observes that the mean square error converges faster with this combination than with ordinary LSM (with MC). [Lemieux \(2004\)](#) compares the empirical performance of various RQMC point sets in LSM on American options; she observed only modest variance reductions, ranging roughly from 1.5 to 15, from using LSM with RQMC relative to LSM with MC. [Lin and Wang \(2008\)](#) and [Wang and Sloan \(2009\)](#) observe larger variance reductions when simulating the sample paths of the underlying asset prices by incorporating change-of-variable techniques to reduce the effective dimension of the integrands, such as bridge sampling and other clever decompositions of the covariance matrix ([Lemieux \(2004\)](#) did not use such techniques). [Lemieux and La \(2005\)](#) and [Wang \(2007\)](#) combine LSM and RQMC with additional variance reduction techniques.

All these authors have considered only the (empirical) variance of the price estimates returned directly by the LSM algorithm. However, this returned price is generally a biased estimate of the expected value associated with the policy returned by LSM, because it is computed from the same sample paths used to optimize it and it also depends on the selected basis functions. An unbiased estimator of this expected value can be obtained from additional *independent* simulation runs with this policy, i.e., from an *out-of-sample* policy evaluation in a second stage.

In this paper, we do that for both TvR and LSM, and we compare various methods in terms of the expected values of the returned policies. We examine the mean, the variance, and the distribution of these expected values. We observe empirically that RQMC reduces significantly both the variance of the returned price estimate and the variance of the expected payoff for a fixed exercise policy, and that reduction is much larger when RQMC is combined with change-of-variable techniques that reduce the effective dimension. The algorithms also return better policies (with higher expected values) when combined with RQMC. This implies that applying RQMC also reduces the bias of the option price obtained in the second stage from an out-of-sample experiment. Interestingly, the policies returned by TvR are roughly as good on average (and better for some examples) as those returned by LSM, even

though the first-stage value estimates typically have more bias for TvR. We also observe that the recently-proposed array-RQMC method brings a bigger improvement for TvR than for LSM. Finally, we compare the RQMC gain with that obtained for European-style options, and find that it is typically smaller for American-style options.

To gain a bit more insight on why and how RQMC works in this context, we perform an empirical ANOVA functional decomposition of the option payoff with a fixed exercise policy, and look at how much of the variance lies in the low-dimensional projections.

In Section 2, we define the problem of pricing an American option. In Section 3, we describe the TvR and LSM methods. We discuss RQMC methods in Section 4. In Section 5, we describe our experimental setting and define the various quantities computed in the experiments. Numerical results are reported in Section 6.

2 AMERICAN OPTION PRICING

We consider a Bermudan *American-style option* with expiration date T and exercise opportunities at the fixed set of dates $0 = t_0 < t_1 < \dots < t_d = T$. The payoff depends on a (possibly multivariate) Markovian stochastic process $\{\mathbf{X}(t), t \geq 0\}$ with state space $\mathcal{X} = \mathbb{R}^c$, where $\mathbf{X}(0) = \mathbf{x}_0 \in \mathcal{X}$ is fixed. We also denote the random variable $\mathbf{X}(t_j)$ by \mathbf{X}_j , for $j = 0, \dots, d$. Thus, $\{\mathbf{X}_j, j = 0, \dots, d\}$ is a discrete-time Markov chain. At each time t_j , the holder observes the state \mathbf{X}_j of the stochastic process, and if he decides to exercise the option, he receives an immediate payoff $g_j(\mathbf{X}_j)$, where $g_j: \mathcal{X} \rightarrow \mathbb{R}$ is a known (measurable) function for each j . When the payoffs are discounted, we assume here that the discounting is contained in $g_j(\mathbf{X}_j)$. An *exercise policy* or *stopping rule* is a sequence $\pi = (\mu_0, \mu_1, \dots, \mu_{d-1})$ of functions $\mu_j: \mathcal{X} \rightarrow \{\text{exercise now, wait}\}$. Such a policy is in fact equivalent to a *stopping time* τ for the stochastic process $\{\mathbf{X}(t), t \geq 0\}$, defined by $\tau = \min\{j \geq 0 : \mu_j(\mathbf{X}_j) = \text{exercise now}\}$.

To any given policy π , or stopping time τ , there corresponds a set of value functions $V_{\pi,j} = V_{\tau,j}$ and $Q_{\pi,j} = Q_{\tau,j}$, from \mathcal{X} to \mathbb{R} , defined recursively by the recurrences

$$V_{\pi,d}(\mathbf{x}) = g_d(\mathbf{x}), \tag{1}$$

$$Q_{\pi,j}(\mathbf{x}) = \mathbb{E}[V_{\pi,j+1}(\mathbf{X}_{j+1}) \mid \mathbf{X}_j = \mathbf{x}], \tag{2}$$

$$V_{\pi,j}(\mathbf{x}) = \begin{cases} g_j(\mathbf{x}) & \text{if } \mu_j(\mathbf{x}) = \text{exercise now,} \\ Q_{\pi,j}(\mathbf{x}) & \text{if } \mu_j(\mathbf{x}) = \text{wait,} \end{cases} \tag{3}$$

for $j = d - 1, \dots, 0$ and for all $\mathbf{x} \in \mathcal{X}$. Everywhere in this paper, all expectations and probabilities are with respect to the risk-neutral measure. Here, $V_{\pi,j}(\mathbf{x})$ represents the expected payoff under policy π when the state is \mathbf{x} at step j , and $Q_{\pi,j}(\mathbf{x})$ is this same expected payoff when we decide not to stop at step j .

The DP equations can be written as

$$V_d(\mathbf{x}) = g_d(\mathbf{x}), \tag{4}$$

$$Q_j(\mathbf{x}) = \mathbb{E}[V_{j+1}(\mathbf{X}_{j+1}) \mid \mathbf{X}_j = \mathbf{x}], \tag{5}$$

$$V_j(\mathbf{x}) = \max[g_j(\mathbf{x}), Q_j(\mathbf{x})], \tag{6}$$

$$\mu_j^*(\mathbf{x}) = \begin{cases} \text{exercise now} & \text{if } g_j(\mathbf{x}) \geq Q_j(\mathbf{x}) \\ \text{wait} & \text{otherwise,} \end{cases} \tag{7}$$

for $j = d - 1, \dots, 0$ and for all $\mathbf{x} \in \mathcal{X}$. The functions Q_j , V_j , and μ_j^* provide the *optimal value*, the *continuation value*, and the *optimal decision*, as a function of the step number j and state \mathbf{x} .

In principle, an optimal policy and the optimal value can be computed by solving these DP equations backward. But unless \mathcal{X} has a finite number of states, the functions Q_j and V_j must be approximated. We usually prefer to approximate Q_j , because it is smoother than V_j . In a linear regression approximation, we select a finite set of *basis functions* $\{\psi_k: \mathcal{X} \rightarrow \mathbb{R}, 1 \leq k \leq m\}$, and we approximate Q_j by

$$\tilde{Q}_j(\mathbf{x}) = \sum_{k=1}^m \beta_{j,k} \psi_k(\mathbf{x})$$

where the vector of coefficients $\beta_j = (\beta_{j,1}, \dots, \beta_{j,m})^t$ is to be determined. In least-squares regression, we obtain (in some way) an approximation of $Q_j(\mathbf{x})$ at a finite set of points $\mathbf{x}_{1,j}, \dots, \mathbf{x}_{n,j}$ in \mathcal{X} , say $W_{i,j}$ is an approximation of $Q_j(\mathbf{x}_{i,j})$ for each i , and we compute the vector of coefficients β_j that minimize the sum of squares

$$\min_{\beta_j \in \mathbb{R}^m} \sum_{i=1}^n (\tilde{Q}_j(\mathbf{x}_{i,j}) - W_{i,j+1})^2.$$

To any given approximation \tilde{Q}_j , $j = 0, \dots, d-1$, there corresponds a stopping policy $\tilde{\pi}$ (or stopping time $\tilde{\tau}$) defined by $\tilde{\tau} = \min\{j \geq 0 : g_j(\mathbf{X}_j) \geq \tilde{Q}_j(\mathbf{X}_j)\}$. In the terminology of Bertsekas (2005), this policy is the *one-step look-ahead* (1SL) policy associated with these \tilde{Q}_j . Let us denote by $V_{\tilde{\pi},j}$ the value function associated with this policy, at step j . Since this policy cannot beat the optimal policy, we obviously have $V_{\tilde{\pi},j}(\mathbf{x}) \leq V_j(\mathbf{x})$ for all j and \mathbf{x} .

In more traditional approximate DP (Haurie and L'Ecuyer 1986, L'Ecuyer 1989, Bertsekas 2005), the states $\mathbf{x}_{i,j}$ are fixed a priori and the basis functions are often taken as piecewise linear in each coordinate or product of splines. In the TvR and LSM method, they are the realizations of random variables; they are taken as the states visited at step j by n independent realizations of the Markov chain. This can be done because the sample paths can be simulated independently of the decision policy, which is not possible for general DP models. This permits one to sidestep the curse of dimensionality and to obtain a set of representative states concentrated in the most frequently visited areas, where things really matter. The TvR and LSM algorithms are given in the next section.

3 SIMULATION-BASED DP ALGORITHMS FOR AMERICAN OPTION PRICING

The TvR regression-based algorithm of Tsitsiklis and Roy (2001) is the following:

1. Simulate n independent trajectories of the Markov chain $\{\mathbf{X}_j, j = 0, \dots, d\}$, and let $\mathbf{X}_{i,j}$ be the state for trajectory i at step j .
2. Let $W_{i,d} = g_d(\mathbf{X}_{i,d})$ for $i = 1, \dots, n$.
3. For $j = d-1, \dots, 0$ do:
 - 3a. Compute the vector of coefficients β_j that minimizes

$$\sum_{i=1}^n \left(\sum_{k=1}^m \beta_{j,k} \psi_k(\mathbf{X}_{i,j}) - W_{i,j+1} \right)^2.$$

// Note: $W_{i,j+1}$ is our estimate of $Q_k(\mathbf{X}_{i,j})$. The function \tilde{Q}_j is now defined everywhere.

- 3b. Let $W_{i,j} = \max[g_j(\mathbf{X}_{i,j}), \tilde{Q}_j(\mathbf{X}_{i,j})]$, $i = 1, \dots, n$.
4. Return $\hat{Q}_0(\mathbf{x}_0) = (W_{1,0} + \dots + W_{n,0})/n$ as an estimate of the option price $Q_0(\mathbf{x}_0)$.

For the LSM method of Longstaff and Schwartz (2001), the only difference is in step 3b, which becomes:

- 3b. For $i = 1, \dots, n$, let

$$W_{i,j} = \begin{cases} g_j(\mathbf{X}_{i,j}) & \text{if } g_k(\mathbf{X}_{j,k}) \geq \tilde{Q}_j(\mathbf{X}_{i,j}); \\ W_{i,j+1} & \text{otherwise.} \end{cases}$$

Here, when we do not exercise, we take $W_{i,j+1}$ instead of the approximation $\tilde{Q}_j(\mathbf{X}_{i,j})$ as the continuation value for sample path i . In this case, $W_{i,j+1}$ is the future payoff for this particular sample trajectory, if we use the policy constructed so far from step $j+1$ onward. Longstaff and Schwartz (2001) also recommend culling the points for which the exercise value is zero (out-of-the-money points) before doing the least-squares fit of the holding value function; they say that the resulting fit of the continuation value is then more accurate in the region where it intersects with the exercise function and the resulting exercise policy is likely to be closer to the optimal policy.

There are two main sources of error for the results returned by these algorithms: (1) the Monte Carlo error, due to the fact that n is finite (it can be reduced by increasing n or using variance reduction methods, including RQMC) and (2) the distance between each exact function Q_j and the linear space generated by the selected basis functions (this error cannot be reduced by increasing n , but only by enriching the set of basis functions). The bias on $Q_0(\mathbf{x}_0)$ can have either

sign for both methods, but most often it is positive for TvR and negative for LSM. Longstaff and Schwartz (2001), Clément, Lamberton, and Protter (2002), Stentoft (2004b), and Zanger (2009) study the convergence of the LSM method.

At the end of the TvR or LSM algorithm, the approximations \hat{Q}_j define a ISL policy $\tilde{\pi}$ or stopping time $\tilde{\tau}$ as explained earlier, and an unbiased estimator of the value $V_{\tilde{\pi}}(\mathbf{x}_0)$ for this (fixed) policy can be obtained by performing simulation runs with it, independent of those used for computing the \hat{Q}_j , and taking the average. We call this the *policy evaluation* stage (or the second stage). It can use a different number of runs (not necessarily n) and also different variance reduction techniques than for the first stage. Since the policy $\tilde{\pi}$ cannot beat the optimal one, this provides a negatively-biased estimator of the true option price.

Let $\hat{Q}_{\tilde{\pi},0}(\mathbf{x}_0)$ be the average returned in the policy evaluation stage. The variance of this random variable can be decomposed as a sum of two terms, namely the variance due to the randomness in the returned policy $\tilde{\pi}$ and the expected variance conditional on $\tilde{\pi}$:

$$\begin{aligned} \text{Var}[\hat{Q}_{\tilde{\pi},0}(\mathbf{x}_0)] &= \text{Var}[\mathbb{E}[\hat{Q}_{\tilde{\pi},0}(\mathbf{x}_0) \mid \tilde{\pi}] + \mathbb{E}[\text{Var}[\hat{Q}_{\tilde{\pi},0}(\mathbf{x}_0) \mid \tilde{\pi}]] \\ &= \text{Var}[V_{\tilde{\pi}}(\mathbf{x}_0)] + \mathbb{E}[\text{Var}[\hat{Q}_{\tilde{\pi},0}(\mathbf{x}_0) \mid \tilde{\pi}]]. \end{aligned} \tag{8}$$

Let us denote by W_{π} the random payoff with policy π . If the second stage uses MC with n' independent runs to evaluate $\tilde{\pi}$, the variance conditional on $\tilde{\pi}$ is $\text{Var}[W_{\tilde{\pi}}]/n'$. If the n' runs are not independent, which happens for instance if we use RQMC in the second stage, this conditional variance has a more complicated expression. In practice, we may want to choose n' large enough so that the second term in (8) is small compared to the first.

For a given total available computing budget, one may allocate part of the budget to run the first stage say r times, then make second-stage evaluations to try to find the best of the r returned policies and estimate the corresponding value. In the second stage, we may allocate (dynamically) different fractions of the budget (different values of n') to the different policies, as in ranking and selection procedures. An interesting problem in this setting (not addressed in this paper) is how to do all of this optimally, at least in the asymptotic sense when the computing budget gets large.

4 RANDOMIZED QUASI-MONTE CARLO (RQMC)

RQMC methods are designed to reduce the variance when estimating smooth integrals over the s -dimensional unit hypercube for some moderate integer $s > 0$ (Niederreiter 1992, Owen 1998, L'Ecuyer 2009, Lemieux 2009). Practically any expectation estimated by Monte Carlo simulation can be written as the integral of some function f over such an hypercube (L'Ecuyer 2009). An RQMC method estimates the integral f by evaluating f over an RQMC point set $P_n = \{\mathbf{U}_1, \dots, \mathbf{U}_n\}$ and taking the average, $\hat{\mu}_{\text{rqmc},n} = (1/n) \sum_{i=1}^n f(\mathbf{U}_i)$. An RQMC point set P_n must satisfy the two conditions:

- (a) it covers $(0,1)^s$ very uniformly when taken as a set and
- (b) each point \mathbf{U}_i has the uniform distribution over $(0,1)^s$ when taken individually.

Condition (b) ensures that the average $\hat{\mu}_{\text{rqmc},n}$ is an unbiased estimator of $\mu = \int_{(0,1)^s} f(\mathbf{u})d\mathbf{u}$. Condition (a) could have several different meanings, depending on the definition of uniformity that we want to adopt, and this gives rise to various point set constructions (Niederreiter 1992, L'Ecuyer 2009).

RQMC can provably increase the convergence rate of the variance as a function of n compared with ordinary MC, if f is smooth enough. For example, for Sobolev classes of functions f whose partial derivative of order 2 with respect to any subset of coordinates is square integrable, there exist RQMC point sets P_n for which the variance converges as $\mathcal{O}(n^{-4+\delta})$ for any $\delta > 0$ (Sloan and Joe 1994, Hickernell 2002). This is much faster than MC.

The price of a European-style option can usually be written as an integral over the unit cube, for which RQMC has been applied very successfully (Glasserman 2004, L'Ecuyer 2009, Lemieux 2009). Here, each point \mathbf{U}_i of P_n is used to simulate one copy of the trajectory of the underlying stochastic process, and the dimension s of these points is the number of uniform random variables required to generate one sample path. But even though much smaller variances are often observed empirically, the corresponding integrands are typically not smooth enough to satisfy the usual conditions under which a faster convergence rate than for MC can be proved.

RQMC methods tend to work much better when f is well approximated by a sum of low-dimensional functions, that depend on just a few coordinates of \mathbf{u} , in the following sense (Owen 1998, L'Ecuyer 2009). If

$$\sigma^2 = \text{Var}[f(\mathbf{U})] = \int_{(0,1)^s} f(\mathbf{u})d\mathbf{u} - \mu^2 < \infty$$

for \mathbf{U} uniformly distributed over $(0,1)^s$, one can make a *functional ANOVA decomposition* of f of the form

$$f(\mathbf{u}) = \mu + \sum_{\mathbf{u} \subseteq \mathcal{S}, \mathbf{u} \neq \emptyset} f_{\mathbf{u}}(\mathbf{u}) \tag{9}$$

where each $f_{\mathbf{u}} : (0,1)^s \rightarrow \mathbb{R}$ depends only on $\{u_i, i \in \mathbf{u}\}$, the $f_{\mathbf{u}}$'s integrate to zero and are orthogonal, and the variance decomposes as $\sigma^2 = \sum_{\mathbf{u} \subseteq \mathcal{S}} \sigma_{\mathbf{u}}^2$ where $\sigma_{\mathbf{u}}^2 = \text{Var}[f_{\mathbf{u}}(\mathbf{U})]$. If $\sum_{\{\mathbf{u}: |\mathbf{u}| \leq s'\}} \sigma_{\mathbf{u}}^2$ is very close to σ^2 for some small s' , we say that f has low effective dimension in the superposition sense, while if this holds for the sum over $\mathbf{u} \subseteq \{1, \dots, s'\}$, we have low effective dimension in the truncation sense (Caflish, Morokoff, and Owen 1997). When this happens, we can focus on constructing the point sets P_n so that their projections over the subsets of coordinates involved in the sum have high uniformity, and give less weight to the other projections.

The effective dimension of f can often be reduced via a change of variables (Moskowitz and Caflish 1996; Glasserman 2004; L'Ecuyer, Parent-Chartier, and Dion 2008; L'Ecuyer 2009). For example, if the trajectory of the Markov chain $\{\mathbf{X}_j, j=0, \dots, d\}$ can be written as a function of a multivariate Brownian motion $\{\mathbf{B}(t), t \geq 0\}$ observed at times t_1, \dots, t_d , then one can sample the vector $(\mathbf{B}(t_1), \dots, \mathbf{B}(t_d))$ by sampling the increments sequentially (Seq) as in a random walk. An alternative is to sample in a Brownian bridge (BB) fashion, by generating first $\mathbf{B}(t_d)$, then $\mathbf{B}(t_{d/2})$, then $\mathbf{B}(t_{d/4})$ and $\mathbf{B}(t_{3d/4})$, and so on recursively (using the conditional distribution at each step, where we assume for simplicity that d is a power of 2). More generally, if $(\mathbf{X}_1, \dots, \mathbf{X}_d)$ can be written as a function an s -dimensional multinormal random vector $\mathbf{Y} = (Y_1, \dots, Y_s)^t$ with mean zero and covariance matrix Σ , one can select a decomposition of the form $\Sigma = \mathbf{C}\mathbf{C}^t$ (in any manner) and return $\mathbf{Y} = \mathbf{C}\mathbf{Z}$ where \mathbf{Z} is an s -dimensional vector of independent standard normal random variables (with mean 0 and variance 1), generated by inversion from a s -dimensional vector \mathbf{U} of independent uniform random variables over $(0,1)$. One special case of the latter is the eigendecomposition of Σ used in principal component analysis (PCA), for which the columns of \mathbf{C} are the eigenvectors of Σ sorted by decreasing order of the corresponding eigenvalues. With this decomposition, the first s' uniform random variables, which correspond to the first few coordinates of the points in the case of RQMC, account for the maximal amount of variance of \mathbf{Y} , for each $s' \leq s$.

The results reported in this paper are for a single type of construction, namely Sobol' nets with a left matrix scramble and a random digital shift (Sobol' 1967, Niederreiter 1992, L'Ecuyer 2009), using the default parameters provided in the SSJ library (L'Ecuyer 2008). We have obtained similar results with other types of RQMC point sets, for example randomly-shifted rank-1 lattice rules with a baker's transformation (Hickernell 2002, L'Ecuyer 2009).

In the case of American-style options, the TvR and LSM algorithms return an estimate $\hat{Q}_0(\mathbf{x}_0)$ which is the average of n function evaluations, $W_{1,0}, \dots, W_{n,0}$, each of which being a function of the uniform random numbers that drive the simulation. These random numbers can be replaced by a set of n RQMC points in the same way as for European-style options. However, here each $W_{i,j}$ depends on all the points of P_n , and not only on \mathbf{U}_i . This should makes the (theoretical) RQMC convergence rate analysis even more complicated. Here we only have empirical results. Note that the fact that the option can be exercised early and at different steps for different trajectories can make the integrand less smooth as a function of the underlying uniforms, and for this reason we may expect RQMC to bring a more modest efficiency improvement for American-style options than for their European-style counterparts. This is what we will observe in our experiments, and this is specially true for the array-RQMC method described below. We will also observe in our experiments that the payoff functions have a higher effective dimension for American options than for the European ones.

When the number of steps d of the Markov chain is large, the RQMC points must be high-dimensional and the method often becomes ineffective. A different RQMC method, called *array-RQMC* has been developed recently for this type of situation (L'Ecuyer, Lécot, and Tuffin 2008; L'Ecuyer, Lécot, and L'Archevêque-Gaudet 2009). The method simulates n copies of the chain at the same time. At each iteration, it advances the states of all copies by one step using an RQMC point set, after matching each copy with an RQMC point using a cleverly selected bijection. The method works by making the empirical distribution of the n states, at each step j , closer to the theoretical distribution of \mathbf{X}_j than if the n copies were simulated independently. Matching the chains with the points at each

step involves a sorting of the states, which adds computational overhead compared with ordinary MC and RQMC, but the variance is sometimes reduced by a much larger factor than this overhead, resulting in a large net gain in efficiency (L'Ecuyer, Lécot, and Tuffin 2008; L'Ecuyer, Lécot, and L'Archevêque-Gaudet 2009). To combine TvR and LSM with array-RQMC, it suffices to simulate the n sample paths with array-RQMC.

5 OUTLINE OF THE NUMERICAL EXPERIMENTS

In this section, we describe how we have made our numerical experiments. The numerical results are given in the next section.

In all our examples, the values of the underlying process at the d observation times is a function of a multinormal vector. For RQMC, we generate this vector by three methods: sequentially (Seq), then with a Brownian bridge approach (BB), then with an eigen-decomposition (PCA). We also try the array-RQMC method, and standard MC. The RQMC methods use Sobol' point sets (Sobol' 1967) with $n = 2^{12}$ or $n = 2^{14}$ points, with a left matrix scramble and a random digital shift (Owen 2003, L'Ecuyer 2009). The parameters of those point sets (the "direction numbers") are taken from Lemieux, Cieslak, and Luttmner (2004) and L'Ecuyer (2008) and were selected by paying more attention to the uniformity of the one- and two-dimensional projections. We also tried removing the left matrix scramble and the results were similar.

For each method, we perform $r = 1000$ independent replications of the TvR and LSM algorithms. This yields r independent replicates of the first-stage estimator $\hat{Q}_0(\mathbf{x}_0)$ and r exercise policies $\tilde{\pi}$. We take a large r because we want to estimate the variance of $\hat{Q}_0(\mathbf{x}_0)$ with meaningful accuracy. We then evaluate each of those policies in an out-of-sample second stage to estimate $V_{\tilde{\pi},0}(\mathbf{x}_0)$. Our main goal here is to estimate $\mathbb{E}[V_{\tilde{\pi},0}(\mathbf{x}_0)]$, where the expectation is with respect to $\tilde{\pi}$. The variance of our estimator has the decomposition (8), and we reduce the second term in this decomposition by using RQMC with PCA in all cases for the second stage, with $n' = n$. Even by doing this, the variance of the second term still dominates in our examples.

In separate experiments with very large sample sizes, we estimated the true option price with sufficient accuracy to be able to make statements about the bias of each method. These separate experiments were made with the same sets of basis functions as the regular runs, which means that our bias estimates do not remove the bias due to the choice of basis functions, but only that due to the nonlinearity of the estimators and the fact that the optimization is not perfect (this is the bias from the first source of error mentioned in Section 3).

For each method, we report the empirical mean of the r replicates of $\hat{Q}_0(\mathbf{x}_0)$ and $V_{\tilde{\pi},0}(\mathbf{x}_0)$ (in which $\tilde{\pi}$ is random), as well as 95% confidence intervals on $\mathbb{E}[\hat{Q}_0(\mathbf{x}_0)]$ and $\mathbb{E}[V_{\tilde{\pi},0}(\mathbf{x}_0)]$. For $\hat{Q}_0(\mathbf{x}_0)$, we also report the variance estimated from the r replicates and the *variance reduction factor* (VRF), defined as the variance of the MC estimator divided by that of the current (RQMC) estimator. Its interpretation is that the number of simulation runs must be multiplied by this factor if we use MC instead of RQMC and want the same variance for the average. Note that this factor generally depends on n . Our reported variances are actually multiplied by n , so they can be loosely interpreted as "variances per sample path". In one case, we plot the empirical distributions of $\hat{Q}_0(\mathbf{x}_0)$ and $V_{\tilde{\pi},0}(\mathbf{x}_0)$.

For the TvR method, we also report the empirical mean, variance, and confidence interval, for what we call the *ex post* value, defined as the value estimate obtained when we run n separate simulations with the retained policy $\tilde{\pi}$, but with the same random numbers (and sample paths) that we used in the algorithm. Those are *in-sample* additional runs. Comparing these *ex-post* values with the out-of-sample estimates may give some idea of how much of the bias originates from the overfitting of the retained policy to the particular sample paths that were generated. Note that for the LSM method, the *ex post* values are exactly the same as those returned by the algorithm.

To assess (and compare) how the RQMC methods perform for American-style options with a fixed policy, and for European-style options, we compute and report similar quantities for these two settings as well. The fixed policy was selected as an arbitrary near-optimal policy. We tried several ones and found very little sensitivity to its choice. Here, there is no optimization and the estimators are unbiased for the value under that policy, but we still have the feature that the different trajectories stop at different steps. For the European-style option, on the other hand, all trajectories stop after the same number of steps, and use all coordinates of each RQMC point. We observe that RQMC tends to provide a larger improvement in this case.

A natural criterion for comparing the performance of the methods would be in terms of the mean squared errors (MSE) of the returned option prices in the second stage, $\hat{Q}_{\tilde{\pi},0}(\mathbf{x}_0)$. This MSE is the squared bias, which can be estimated as the difference between the $\mathbb{E}[V_{\tilde{\pi},0}(\mathbf{x}_0)]$ entries in the table and the exact option value, plus the variance given in (8). This variance can be estimated by the empirical variance of $\hat{Q}_{\tilde{\pi},0}(\mathbf{x}_0)$ over the r independent replications. Note that the second variance term in (8) depends on the evaluation method used in the second stage.

We may also be interested in estimating only the first term in (8), which represents the variance of the exact value associated with the retained policy. One way of doing this is by using a second-stage evaluation method that would make the second variance term negligible with respect to the first. We did not do that for our examples because it would have required an excessively large amount of work in the second stage. Another way is to estimate the total variance of the second-stage price, then the second variance term (for a fixed policy), and subtract.

In the cases where the state of the Markov chain is one-dimensional, and where the exercise policy can be characterized by a single threshold value at each time step, we examine and compare the mean and variance of this threshold for the TvR and LSM methods, and their combination with RQMC, at a fixed (selected) time step. The idea is to see how RQMC affects the stability of the retained policy.

To get further insight on the performance of RQMC for these examples, we perform a functional ANOVA decomposition of the payoff function under the same fixed near-optimal policy as earlier, and also for the European-style option, for the different sampling schemes (Seq, BB, and PCA). We estimate the variance contributions of all the one- and two-dimensional ANOVA components, using the method of Sobol' and Myshetskaya (2007), from 20 replications with 2^{16} RQMC points. For this, we also used Sobol' point sets with random digital shifts. This variance contribution turns out to be drastically reduced (to a nearly negligible quantity) when the TvR and LSM algorithms are combined with RQMC. We could estimate the higher-dimensional ANOVA components as well, but this would be more computationally expensive. In our experiments, their variance contributions is not reduced as much.

6 NUMERICAL RESULTS

6.1 A Simple American Put

We first consider an American put option on an underlying asset price that obeys a geometric Brownian motion $\{X(t), t \geq 0\}$ with drift parameter (riskless interest rate) $\mu = 0.05$, volatility $\sigma = 0.08$, and initial value $X(0) = x_0 = 100$. The contract is valid for one year and the exercise dates are $t_j = j/16$ for $j = 1, \dots, 16$ (in years). Putting $X_j = X(t_j)$, a trajectory at these observation dates can be generated as

$$X_{j+1} = X_j \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)(t_{j+1} - t_j) + \sigma\sqrt{t_{j+1} - t_j}Z_{j+1}\right),$$

for $j = 0, \dots, d-1$, where Z_1, \dots, Z_d are independent standard normals. The payoff at step j is then $g_j(X_j) = \exp(-0.05t_j)\max(0, K - X_j)$, where $K = 101$ is the strike price.

For LSM, the basis functions are simple polynomials, as recommended by Stentoft (2004a), up to order 4: $\psi_1(x) = 1$ and $\psi_k(x) = (x - 101)^{k-1}$ for $k = 2, \dots, 5$. These polynomials are centered at the strike price to improve the numerical stability of the least-squares fit. For TvR, we add two more basis functions: $\psi_6(x) = \max(0, x - 101)$ and $\psi_7(x) = (\max(0, x - 101))^2$. With TvR, the continuation value functions are fitted over the entire state space and are somewhat less simple than the truncated continuation value functions of LSM.

The true value of the option has been estimated at 2.1690, from 4000 replications of LSM with RQMC, PCA, and $n' = 2^{20}$, while keeping the same basis functions as above. This value is accurate to the given digits, given the choice of basis functions. The estimated values in Table 1 can be compared to this number to get an idea of the part of the bias not due to the choice of basis functions. We also observed that increasing the number of basis functions to include polynomials up to order 6 increases the estimated value to 2.1691 with LSM. For the TvR method, with the basis functions given above, the value resulting from the policy in the limit of an infinite number of simulations is estimated at 2.1674. The best policy that can be found with TvR is therefore not as good as the one from LSM, for the given set of basis functions. This gap is only due to the fact that the continuation values Q_j cannot be described accurately over their whole domain by a linear combination of the basis functions.

The $\hat{Q}_0(x_0)$ values returned by LSM have slight positive bias (their expectation appears larger than 2.1690). The out-of-sample values $V_{\hat{\pi},0}(x_0)$ are always biased low relatively to the optimal value $V_0(\mathbf{x}_0)$, as expected. For TvR, the $\hat{Q}_0(x_0)$ values are slightly negatively biased (relatively to their best possible value of 2.1674), the ex post values have no perceptible bias (high and low sources of bias mostly cancelling one another). The TvR out-of-sample values are always biased low and this bias is on average smaller than the corresponding bias for LSM. However it is a much larger bias relatively to the true option value (2.1691). The RQMC methods also provide higher out-of-sample values than standard MC, which means that they return better policies and second-stage values with smaller bias.

	MC	RQMC			array-RQMC
		Seq	BB	PCA	
LSM					
$\mathbb{E}[\hat{Q}_0(x_0)]$	2.170 ± 0.001	2.1697 ± 0.0005	2.1694 ± 0.0004	2.1692 ± 0.0004	2.1693 ± 0.0005
variance	5.5	1.0	0.55	0.65	0.92
VRF	–	6	10	9	6
$\mathbb{E}[V_{\tilde{\pi},0}(x_0)]$	2.1671 ± 0.0004	2.1685 ± 0.0004	2.1682 ± 0.0004	2.1682 ± 0.0004	2.1686 ± 0.0004
TvR					
$\mathbb{E}[\hat{Q}_0(x_0)]$	2.167 ± 0.001	2.1667 ± 0.0002	2.1666 ± 0.0001	2.1665 ± 0.0001	2.1665 ± 0.00003
variance	4.5	0.21	0.028	0.066	0.0059
VRF	–	20	160	70	750
ex post value	2.167 ± 0.001	2.1678 ± 0.0005	2.1675 ± 0.0004	2.1674 ± 0.0004	2.1671 ± 0.0005
variance	5.7	1.0	0.55	0.61	0.97
VRF	–	5	10	9	6
$\mathbb{E}[V_{\tilde{\pi},0}(x_0)]$	2.1664 ± 0.0004	2.1672 ± 0.0004	2.1671 ± 0.0004	2.1671 ± 0.0004	2.1675 ± 0.0004
Fixed policy					
$V_{\pi,0}(x_0)$	2.168 ± 0.001	2.1693 ± 0.0005	2.1687 ± 0.0004	2.1691 ± 0.0004	2.1684 ± 0.0005
variance	5.9	0.94	0.57	0.63	0.92
VRF	–	6	10	9	6
European put					
variance	9.4	0.51	0.00010	0.013	0.00016
VRF	–	20	90000	700	60000

Table 1: Evaluation of an American put option with 1000 independent replications with $n = 2^{14}$. The value of this option is 2.1691, accurate to the given digits. The variances reported are the sample variances over the 1000 replications, multiplied by 2^{14} . The rows marked $\mathbb{E}[\hat{Q}_0(x_0)]$ give unbiased estimates of this expectation from the average over the 1000 replications of $\hat{Q}_0(x_0)$, together with a 95% confidence interval for this expectation. The rows marked $\mathbb{E}[V_{\tilde{\pi},0}(x_0)]$ give out-of-sample unbiased estimates of this expectation, in which $\tilde{\pi}$ must be interpreted as a random variable, all done with RQMC, PCA and $n' = n$ for each $\tilde{\pi}$. We denote by π the selected near-optimal policy; its value is also estimated with 1000 replications with $n = 2^{14}$.

Looking at the variances and VRFs in Table 1, we see that RQMC provides a significant variance reduction, compared to MC, in all the settings considered, and that the variance reduction on the first-stage value is much larger for TvR than for LSM, especially for array-RQMC. The reported variances are accurate to roughly 10%. The variances and VRFs for the different RQMC samplings are about the same for LSM, for the ex post value of TvR, and for a fixed policy. For those, RQMC with BB and PCA are the best performers, and array-RQMC provides about the same reduction variance as RQMC with Seq, while its computational cost in this particular example is nearly twice that of the other methods. Note that RQMC with the Seq, BB, and PCA sampling schemes all take about the same time as MC, approximately 0.3 second per replication of 2^{14} simulations on a common desktop computer. For TvR, array-RQMC provides the best VRF, followed by RQMC with BB.

Note that for $\hat{Q}_0(x_0)$ from LSM and TvR ex post values, part of the variance originates from the variation of $\tilde{\pi}$ and part of it comes from the variance given $\tilde{\pi}$, and this second contribution should be roughly the same as the variance of the estimator for a fixed policy. The fact that the latter variance is about the same as for LSM and for the TvR ex post value means that the variation of $\tilde{\pi}$ has a negligible variance contribution in this example.

Figure 1 provides histograms of the 1000 price estimates returned by LSM with MC in the first stage (left) and in the second stage (right). For each of the 1000 policies, the second-stage policy evaluation was performed with RQMC and PCA as earlier, but this time with $n' = 2^{16}$ replicated 16 times, to get a better estimate of $V_{\tilde{\pi},0}(x_0)$ for each $\tilde{\pi}$ and to estimate its variance. The distribution on the right is visibly non-normal and negatively skewed. This means that $V_{\tilde{\pi},0}(x_0)$ is often close to its upper bound $V_0(x_0)$, but once in a while it is much smaller because the algorithm returns a poor policy. The estimated values of $V_{\tilde{\pi},0}(x_0)$ for the 1000 retained policies range from 2.160 to 2.170. The standard error on the estimation of each $V_{\tilde{\pi},0}(x_0)$ is around 0.00075; its square corresponds to the last term in (8). This source of variance is significant in Figure 1, but it does not obfuscate the variation due to the different quality of the policies $\tilde{\pi}$, represented by the first term on the right in (8). In similar plots with RQMC, the right histogram has a smaller spread and is more symmetric; this corresponds to the fact that the variance from

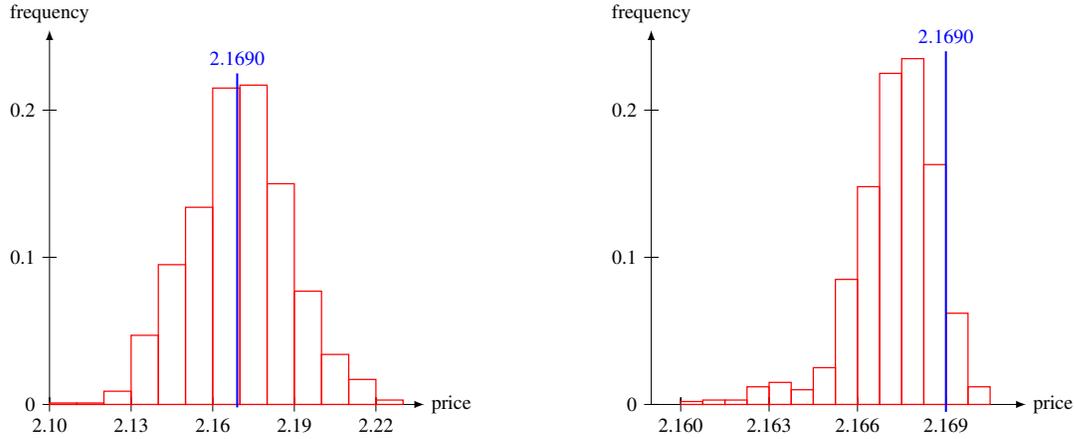


Figure 1: Left: histogram of the 1000 replications of $\hat{Q}_0(x_0)$ for the LSM method with MC. The blue vertical line indicates the exact option value (2.1690). Right: histogram of the 1000 second-stage (out-of-sample) price estimates for LSM with MC. The standard error on $V_{\hat{\pi},0}(x_0)$ for each policy $\hat{\pi}$ is 0.00075, and this value matches the width of the rectangles, so a 95% confidence interval would span roughly four rectangles horizontally.

	MC	RQMC			array-RQMC
		Seq	BB	PCA	
LSM					
Policy threshold	97.67 ± 0.04	97.71 ± 0.01	97.712 ± 0.007	97.719 ± 0.006	97.712 ± 0.006
variance	0.38	0.036	0.014	0.010	0.010
VRF	–	10	30	40	40
TvR					
Policy threshold	97.20 ± 0.09	97.33 ± 0.06	97.445 ± 0.003	97.444 ± 0.003	97.445 ± 0.002
variance	2.1	1.0	0.0025	0.0025	0.0009
VRF	–	2	800	800	2000

Table 2: Empirical mean and variance of the estimated threshold γ_j at step $j = 10$, for the 1000 replications of Table 1, and 95% confidence intervals on the expected threshold returned by the algorithm. The VRFs are for RQMC compared with MC.

the varying quality of the policies in that case is smaller than the variance from the estimation of the $V_{\hat{\pi},0}(x_0)$'s by simulation.

For the corresponding European option, where early exercise opportunities are removed, the exact price can be computed from the standard Black-Scholes formula (1.5489). We simulated this European option over the same 16 time steps as the American one (although a single time step to the final time would suffice), and computed the VRFs with respect to MC, for comparison. These VRFs are much larger here, in particular for array-RQMC and for RQMC with BB (approximately 90,000). This last result is not surprising because the payoff depends only on the final asset value, which is generated directly from the first coordinate with BB, and this makes the problem a one-dimensional integration.

For this example, the state space of the stochastic process is one-dimensional and all considered policies are defined by a threshold γ_j at each step j ; we exercise if and only if $X_j \geq \gamma_j$. Table 2 reports the empirical mean and variance of the estimated threshold at step $j = 10$ (out of 16), over the 1000 replications, as well as 95% confidence intervals of the expected threshold returned by the algorithm. An accurate estimate of the optimal threshold is 97.707, with the LSM basis functions, and 97.441, with the TvR basis functions. We observe that RQMC reduces the variance of the threshold estimator, i.e., the variance of the policy estimate, by large factors. The array-RQMC

method performs very well here. For other values of j , the variance reduction provided by RQMC is smaller [larger] for the smaller [larger] values of j .

	American with fixed policy π			European		
	Seq	BB	PCA	Seq	BB	PCA
one-dimensional components	65%	72%	76%	60%	100%	86%
one- and two-dimensional components	73%	84%	84%	93%	100%	99%

Table 3: Proportion of the variance captured by the one-dimensional ANOVA components, and by the one- and two-dimensional ANOVA components, for an American and European put option.

Seq		BB		PCA	
dim	proportion	dim	proportion	dim	proportion
{1}	16%	{1}	45%	{1}	65%
{2}	14%	{2}	14%	{2}	7%
{3}	10%	{3}	6%	{1,2}	4%
{4}	7%	{1,2}	5%	{3}	2%
{5}	5%	{5}	2%	{1,3}	1%
{6}	4%	{1,3}	1%	{2,3}	1%
{7}	3%	{3,5}	1%	{4}	1%
{1,2}	2%	{2,3}	1%	{5}	0.3%
{8}	2%	{4}	1%	{1,4}	0.3%
{9}	2%	{1,4}	1%	{3,4}	0.3%

Table 4: Specific variance contributions of the most important one- and two-dimensional ANOVA components, by decreasing order, for the American option with the Seq, BB, and PCA sampling schemes.

Tables 3 and 4 give the variance contributions of the one- and two-dimensional terms in the ANOVA decomposition of the integrand for a fixed policy π . Table 4 shows that for BB and PCA, most of this variance contribution comes from the first three coordinates. We made a similar decomposition for several good policies π and the results were similar for all π 's. Here, the nominal dimension of the integral is 16 (the number of time steps). If we compare these variance contributions to the VRFs in Table 1, we find that the amount of variance captured by the one- and two-dimensional components is a good indicator of the VRF. The Sobol' point sets that we use integrate these projections quite well because their parameters were selected by paying more attention to these particular projections. Of course, there still remains some integration error for these projections (this error may be more important when the integrand is not smooth), and the variance is also reduced (but perhaps not as much) for the other projections.

We expect the efficiency of RQMC to improve relatively to standard MC as n is increased. We estimated the VRFs of the LSM $\hat{Q}_0(x_0)$ with $n = 2^{20}$ for RQMC Seq, BB, PCA, and array-RQMC as 10, 21, 12, and 5, respectively. The VRF of BB is twice its value with $n = 2^{14}$, in Table 1, and the variance reductions provided by Seq and PCA improve slightly. The VRF of array-RQMC is basically unchanged. We also tried the same example, with $n = 2^{14}$, but with 256 equally split observation times in the year instead of 16, and the VRFs were about the same for BB and PCA, and just slightly smaller for Seq.

6.2 A Callable Bond

We consider an example of a callable bond taken from Ben Ameur et al. (2007). The bond is issued at time $t_0 = 0$, pays a coupon value c at each of the coupon dates $t_1^c < \dots < t_d^c$, and returns a principal value of 1 at the maturity date t_d^c . The bond is *callable* if it can be purchased back (called back) by the issuer before maturity. These purchasing decisions can be made by the issuer at dates $t_j = t_j^c - \Delta$ for $j = d_0, \dots, d$, for some fixed positive integer $d_0 < d$, where $\Delta > 0$ is the *notice period*. The delay t_{d_0} before the first purchase opportunity is the *protection period*. If a call back decision is made at time t_j , the owner receives a final payment of value $c + C_j$ at t_j^c , where each C_j is a predetermined *call price*.

The payment structure described above is deterministic, but the value of those payments depends on the interest rate, which evolves as a stochastic process $\{R(t), t \geq 0\}$. Here we model it as in Vasicek (1977), by the stochastic

differential equation:

$$dR(t) = \kappa(\bar{r} - R(t))dt + \sigma dB(t), \tag{10}$$

where $\{B(t), t \geq 0\}$ is a standard Brownian motion. We denote $R_j = R(t_j)$ for $j = 0, \dots, d$. The (stochastic) discount factor from t_j to t_{j-1} is $e^{-\Lambda_j}$, where

$$\Lambda_j = \int_{t_{j-1}}^{t_j} R(y)dy. \tag{11}$$

Conditional on R_{j-1} , the pair (R_j, Λ_j) has a known distribution given in Ben Ameur et al. (2007), which can be simulated easily by applying a simple transformation to a bivariate normal vector, and an explicit expression is available for $\mathbb{E}[\exp(-\Lambda_j) | R_{j-1}]$.

We give the DP formulation of this optimal stopping problem from the viewpoint of the issuer, who wants to minimize his cost. Given the initial interest rate $R_0 = r_0$, the expected discounted value of all coupons paid in the protection period is

$$V^P(r_0) = c \sum_{j=1}^{d_0-1} \mathbb{E} \left[\exp \left(- \int_{t_0}^{t_j^c} R(y)dy \right) \middle| R_0 = r_0 \right].$$

The expected discounting over the notice period at step j , conditional on R_j , is $\rho_j(r) = \mathbb{E} \left[\exp \left(- \int_{t_j}^{t_j^c} dy R(y) \right) \middle| R_j = r \right]$, for $d_0 \leq j \leq d$. Note that $V^P(r_0)$ and $\rho_j(r)$ can be computed analytically. The expected payoff at step j if the issuer decides to recall at that step, discounted to time t_j , is

$$g_j(r) = \mathbb{E} \left[\exp \left(- \int_{t_j}^{t_j^c} R(y)dy \right) (c + C_j) \middle| R_j = r \right] = \rho_j(r)(c + C_j).$$

Here we do not include the discounting from t_j to t_0 in g_j , in contrast to the general DP formulation in Section 2, to avoid keeping it in the state description (this would make the state two-dimensional), and because the optimal decision at step j does not depend on it.

Let $V_j(r)$ represent the expected value of the bond at time t_j , conditional on $R_j = r$. The DP equations can be written as

$$\begin{aligned} V_d(r) &= \rho_d(r)(c + 1), \\ Q_j(r) &= \mathbb{E} \left[e^{-\Lambda_{j+1}} V_{j+1}(R_{j+1}) \middle| R_j = r \right], \\ V_j(r) &= \min [\rho_j(r)(c + C_j), \rho_j(r)c + Q_j(r)], \\ V_0(r) &= \min [\rho_{d_0}(r)(c + C_j), \rho_{d_0}(r)c + Q_{d_0}(r)] + V^P(r). \end{aligned}$$

To apply TvR or LSM to this example, we need to estimate the (expected) continuation value, Q_j , at each purchase opportunity. For this, we need to simulate realizations of the pairs (R_j, Λ_j) for $j = d_0, \dots, d$, which requires the generation of (a vector of) $2(d - d_0 + 1)$ normal random variables, and can be done by the Seq, BB, and PCA methods described earlier. In the Seq and BB methods, each step involves the generation of a bivariate normal, which we always do via a PCA decomposition of the bivariate covariance matrix. Note that no trajectory is culled back in LSM, because g_j is always strictly positive here.

For our numerical experiment, we take the same parameters as in Ben Ameur et al. (2007). We have $d = 21$, $t_i^c = (i - 1) + 0.172$ for $i = 1, \dots, d$, $\Delta = 0.1666$, $d_0 = 11$, $c = 0.0425$, $C_{11} = 1.025$, $C_{12} = 1.02$, $C_{13} = 1.015$, $C_{14} = 1.01$, $C_{15} = 1.005$, and $C_j = 1$ for $j = 16, \dots, 21$. Thus, the first call back opportunity is at time $t_{11} = 10.0054$ and the corresponding payment (of c or $c + C_{11}$) is at time $t_{11}^c = 10.172$. The Vasicek model is parameterized with $\bar{r} = 0.098397028$, $\kappa = 0.44178462$, $\sigma = 0.13264223$, and $r_0 = 0.05$.

Table 5 reports the results of our simulation experiments with TvR and LSM. The basis functions were the simple polynomials of order up to 3, namely $\psi_k(x) = x^{k-1}$ for $k = 1, \dots, 4$. We took $n = 2^{12}$ and it was already sufficient to reach a nearly-optimal policy, in the sense that the average from the second stage evaluations did not differ significantly (statistically) from the exact value (for the given basis functions). This exact value is estimated

	Standard MC	RQMC			array-RQMC
		Seq	BB	PCA	
LSM					
$\mathbb{E}[\hat{Q}_0(r_0)]$	0.7787 ± 0.0005	0.7786 ± 0.0001	0.7787 ± 0.0001	0.7786 ± 0.0001	0.7787 ± 0.0002
variance	0.28	0.0081	0.0088	0.0094	0.032
VRF	–	34	30	30	9
$\mathbb{E}[V_{\tilde{\pi},0}(r_0)]$	0.77881 ± 0.00004	0.77874 ± 0.00004	0.77873 ± 0.00004	0.77867 ± 0.00004	0.77877 ± 0.00004
TvR					
$\mathbb{E}[\hat{Q}_0(r_0)]$	0.7790 ± 0.0005	0.77879 ± 0.00003	0.77881 ± 0.00003	0.77881 ± 0.00004	0.77879 ± 0.00002
variance	0.25	0.00090	0.0012	0.0015	0.00057
VRF	–	280	200	170	440
ex post value	0.7788 ± 0.0005	0.7786 ± 0.0001	0.7788 ± 0.0001	0.7786 ± 0.0001	0.7787 ± 0.0002
variance	0.28	0.0081	0.0095	0.0096	0.032
VRF	–	34	30	30	9
$\mathbb{E}[V_{\tilde{\pi},0}(r_0)]$	0.77876 ± 0.00004	0.77871 ± 0.00004	0.77873 ± 0.00004	0.77867 ± 0.00004	0.77874 ± 0.00004
Fixed policy					
$V_{\pi,0}(r_0)$	0.7789 ± 0.0005	0.7787 ± 0.0001	0.7788 ± 0.0001	0.7786 ± 0.0001	0.7787 ± 0.0002
variance	0.28	0.0082	0.0094	0.0097	0.032
VRF	–	34	30	30	9
Uncallable bond					
variance	0.69	0.028	0.018	0.018	0.19
VRF	–	25	40	40	4

Table 5: A callable bond evaluated over 1000 replications with $n = 2^{12}$. We estimate the bond value at 0.77871, accurate to the given digits with the given basis functions. Since this is a minimization problem, a lower out-of-sample bond value for a given policy indicates that the policy is closer to the optimal policy. The variances reported are the sample variances over the 1000 replications, multiplied by 2^{12} . The rows marked $\mathbb{E}[\hat{Q}_0(r_0)]$ give unbiased estimates of this expectation from the average over the 1000 replications of $\hat{Q}_0(r_0)$, together with a 95% confidence interval for this expectation. The rows marked $\mathbb{E}[V_{\tilde{\pi},0}(r_0)]$ give out-of-sample unbiased estimates of this expectation, in which $\tilde{\pi}$ must be interpreted as a random variable, all obtained with RQMC, PCA and $n' = 2^{14}$ for each $\tilde{\pi}$. We denote by π the selected near-optimal policy; its value is estimated with 1000 replications with $n = 2^{12}$.

to be 0.77871 (obtained from 5000 replications of RQMC with PCA, with $n' = 16$) and is identical to the value reported in Ben Ameur et al. (2007).

For the first stage LSM values, the accuracy of the estimated expectations does not permit us to detect a significant bias with respect to the exact value of 0.77871. The second stage standard MC and array-RQMC values are however biased high (less optimal policies). For TvR, the first stages values are biased high, but the bias on the second stage values of standard MC and array-RQMC is reduced compared to the second stage values of LSM. This indicates that the policies found from TvR are on average closer to the optimal policy than LSM.

The VRFs provided by RQMC are about the same for the three sampling methods, around 30, except for the first-stage values of TvR, where they are much larger. BB and PCA provide no improvement compared with Seq. For array-RQMC, the VRFs are only around 10, except again for the first-stage values of TvR, where the VRF exceeds 400. As in the previous example, the fact that the variances are about the same for a fixed policy, for LSM, and for the ex post values of TvR, indicates that the variance coming from the variation of $\tilde{\pi}$ has a small relative contribution to the total variance.

The value of the *uncallable bond*, for which the call back options have been removed, can be computed analytically (0.8559), but we also estimated its price via simulations over the same time steps t_{d_0}, \dots, t_d , for comparison. Somewhat surprisingly, we found that RQMC did not provide better variance reductions than for the callable bond. We tried to explain this by performing a functional ANOVA decomposition, summarized in Table 6.

	Callable bond with fixed π			Uncallable bond		
	Seq	BB	PCA	Seq	BB	PCA
one-dimensional components	87%	79%	90%	65%	99.4%	93%
one- and two-dimensional components	95%	95%	95%	91%	100%	99.9%

Table 6: Proportion of the variance captured by the one-dimensional ANOVA components, and by the one- and two-dimensional ANOVA components, for the callable and the uncallable bond.

From the table, for the callable bond, the one- and two-dimensional terms of the ANOVA decomposition account for 95% of the variance, for any of the three sampling methods. This corresponds very roughly to the fraction of the variance removed by the RQMC method (VRFs around 30 in Table 5), suggesting that RQMC integration with $n = 2^{14}$ removed most of the variance associated with the one- and two-dimensional terms. RQMC reduces the variance even further when the size of the integration point set is increased: with $n = 2^{20}$, the estimated VRFs of the LSM for Seq, BB and PCA sampling increase, respectively, to 96, 85 and 70 (the VRF for array-RQMC increases to 13).

On the other hand, for the uncallable bond, with BB and PCA, almost all of the variance is captured by the ANOVA components in one and two dimensions, and yet the VRFs for RQMC are about the same as for the callable bond. For BB, 98.7% of the variance is explained solely from the first coordinate, $u = \{1\}$, while for PCA, coordinates 1 and 2 (including $u = \{1, 2\}$) account for 98% of the variance. This suggests that the one- and two-dimensional ANOVA components are not very smooth functions. When the cardinality of the QMC point set is increased to $n = 2^{20}$, the VRFs of Seq, BB and PCA increase to 400, 1100 and 1100 (the VRF for array-RQMC is relatively unchanged at 5).

6.3 Other Options

We made additional experiments with an Asian American call option example, taken from Ben Ameur et al. (2002), and an American call over the maximum of 5 independent assets, taken from Broadie and Glasserman (1997) and Longstaff and Schwartz (2001). For the Asian American option, the VRFs were over 100 for RQMC with BB or PCA, whereas those factors were around 7 for the maximum call option. In both cases, the TvR method gave better out-of-sample values than LSM. Note that for the American Asian call, we added a few basis functions for TvR, to capture the non-smooth behavior, in the same spirit as in the simple put example.

ACKNOWLEDGMENTS

This work has been supported by an NSERC-Canada Grant and a Canada Research Chair to the second author. We are grateful to David Munger, who implemented the ANOVA functional decomposition.

REFERENCES

- Ben Ameur, H., M. Breton, L. Karoui, and P. L'Ecuyer. 2007. A dynamic programming approach for pricing options embedded in bonds. *Journal of Economic Dynamics and Control* 31:2212–2233.
- Ben Ameur, H., M. Breton, and P. L'Ecuyer. 2002. A dynamic programming procedure for pricing american-style asian options. *Management Science* 48:625–643.
- Bertsekas, D. P. 2005. *Dynamic programming and optimal control, volume I*. third ed. Belmont, Mass.: Athena Scientific.
- Bertsekas, D. P. 2007. *Dynamic programming and optimal control, volume II*. third ed. Belmont, Mass.: Athena Scientific.
- Bertsekas, D. P., and J. N. Tsitsiklis. 1996. *Neuro-dynamic programming*. Belmont, Mass.: Athena Scientific.
- Broadie, M., and P. Glasserman. 1997. A stochastic mesh method for pricing high-dimensional American options. Technical report, Graduate School of Business, Columbia University.
- Caffisch, R. E., W. Morokoff, and A. Owen. 1997. Valuation of mortgage-backed securities using Brownian bridges to reduce effective dimension. *The Journal of Computational Finance* 1 (1): 27–46.
- Chang, H. S., M. C. Fu, J. Hu, and S. I. Marcus. 2007. *Simulation-based algorithms for Markov decision processes*. London: Springer-Verlag.
- Chaudhary, S. K. 2005. American options and the LSM algorithm: Quasi-random sequences and Brownian bridges. *Journal of Computational Finance* 8:101–115.
- Chow, Y. S., H. Robbins, and D. Siegmund. 1971. *Great expectations: The theory of optimal stopping*. Boston: Houghton Mifflin Company.
- Clément, E., D. Lamberton, and P. Protter. 2002. An analysis of a least squares regression method for American option pricing. *Finance and Stochastics* 6:449–471.
- Glasserman, P. 2004. *Monte Carlo methods in financial engineering*. New York: Springer-Verlag.

- Haurie, A., and P. L'Ecuyer. 1986. Approximation and bounds in discrete event dynamic programming. *IEEE Trans. on Automatic Control* AC-31 (3): 227–235.
- Hickernell, F. J. 2002. Obtaining $O(N^{-2+\epsilon})$ convergence for lattice quadrature rules. In *Monte Carlo and Quasi-Monte Carlo Methods 2000*, ed. K.-T. Fang, F. J. Hickernell, and H. Niederreiter, 274–289. Berlin: Springer-Verlag.
- L'Ecuyer, P. 1989. Computing approximate solutions to Markov renewal programs with continuous state spaces. Technical Report DIUL-RT-8912, Département d'informatique, Univ. Laval, Canada. available at <http://www.iro.umontreal.ca/~lecuyer/papers.html>.
- L'Ecuyer, P. 2008. *SSJ: A Java library for stochastic simulation*. Software user's guide, available at <http://www.iro.umontreal.ca/~lecuyer>.
- L'Ecuyer, P. 2009. Quasi-Monte Carlo methods with applications in finance. *Finance and Stochastics* 13 (3): 307–349.
- L'Ecuyer, P., C. Lécot, and A. L'Archevêque-Gaudet. 2009. On array-RQMC for Markov chains: Mapping alternatives and convergence rates. In *Monte Carlo and Quasi-Monte Carlo Methods 2008*, ed. P. L'Ecuyer and A. B. Owen, 485–500. Berlin: Springer-Verlag.
- L'Ecuyer, P., C. Lécot, and B. Tuffin. 2008. A randomized quasi-Monte Carlo simulation method for Markov chains. *Operations Research* 56 (4): 958–975.
- L'Ecuyer, P., J.-S. Parent-Chartier, and M. Dion. 2008. Simulation of a Lévy process by PCA sampling to reduce the effective dimension. In *Proceedings of the 2008 Winter Simulation Conference*, 436–443: IEEE Press.
- Lemieux, C. 2004. Randomized quasi-Monte Carlo: A tool for improving the efficiency of simulations in finance. In *Proceedings of the 2004 Winter Simulation Conference*, 1565–1573. Piscataway, NJ: IEEE Press.
- Lemieux, C. 2009. *Monte Carlo and quasi-Monte Carlo sampling*. New York, NY: Springer-Verlag.
- Lemieux, C., M. Cieslak, and K. Luttmner. 2004. *RandQMC user's guide: A package for randomized quasi-Monte Carlo methods in C*. Software user's guide, available at <http://www.math.uwaterloo.ca/~clemieux/randqmc.html>.
- Lemieux, C., and J. La. 2005. A study of variance reduction techniques for American option pricing. In *Proceedings of the 2005 Winter Simulation Conference*, ed. M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, 1884–1891: IEEE Press.
- Lin, J., and X. Wang. 2008. New Brownian bridge construction in quasi-Monte Carlo methods for computational finance. *Journal of Complexity* 24:109–133.
- Longstaff, F. A., and E. S. Schwartz. 2001. Valuing american options by simulation: A simple least-squares approach. *The Review of Financial Studies* 14 (1): 113–147.
- Moskowitz, B., and R. E. Caflisch. 1996. Smoothness and dimension reduction in quasi-Monte Carlo methods. *Journal of Mathematical and Computer Modeling* 23:37–54.
- Niederreiter, H. 1992. *Random number generation and quasi-Monte Carlo methods*, Volume 63 of *SIAM CBMS-NSF Regional Conference Series in Applied Mathematics*. Philadelphia, PA: SIAM.
- Owen, A. B. 1998. Latin supercube sampling for very high-dimensional simulations. *ACM Transactions on Modeling and Computer Simulation* 8 (1): 71–102.
- Owen, A. B. 2003. Variance with alternative scramblings of digital nets. *ACM Transactions on Modeling and Computer Simulation* 13 (4): 363–378.
- Sloan, I. H., and S. Joe. 1994. *Lattice methods for multiple integration*. Oxford: Clarendon Press.
- Sobol', I. M. 1967. The distribution of points in a cube and the approximate evaluation of integrals. *U.S.S.R. Comput. Math. and Math. Phys.* 7:86–112.
- Sobol', I. M., and E. E. Myshetskaya. 2007. Monte Carlo estimators for small sensitivity indices. *Monte Carlo Methods and Applications* 13 (5–6): 455–465.
- Stentoft, L. 2004a. Assessing the Least Squares Monte-Carlo approach to American option valuation. *Review of Derivatives Research* 7:129–168.
- Stentoft, L. 2004b. Convergence of the Least Squares Monte Carlo approach to American option valuation. *Management Science* 50 (9): 1193–1203.
- Tsitsiklis, J. N., and B. V. Roy. 2001. Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks* 12 (4): 694–703.
- Vasicek, O. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5:177–188.
- Wang, X. 2007. Constructing robust good lattice rules for computational finance. *SIAM Journal on Scientific Computing* 29 (2): 598–621.
- Wang, X., and I. H. Sloan. 2009. Quasi-Monte Carlo methods in financial engineering: An equivalence principle and dimension reduction. see www.maths.unsw.edu.au/applied/files/2008/amr08_10.pdf.

Zanger, D. Z. 2009. Convergence of a least-squares Monte Carlo algorithm for bounded approximating sets. *Applied Mathematical Finance* 16 (2): 123–150.

AUTHOR BIOGRAPHIES

MAXIME DION has a Ph.D. in physics from Rutgers University and a M.Sc. in computational finance from Université de Montréal. He is currently a postdoctoral fellow at the Université de Montréal. He is interested in stochastic optimization. His email address is dionmaxi@iro.umontreal.ca.

PIERRE L'ECUYER is Professor in the Département d'Informatique et de Recherche Opérationnelle, at the Université de Montréal, Canada. He holds the Canada Research Chair in Stochastic Simulation and Optimization. He is a member of the CIRRELT and GERAD research centers. His main research interests are random number generation, quasi-Monte Carlo methods, efficiency improvement via variance reduction, sensitivity analysis and optimization of discrete-event stochastic systems, and discrete-event simulation in general. He is currently Editor-in-Chief for *ACM Transactions on Modeling and Computer Simulation*, and Associate/Area Editor for *ACM Transactions on Mathematical Software, Statistics and Computing, International Transactions in Operational Research, Cryptography and Communications*. He obtained the *E. W. R. Steacie* fellowship in 1995-97, a *Killam* fellowship in 2001-03, and became an INFORMS Fellow in 2006. More information and his recent research articles are available on-line from his web page: <http://www.iro.umontreal.ca/~lecuyer>.