Version: May 17, 2000

A Numerical Procedure for Pricing American-style Asian Options

Hatem Ben Ameur
École des Sciences de Gestion, Université du Québec à Montréal, Montréal, Canada
E-mail: Ben_Ameur.Hatem@uqam.ca

Michèle Breton
GERAD and Département des Méthodes Quantitatives de Gestion, École des H.E.C., Montréal, Canada
E-mail: Michele.Breton@hec.ca.

Pierre L'Ecuyer
GERAD and Département d'Informatique et de Recherche Opérationnelle
Université de Montréal, C.P. 6128, Succ. Centre-Ville, Montréal, H3C 3J7, Canada
URL: http://www.iro.umontreal.ca/~lecuyer

Abstract. Pricing Asian options based on the arithmetic average, under the Black and Scholes model, involves estimating an integral (a mathematical expectation) for which no analytical solution is available. Pricing their American-style counterparts, which provide early exercise opportunities, poses the additional difficulty of solving a dynamic optimization problem to determine the optimal exercise strategy. We develop a numerical method for pricing American-style Asian options based on dynamic programming combined with finite-element piecewise-polynomial approximation of the value function. Numerical experiments show convergence, consistency, and efficiency. Some theoretical properties of the value function and of the optimal exercise strategy are also established.
(*Option pricing, Asian Options, Path-dependent options, American Options, Dynamic Programming, Piecewise Polynomials*)

---

1 Introduction

A financial *derivative* is a contract which provides its holder a future payment that depends on the price of one or more primitive asset(s), such as stocks or currencies. In a frictionless market, the no-arbitrage principle allows one to express the value of a derivative as the mathematical expectation of its discounted future payment, with respect to a so-called risk-neutral probability measure. *Options* are particular derivatives characterized by non-negative payoffs. *European-style* options can be exercised at the expiration date only, whereas *American-style* ones offer early exercise opportunities to the holder.

For simple cases, such as for European call and put options written on a stock whose price is modeled as a geometric Brownian motion (GBM), as studied by Black and Scholes (1973), analytic formulas are available for the fair price of the option. For more complicated derivatives, however, which may involve multiple assets, complex payoff functions, possibilities of early exercise, stochastic time-varying model parameters, etc., analytic formulas are unavailable. These derivatives are usually priced either via Monte Carlo simulation or via numerical methods. (e.g., Boyle, Broadie, and Glasserman 1997, Hull 2000, Wilmott, Dewynne, and Howison 1993, and other references given there).

An important class of options for which no analytic formula is available even under the standard Black-Scholes GBM model is the class of *Asian options*, for which the payoff is a function of the arithmetic average of the price of a primitive asset over a certain time period. These options are often used for protection against brutal and unexpected changes of prices. An Asian option can hedge the risk exposure of a firm that sells or buys certain types of resources (raw materials, energy, foreign currency, etc.), on a regular basis over some period of time. The average being in general less volatile than the underlying asset price itself, these contracts are less expensive than their standard versions. Asian options are heavily traded over-the-counter and, because of the possible lack of depth of these markets, their theoretical values often need to be computed *on-the-fly* for fair negotiations.

Asian options come in various flavors. For example, the average can be arithmetic or it can be geometric. One talks of a *plain vanilla* Asian option if the average is computed over the full trading period, and a *backward-starting* option if it is computed over a right subinterval of the trading period. This interval usually has a fixed starting point in time. The Asian option can be *fixed-strike* (if the strike price is a fixed constant) or *floating-strike* (if the strike is itself an average). It is called *flexible* when the payoff is a weighted average, and *equally weighted* when all the weights are equal. The prices are *discretely sampled* if the payoff is the average of a discrete set of values of the underlying asset (observed at discrete epochs), and *continuously sampled* if the payoff is the integral of the asset value over some time interval, divided by the length of that interval. The options considered in this paper are the most common: *Fixed-strike, equally-weighted, discretely-sampled Asian options based on arithmetic averaging*. Our method could also be adapted to price other kinds of discretely-sampled Asian options.

European-style Asian (*Eurasian*) options can be exercised at the expiration date only, whereas *American-style* ones (named *Amerasian*) offer earlier exercise opportunities, which may become attractive intuitively when the current asset price is below the current running average (i.e., is pulling down the average) for a call option, and when it is above the running average for a put. Here, *we focus on Amerasian call options*, whose values are harder to compute than the Eurasian ones, because an optimization problem must be solved at the same time as computing the mathematical expectation giving the option's value.

There is an extensive literature on the pricing of Eurasian options. In the context of the GBM model, there is a closed-form analytic solution for the value of discretely-sampled Eurasian options only when they are based on the *geometric average* (Turnbull and Wakeman 1991, Zhang 1995). The idea is that under the GBM model, the asset price at any given time has the lognormal distribution, and the geometric average of lognormals is a lognormal. Geman and Yor (1993) used Bessel processes and derived exact formulas for the Laplace transform of the value of a continuous-time Eurasian option. For options based on the arithmetic average, solution approaches include *quasi-analytic* approximation methods based on Fourier transforms, Edgeworth and Taylor expansions, and the like (e.g., Bouaziz, Briys, and Crouhy 1994, Carverhill and Clewlow 1990, Curran 1994, Levy 1992, Ritchken, Sankarasubramanian, and Vijh 1993, Turnbull and Wakeman 1991), methods based on partial differential equations (PDEs) and their numerical solution via *finite-difference* techniques (e.g., Alziary, Décamps, and Koehl 1997, Rogers and Shi 1995, Zvan, Forsyth, and Vetzal 1998), and *Monte Carlo simulation* coupled with variance-reduction techniques (e.g., Glasserman, Heidelberger, and Shahabuddin 1999, Kemna and Vorst 1990, L'Ecuyer and Lemieux 2000, Lemieux and L'Ecuyer 1998, Lemieux and L'Ecuyer 2000).

Techniques for pricing Amerasian options are surveyed by Barraquand and Pudet (1996), Grant, Vora, and Weeks (1997), and Zvan, Forsyth, and Vetzal (1998, 1999). Hull and White (1993) have adapted binomial lattices (from the binomial tree model of Cox, Ross, and Rubinstein 1979) to the pricing of Amerasian options. But these methods remain limited in their application and do not give a clear insight on the optimal exercising region. Broadie and Glasserman (1997a) proposed a simulation method based on nonrecombining trees in the lattice model, and which produces two estimators of the option value, one with positive bias and one with negative bias. By taking the union of the confidence intervals corresponding to these two estimators, one obtains a conservative confidence interval for the true value. However, the work and space requirements of their approach increases exponentially with the number of exercise opportunities. Broadie and Glasserman (1997b) then developed a simulation-based stochastic mesh method that accommodates a large number of exercise dates and high-dimensional American options. Their method appears adaptable to Amerasian options, although this is not the route we take here.

Zvan, Forsyth, and Vetzal (1998) have developed stable numerical PDE methods techniques, adapted from the field of computational fluid dynamics, for pricing Amerasian options with continuously sampled prices. Zvan, Forsyth, and Vetzal (1999) have also adapted these PDE methods to options with discretely sampled asset prices, and with barriers. The numerical approach introduced in this paper is formulated in discrete time directly.

Pricing American-style options is naturally formulated as a Markov Decision process, i.e., a stochastic dynamic programming (DP) problem, as pointed out by Barraquand and Martineau (1995) and Broadie and Glasserman (1997b), for example. The DP *value function* expresses the value of an Amerasian option as a function of the current time, current price, and current average. This value function satisfies a DP recurrence (or Bellman equation), written as an integral equation. Solving this equation yields both the option value and the optimal exercise strategy. For a general overview of stochastic DP, we refer the reader to Bertsekas (1987).

In this paper, we write the DP equation for Amerasian options under the GBM model. Using this equation, we prove by induction certain properties of the value function and of the optimal *exercise frontier* (which delimits the region where it is optimal to exercise the option). We then propose a numerical solution approach for the DP equation, based on piecewise bilinear interpolation over rectangular finite elements. This kind of approach has been used in other application contexts, e.g. by Haurie and L'Ecuyer (1986), L'Ecuyer and Malenfant (1988). In fact, we reformulate the DP equation in a way that simplifies significantly the numerical integration at each step. This is a key ingredient for improving the efficiency of the procedure. Convergence and consistency of the method, as the discretization gets finer, follow from the monotonicity properties of the value function. Numerical experiments indicate that the method is competitive and efficient; it provides precise results in a reasonable computing time. It could also be easily adapted to price most low-dimensional American-style products such as calls with dividends, puts, lookback options, and options with barriers. The general methodology would also work for other types of models than the GBM for the evolution of the underlying asset, e.g., for a CEV process (Cox 1996, Boyle and Tian 1999). The properties of the value function that we derive in Section 4 are easy to generalize. On the other hand, the implementation details that we develop in Section 5 are specific to the GBM model. These details would have to be reworked for other types of models.

The idea of this paper came after reading Grant, Vora, and Weeks (1997). These authors also formulate the problem of pricing an Amerasian option in the dynamic programming framework, but use Monte Carlo simulation to estimate the value function at each point of some discretization of the state space, and identify a "good" exercise frontier by interpolation. Their estimate of the value function at the initial date is an estimate of the option value. These authors also propose to restrict the strategy of exercise to a class of suboptimal rules where the exercise frontier is approximated by two linear segments, at each date of exercise opportunity. They observed on a few numerical examples that restricting the class of strategies in this way did not seem to diminish the value of the option significantly, but they provided no proof that this is true in general.

Here, we suggest replacing simulation at both stages by numerical integration, which is obviously less noisy, and we do not assume a priori a shape of the exercise frontier. For both the simulation approach and our approach, an approximation of the value function must be memorized, so the storage requirement is essentially the same for the two methods.

While finalizing the revision of this paper, we learned about related work by Wu and Fu (2000), who also proved some properties of the exercise frontier for Amerasian options and proposed a different approach, that parameterizes the exercise frontier and optimizes the parameters by a stochastic approximation algorithm combined with a simulation-based perturbation analysis gradient estimation method.

The rest of the paper is organized as follows. Section 2 presents the model and notation. In Section 3, we develop the DP formulation. In Section 4, we establish certain properties of the value function and of the optimal region of exercise. Our numerical approximation approach is detailed in Section 5. In Section 6, we report on numerical experiments. Section 7 is a conclusion.

2 Model and notation

2.1 Evolution of the Primitive Asset

We assume a single primitive asset whose price process $\{S(t),\, t \in [0,T]\}$ is a GBM, in a world that satisfies the standard assumptions of Black and Scholes (1973). Under these assumptions (see, e.g., Karatzas and Shreve 1998), there is a probability measure Q called *risk-neutral*, under which the primitive asset price $S(\cdot)$ satisfies the stochastic differential equation

$$dS(t) = rS(t)dt + \sigma S(t)dW(t), \quad \text{for } 0 \le t \le T, \tag{1}$$

where $S(0) > 0$, r is the risk-free rate, σ is the volatility parameter, T is the maturity date, and $\{W(t),\, t \in [0,T]\}$ is a standard Brownian motion. The solution of (1) is given by

$$S(t'') = S(t')e^{\mu(t''-t')+\sigma[W(t'')-W(t')]}, \quad \text{for } 0 \le t' \le t'' \le T, \tag{2}$$

where $\mu = r - \sigma^2/2$. An important feature is that the random variable $S(t'')/S(t')$ is lognormal with parameters $\mu(t'' - t')$ and $\sigma\sqrt{t'' - t'}$, and independent of the σ-field $\mathcal{F}(t') = \sigma\{S(t),\, t \in [0,t']\}$, i.e., of the trajectory of $S(t)$ up to time t'. This follows from the independent-increments property of the Brownian motion. In addition, from the no-arbitrage property of the Black-Scholes model, the discounted price of the primitive asset is a Q-martingale:

$$\rho(t')S(t') = E\left[\rho(t'')S(t'') \mid \mathcal{F}(t')\right], \quad \text{for } 0 \le t' \le t'' \le T, \tag{3}$$

where $\{\rho(t) = e^{-rt},\, t \in [0,T]\}$ is the discount factor process and E is (all along this paper) the expectation with respect to Q. Details about risk-neutral evaluation can be found in Karatzas and Shreve (1998).

2.2 The Amerasian Contract

We consider a model similar to that of Grant, Vora, and Weeks (1997). Let $0 = t_0 \le t_1 < t_2 < \cdots < t_n = T$ be a fixed sequence of *observation dates*, where T is the *time horizon*, and let m^* be an integer satisfying $1 \leq m^* \leq n$. The *exercise opportunities* are at dates t_m, for $m^* \leq m \leq n$. If the option is exercised at time t_m, the *payoff* of the Amerasian call option is $(\overline{S}_m - K)^+ \stackrel{\text{def}}{=} \max(0, \overline{S}_m - K)$, where $\overline{S}_m = (S(t_1) + \cdots + S(t_m))/m$ is the arithmetic average of the asset prices at the observation dates up to time t_m. This model is quite flexible. For $n = 1$, we get a standard European call option. For $m^* = n > 1$, we have an Eurasian option. Here, we are not really interested in these degenerate cases, but in the case where $m^* < n$. To simplify the exposition we will assume that the observation dates are equally spaced: $t_i - t_{i-1} = h$ for $i = 1, \ldots, n$, for some constant h.

3 The Dynamic Programming Formulation

3.1 Value Function and Recurrence Equations

For $m = 0, \ldots, n$, denote by $v_m(s, \overline{s})$ the value of the option at the observation date t_m when $S(t_m) = s$ and $\overline{S}_m = \overline{s}$, assuming that the decisions of exercising the option or not, from time t_m onwards, are always made in an optimal way (i.e., in a way that maximizes the option value). This optimal value is a function of the *state variables* $(s, \overline{s})$ and of the time t_m. We take the *state space* as $[0, \infty)^2$ for convenience, although at each time step, only a subset of this space is reachable: Because $S(\cdot)$ is always positive, at time t_m one must have $\overline{s} = s > 0$ if $m = 1$ and $\overline{s} > s/m > 0$ if $m > 1$. At time t_n, $v_n(s, \overline{s}) \equiv v_n(\overline{s})$ does not depend on s, whereas at time t_0, $\overline{s}$ is just a dummy variable in $v_0(s) \equiv v_0(s, \overline{s})$, which depends only on s.

At time t_m, if $S(t_m) = s$ and $\overline{S}_m = \overline{s}$, the *exercise value* of the option (for $m \geq m^*$) is

$$v_m^e(\overline{s}) = (\overline{s} - K)^+, \tag{4}$$

whereas the *holding value* (i.e., the value of the option if it is not exercised at time t_m and if we follow an optimal exercise strategy thereafter) is

$$v_m^h(s, \overline{s}) = \Big\{ \; \rho\, E_{m,s,\overline{s}}\left[v_{m+1}(S(t_{m+1}), (m\overline{s} + S(t_{m+1}))/(m+1))\right] \quad \text{for } 0 \leq m \leq n-1, \tag{5}$$

where $E_{m,s,\overline{s}}[\cdot]$ represents the conditional expectation $E[\cdot \mid \mathcal{F}(t_m),\ S(t_m) = s,\ \overline{S}_m = \overline{s}]$, and $\rho = e^{-rh}$ is the discount factor over the period $[t_m, t_{m+1}]$. This holding value $v_m^h(s, \overline{s})$ is the (conditional) expected value of the option at time t_{m+1}, discounted to time t_m.

The optimal value function satisfies:

$$v_m(s, \overline{s}) = \begin{cases} v_m^h(s, \overline{s}) & \text{if } 0 \le m \le m^* - 1, \\ \max\{v_m^e(\overline{s}),\ v_m^h(s, \overline{s})\} & \text{if } m^* \le m \le n - 1, \\ v_m^e(\overline{s}) & \text{if } m = n. \end{cases} \tag{6}$$

The optimal exercise strategy is defined as follows: In state $(s, \overline{s})$ at time t_m, for $m^* \le m < n$, exercise the option if $v_m^e(\overline{s}) > v_m^h(s, \overline{s})$, and hold it otherwise. The value of the Amerasian option at the initial date t_0, under an optimal exercise strategy, is $v_0(s) = v_0(s, \overline{s})$. The functions v_m^e, v_m^h, and v_m are defined over the entire state space $[0, \infty)^2$ for all m, via the above recurrence equations, even if we know that part of the state space is not reachable. (We do this to simplify the notation and to avoid considering all sorts of special cases.)

The natural way of solving (6) is via backward iteration: From the known function v_n and using (4)–(6), compute the function v_{n-1}, then from v_{n-1} compute v_{n-2}, and so on, down to v_0. Here, unfortunately, the functions v_m for $m \le n-2$ cannot be obtained in closed form (we will give a closed-form expression for v_{n-1} in a moment), so they must be approximated in some way. We propose an approximation method in Section 5. In the next section, we establish some properties of v_m and of the optimal strategy of exercise, which are interesting per se and are also useful for analyzing the numerical approximation techniques.

4 Characterizing the Value Function and the Optimal Exercise Strategy

4.1 The Value Function v_{n-1}

Recall that the value function v_n at the horizon $T = t_n$ has the simple form $v_n(s, \overline{s}) = (\overline{s} - K)^+$. We now derive a closed-form expression for the value function at time t_{n-1}, the last observation date before the horizon. We assume that $1 \le m^* \le n-1$ (otherwise one has $v_{n-1} = v_{n-1}^h$ and the argument simplifies). From (5) we have

$$v_{n-1}^h(s,\overline{s}) \;=\; \rho\, E_{n-1,s,\overline{s}}\left[\left(\frac{(n-1)\overline{s}+S(t_n)}{n} - K\right)^+\right] \;=\; \frac{\rho}{n} E_{n-1,s,\overline{s}}\left[\left(S(t_n) - \overline{K}\right)^+\right],$$

where $\overline{K} = nK - (n-1)\overline{s}$.

We first consider the case where $\overline{K} \le 0$, i.e., $\overline{s} \ge Kn/(n-1)$. The holding value can then be derived from (3) as the linear function

$$v_{n-1}^h(s,\overline{s}) \;=\; v^{\text{lin}}(s,\overline{s}) \;=\; \frac{s}{n} - \frac{\rho}{n}\overline{K} \;=\; \frac{s}{n} + \rho\frac{n-1}{n}\overline{s} - \rho K,$$

and the exercise value equals $\overline{s}-K > 0$. One can easily identify the optimal decision (exercise or not) for any given state $(s,\overline{s})$ by comparing this exercise value with the holding value v^{lin}. This yields an explicit expression for the value function. Consider the line defined in the $(s,\overline{s})$ plane by $\overline{s} - K = v^{\text{lin}}(s,\overline{s})$, that is,

$$s - (n-(n-1)\,\rho)\,\overline{s} + nK(1-\rho) = 0. \tag{7}$$

The optimal strategy here is: Exercise the option if and only if $(s,\overline{s})$ is above the line (7). This line passes through the point $(K,K)n/(n-1)$ and has a slope of $1/(n-(n-1)\rho) < 1$, so it is optimal to exercise for certain pairs $(s,\overline{s})$ with $s > \overline{s}$, a possibility which was neglected by Grant, Vora, and Weeks (1997). A partial intuition behind this optimal strategy is that for sufficiently large $\overline{s}$ and for $s < \overline{s}$, the average price will most likely decrease in the future (it is pressured down by the current value), so it is best to exercise right away.

We now consider the case $\overline{K} > 0$, i.e., $\overline{s} < Kn/(n-1)$. In this case, the holding value is equivalent to the value of an European call option under the GBM model, with strike price $\overline{K}$, initial price s for the primitive asset, maturity horizon $T - t_{n-1} = h$, volatility σ, and risk-free rate r. This value is given by the well-known Black-Scholes formula:

$$v^{\text{BS}}(s,\overline{s}) \;=\; \frac{1}{n}\left(\Phi(d_1)s - \rho\overline{K}\Phi(d_1 - \sigma\sqrt{h})\right)$$

where

$$d_1 = \frac{\ln(s/\overline{K}) + (r + \sigma^2/2)h}{\sigma\sqrt{h}}$$

and Φ denotes the standard normal distribution function. If $\bar{s} \leq K$, one must clearly hold the option because the exercise value is 0. For $\bar{s} > K$, the optimal decision is obtained by comparing $v^{\mathrm{BS}}(s, \bar{s})$ with $\bar{s} - K$, similar to what we did for the case where $\overline{K} \leq 0$. We now have a closed-form expression for v_{n-1}:

$$v_{n-1}(s, \bar{s}) = \begin{cases} \max\{\bar{s} - K,\ [s + (n-1)\rho\bar{s}]/n - \rho K\} & \text{if } \ \bar{s} \geq Kn/(n-1), \\ \max\{\bar{s} - K,\ v^{\mathrm{BS}}(s, \bar{s})\} & \text{if } \ \bar{s} < Kn/(n-1). \end{cases}$$

We could (in principle) compute an expression for v_{n-2} by placing our expression for v_{n-1} in the DP equations (5) and (6), although this becomes quite complicated. The functions v_n and v_{n-1} are obviously continuous, but are not differentiable (v_n is not differentiable with respect to $\bar{s}$ at $\bar{s} = K$). These functions are also monotone non-decreasing with respect to both s and $\bar{s}$. Finally, the optimal exercise region at t_{n-1} is the epigraph of some function φ_{n-1}, i.e., the region where $\bar{s} > \varphi_{n-1}(s)$, where $\varphi_{n-1}(s)$ is defined as the value of $\bar{s}$ such that $v^h_{n-1}(s, \bar{s}) = \bar{s} - K$. In the next subsection, we show that these general properties hold as well at time t_m for $m \leq n - 1$.

4.2 General Properties of the Value Function and of the Exercise Frontier

We now prove certain monotonicity and convexity properties of the value function at each step, and use these properties to show that the optimal strategy of exercise at each step is characterized by a function φ_m whose graph partitions the state space in two pieces: At time t_m, for $m^* \leq m < n$, if $\bar{s} \geq \varphi_m(s)$ it is optimal to exercise the option now, whereas if $\bar{s} \leq \varphi_m(s)$ it is optimal to hold it for at least another step. We derive these properties under the GBM model, but the proofs work as well if $S(t_{m+1})/S(t_m)$ has a different distribution than the lognormal, provided that it is independent of $\mathcal{F}(t_m)$.

Proposition 1. *At each observation date t_m, for $1 \leq m < n$, the holding value $v_m^h(s, \overline{s})$ is a continuous, strictly positive, strictly increasing, and convex function of both s and $\overline{s}$, for $s > 0$ and $\overline{s} > 0$. The function $v_0(s)$ enjoys the same properties as a function of s, for $s > 0$. For $1 \leq m < n$, the value function $v_m(s, \overline{s})$ also has these properties except that it is only non-decreasing (instead of strictly increasing) in s.*

Proof. The proof proceeds by backward induction on m. At each step, we define the auxiliary random variable $\tau_{m+1} = S(t_{m+1})/S(t_m)$, which has the lognormal distribution with parameters μh and $\sigma\sqrt{h}$, independently of $\mathcal{F}(t_m)$, as in (2). A key step in our proof will be to write the holding value $v_m^h(s, \overline{s})$ as a convex combination of a continuous family of well-behaved functions indexed by τ_{m+1}.

For $m = n - 1$, the holding value is

$$v_{n-1}^h(s, \overline{s}) \;=\; \rho E_{n-1,s,\overline{s}}\left[v_n\left(\overline{S}_n\right)\right] \;=\; \rho \int_0^\infty \left(\frac{(n-1)\,\overline{s} + s\tau}{n} - K\right)^+ f(\tau)d\tau,$$

where f is the density of τ_n. The integrand is continuous and bounded by an integrable function of τ over any bounded interval for s and $\overline{s}$. Therefore the holding value v_{n-1}^h is also continuous by Lebesgue's dominated convergence theorem (Billingsley 1986). The integral is strictly positive because, for instance, the lognormal distribution always gives a strictly positive measure to the event $\{(n-1)\overline{s} + s\tau_n - nK \geq n\}$, on which the integrand is ≥ 1.

To show that $v_{n-1}^h(s, \overline{s})$ is strictly increasing in $\overline{s}$, let $\overline{s} > 0$ and $\delta > 0$. One has

$$\begin{aligned}
& v_{n-1}^h(s, \overline{s} + \delta) - v_{n-1}^h(s, \overline{s}) \\
= \;& \rho \int_0^\infty \left[\left(\frac{(n-1)\left(\overline{s} + \delta\right) + s\tau}{n} - K\right)^+ - \left(\frac{(n-1)\,\overline{s} + s\tau}{n} - K\right)^+\right] f(\tau)d\tau \\
\geq \;& \rho \int_{(nK-(n-1)\overline{s})/s}^\infty \left[\frac{(n-1)\left(\overline{s} + \delta\right) + s\tau}{n} - \frac{(n-1)\,\overline{s} + s\tau}{n}\right] f(\tau)d\tau \\
\geq \;& \frac{n-1}{n}\,\delta \;>\; 0.
\end{aligned}$$

The same argument can be used to prove that $v_{n-1}^h(s, \overline{s})$ is strictly increasing in s. The convexity of $v_{n-1}^h(s, \overline{s})$ follows from the fact that this function is a positively weighted average (a convex combination), over all positive values of τ, of the values of $((n-1)\overline{s}+s\tau)/n-K)^+$, which are (piecewise linear) convex functions of s and $\overline{s}$ for each τ. It is also straightforward to verifying directly the definition of convexity for v^h_{n-1} (e.g., Rockafellar 1970): For any pair $(s_1, \overline{s}_1)$ and $(s_2, \overline{s}_2)$, and any $\lambda \in (0,1)$, $v^h_{n-1}(\lambda s_1 + (1-\lambda)s_2, \lambda\overline{s}_1 + (1-\lambda)\overline{s}_2) \leq \lambda v^h_{n-1}(s_1, \overline{s}_1) + (1-\lambda)v^h_{n-1}(s_2, \overline{s}_2)$.

Because the holding function is continuous and strictly positive, the value function

$$v_{n-1}(s, \overline{s}) = \max\left((\overline{s}-K)^+,\, v^h_{n-1}(s, \overline{s})\right)$$

is also continuous and strictly positive. It is also convex, non-decreasing in s, and strictly increasing in $\overline{s}$, because it is the maximum of two functions that satisfy these properties. (The maximum can be reached at $(\overline{s}-K)^+$ only if $\overline{s} > K$, since $v^h_{n-1}(s, \overline{s}) > 0$.)

We now assume that the result holds for $m+1$, where $1 \leq m \leq n-2$, and show that this implies that it holds for m. The holding value at t_m is

$$\begin{aligned} v^h_m(s, \overline{s}) &= \rho\, E_{m,s,\overline{s}}\left[v_{m+1}(s\tau_{m+1}, (m\overline{s}+s\tau_{m+1})/(m+1))\right] \\ &= \rho \int_0^\infty v_{m+1}(s\tau, (m\overline{s}+s\tau)/(m+1)) f(\tau) d\tau. \end{aligned} \tag{8}$$

where f is the lognormal density of τ_m. The function $v^h_m(s, \overline{s})$ is also continuous and strictly positive because the integrand is continuous, strictly positive, and bounded by an integrable function of τ over every bounded interval for s and $\overline{s}$. The other properties can be proved via similar arguments as for the case of $m = n-1$. The proof for v_0 is also similar as for $m > 0$. We omit the details. ∎

LEMMA 2. *For $s > 0$ and $0 < \overline{s}_1 < \overline{s}_2$, one has*

$$v^h_m(s, \overline{s}_2) - v^h_m(s, \overline{s}_1) < \frac{m}{m+1}(\overline{s}_2 - \overline{s}_1)\rho \qquad \text{for } 1 \leq m < n \tag{9}$$

and

$$v_m(s, \overline{s}_2) - v_m(s, \overline{s}_1) \leq \overline{s}_2 - \overline{s}_1 \qquad \text{for } 1 \leq m \leq n. \tag{10}$$

PROOF. The proof proceeds again by backward induction on m. We will use the property that $b^+ - a^+ \leq b - a$ when $a \leq b$. For $m = n$, we have $v_n(s, \overline{s}_2) - v_n(s, \overline{s}_1) = (\overline{s}_2 - K)^+ - (\overline{s}_1 - K)^+ \leq \overline{s}_2 - \overline{s}_1$, so (10) holds for $m = n$. We now assume that (10) holds for $m+1$, where $m < n$, and show that this implies (9) and (10) for m. From (8), we have

$$\begin{aligned}
& v_m^h(s, \overline{s}_2) - v_m^h(s, \overline{s}_1) \\
= \; & \rho \int_0^\infty \left(v_{m+1}\left(s\tau, \frac{m\overline{s}_2 + s\tau}{m+1}\right) - v_{m+1}\left(s\tau, \frac{m\overline{s}_1 + s\tau}{m+1}\right)\right) f(\tau) d\tau \\
\leq \; & \rho \int_0^\infty \left(\frac{m\overline{s}_2 + s\tau}{m+1} - \frac{m\overline{s}_1 + s\tau}{m+1} \right) f(\tau) d\tau \\
\leq \; & \frac{m}{m+1}(\overline{s}_2 - \overline{s}_1)\rho.
\end{aligned}$$

Moreover, $v_m^e(\overline{s}_2) - v_m^e(\overline{s}_1) = (\overline{s}_2 - K)^+ - (\overline{s}_1 - K)^+ \leq \overline{s}_2 - \overline{s}_1$. Now,

$$\begin{aligned}
v_m(s, \overline{s}_2) - v_m(s, \overline{s}_1) &= \max(v_m^e(\overline{s}_2),\, v_m^h(s, \overline{s}_2)) - \max(v_m^e(\overline{s}_1),\, v_m^h(s, \overline{s}_1)) \\
&\leq \max(v_m^e(\overline{s}_2) - v_m^e(\overline{s}_1),\, v_m^h(s, \overline{s}_2) - v_m^h(s, \overline{s}_1)) \\
&\leq \overline{s}_2 - \overline{s}_1.
\end{aligned}$$

This completes the proof. ∎

PROPOSITION 3. *For $m = m^*, \ldots, n-1$, there exists a continuous, strictly increasing, and convex function $\varphi_m : (0, \infty) \to (K, \infty)$ such that*

$$v_m^h(s, \overline{s}) \begin{cases} > v_m^e(\overline{s}) & \textit{for } \overline{s} < \varphi_m(s) \\ = v_m^e(\overline{s}) & \textit{for } \overline{s} = \varphi_m(s) \\ < v_m^e(\overline{s}) & \textit{for } \overline{s} > \varphi_m(s). \end{cases} \tag{11}$$

PROOF. Let $s > 0$ and $m^* \leq m \leq n-1$. We know from Proposition 1 and Lemma 2 that $v_m^h(s, \overline{s})$ is always strictly positive and strictly increasing in $\overline{s}$, with a growth rate always strictly less than $\rho m/(m+1) < 1$. On the other hand, $v_m^e(\overline{s}) = (\overline{s} - K)^+$ is 0 for $\overline{s} \leq K$ and increases at rate 1 for $\overline{s} > K$. Therefore, there is a unique value of $\overline{s} > K$, denoted $\varphi_m(s)$, such that (11) is satisfied.

To show that $\varphi_m(s)$ is strictly increasing in s, let $0 < s_1 < s_2$. If we suppose that $\varphi_m(s_1) \geq \varphi_m(s_2)$, we obtain the contradiction

$$\begin{aligned} 0 &< v_m^h(s_2, \varphi_m(s_2)) - v_m^h(s_1, \varphi_m(s_2)) \\ &= v_m^h(s_2, \varphi_m(s_2)) - v_m^h(s_1, \varphi_m(s_1)) + v_m^h(s_1, \varphi_m(s_1)) - v_m^h(s_1, \varphi_m(s_2)) \\ &= \varphi_m(s_2) - K - [\varphi_m(s_1) - K] + v_m^h(s_1, \varphi_m(s_1)) - v_m^h(s_1, \varphi_m(s_2)) \\ &= v_m^h(s_1, \varphi_m(s_1)) - v_m^h(s_1, \varphi_m(s_2)) - [\varphi_m(s_1) - \varphi_m(s_2)] \;\leq\; 0 \end{aligned}$$

where the first inequality is by Proposition 1 and the last inequality is by (9) in Lemma 2. Therefore, $\varphi_m(s_1) < \varphi_m(s_2)$, i.e., $\varphi_m(s)$ is strictly increasing in s.

Now consider two states $(s_1, \overline{s}_1)$ and $(s_2, \overline{s}_2)$ where it is optimal to exercise the option at time t_m, i.e., for which $\overline{s}_1 \geq \varphi_m(s_1)$ and $\overline{s}_2 \geq \varphi_m(s_2)$. Let $(s_\lambda, \overline{s}_\lambda) = \lambda(s_1, \overline{s}_1) + (1-\lambda)(s_2, \overline{s}_2)$ and $\varphi_m(s)_\lambda = \lambda\varphi_m(s_1) + (1-\lambda)\varphi_m(s_2)$, for $0 \leq \lambda \leq 1$. Because v_m^h is convex by Proposition 1,

$$\begin{aligned} v_m^h(s_\lambda, \overline{s}_\lambda) &\leq \lambda v_m^h(s_1, \overline{s}_1) + (1-\lambda) v_m^h(s_2, \overline{s}_2) \\ &\leq \lambda\left(\overline{s}_1 - K\right) + (1-\lambda)\left(\overline{s}_2 - K\right) \\ &= \overline{s}_\lambda - K. \end{aligned}$$

Therefore, in state $(s_\lambda, \overline{s}_\lambda)$, it is also optimal to exercise the option, i.e., $\overline{s}_\lambda \geq \varphi_m(s_\lambda)$. We have just proved that for any two points lying above the function φ_m, the straight line joining these two points is also above the function φ_m. This implies that φ_m is a convex function. The convexity then implies the continuity. ∎

For $m = m^*, \ldots, n-1$, we define the (optimal) *exercise frontier* at time t_m as the graph of φ_m, i.e., the locus of points $(s, \overline{s})$ such that $v_m^h(s, \overline{s}) = v_m^e(\overline{s})$. The function $\varphi_m(s)$ is the *optimal exercise function* and its epigraph is the (optimal) exercise region. It is optimal to exercise the option at time t_m if $\overline{s} \geq \varphi_m(s)$, and hold it until the next exercise date t_{m+1} if $\overline{s} \leq \varphi_m(s)$. The optimal exercise function is illustrated in Section 6 for a numerical example.

5 Numerical Solution of the DP Equation

We now elaborate the numerical approach that we suggest for approximating the solution of the DP equations and the optimal exercise function. The general idea is to partition the positive quadrant of the plane $(s, \overline{s})$ by a rectangular grid and to approximate the value function, at each observation date, by a function which is bilinear on each rectangle of the grid (i.e., piecewise bilinear). However, instead of fitting the approximation to v_m directly, we will make the change of variable $\overline{s}' = (m\overline{s} - s)/(m-1)$ at time t_m and redefine the value function in terms of $(s, \overline{s}')$, by setting $w_m(s, \overline{s}') = v_m(s, \overline{s})$, before fitting a piecewise bilinear approximation to this new function w_m. This change of variable greatly simplifies the integration when the piecewise linear approximation is incorporated into Eq. (5): It allows us to compute the integral formally (explicitly) instead of numerically. Other types of approximations could also be used for w_m, such as a piecewise constant function, or a piecewise linear function over triangles, or bidimensional splines, etc. (see, e.g., de Boor 1978), but we found that the technique proposed here gives a good compromise in terms of the amount of work required to achieve a given precision. Simpler methods (e.g., piecewise constant) require much finer partitions to reach an equivalent precision, whereas more elaborate methods (e.g., higher-dimensional splines) bring significant overhead, especially for performing the integration in (5).

5.1 The Piecewise Bilinear Approximation

To define the grid, let $0 = a_0 < a_1 < \cdots < a_p < a_{p+1} = \infty$ and $0 = b_0 < b_1 < \cdots < b_q < b_{q+1} = \infty$. The grid points are

$$G = \{(a_i, b_j) : 0 \le i \le p \text{ and } 0 \le j \le q\}.$$

These points define a partition of the positive quadrant $[0, \infty) \times [0, \infty)$ into the $(p+1)(q+1)$ rectangles

$$R_{ij} = \{(s, \overline{s}) : a_i \le s < a_{i+1} \text{ and } b_j \le \overline{s} < b_{j+1}\}, \qquad (12)$$

for $i = 0, \ldots, p$ and $j = 0, \ldots, q$.

At time t_m, let

$$\bar{s}' = \begin{cases} (m\bar{s} - s)/(m-1) & \text{if } m > 1, \\ 0 & \text{if } m \le 1, \end{cases} \tag{13}$$

which is the value of $\bar{S}_{m-1}$ if $S(t_m) = s$ and $\bar{S}_m = \bar{s}$, and define

$$w_m(s, \bar{s}') = v_m(s, ((m-1)\bar{s}' + s)/m) = v_m(s, \bar{s}) \tag{14}$$

where $\bar{s} = ((m-1)\bar{s}' + s)/m$ if $m \ge 1$ and $\bar{s} = 0$ if $m = 0$. The function w_m has the same properties as stated for v_m in Proposition 1, except that w_1 does not depend on $\bar{s}$. The recurrence (5)–(6) can be rewritten in terms of w_m as

$$w_m^e(s, \bar{s}') = v_m^e(s, \bar{s}) = (\bar{s} - K)^+, \tag{15}$$

$$w_m^h(s, \bar{s}') = v_m^h(s, \bar{s}) = \rho\, E_{m,s,\bar{s}}\left[w_{m+1}(s\tau_{m+1}, \bar{s})\right] \qquad \text{for } 0 \le m \le n-1, \tag{16}$$

$$w_m(s, \bar{s}') = \begin{cases} w_m^h(s, \bar{s}') & \text{if } 0 \le m \le m^* - 1, \\ \max\{w_m^e(s, \bar{s}'),\, w_m^h(s, \bar{s}')\} & \text{if } m^* \le m \le n-1, \\ w_m^e(s, \bar{s}') & \text{if } m = n. \end{cases} \tag{17}$$

The idea now is to approximate each value function w_m by a bilinear function of $(s, \bar{s}')$ over each rectangle R_{ij}, and continuous at the boundaries. More specifically, the approximation $\widehat{w}_m$ of w_m is written as

$$\widehat{w}_m(s, \bar{s}') = \alpha_{ij}^m + \beta_{ij}^m s + \gamma_{ij}^m \bar{s}' + \delta_{ij}^m s\bar{s}' \tag{18}$$

for $(s, \bar{s}') \in R_{ij}$. To determine the coefficients of these bilinear pieces, we first compute an approximation of w_m at each point of G. This is done via the DP equations (4)–(6), using an available approximation $\widehat{w}_{m+1}$ of the function w_{m+1} (in a manner to be detailed in a moment). Now, given an approximation $\tilde{w}_m(a_i, b_j)$ of $w_m(a_i, b_j)$ for each $(a_i, b_j) \in G$, we impose $\widehat{w}_m(a_i, b_j) = \tilde{w}_m(a_i, b_j)$ at each of these points. For each bounded rectangle R_{ij}, this gives one equation for each corner of the rectangle, that is, a system of 4 linear equations in the 4 unknown (α_{ij}^m, β_{ij}^m, γ_{ij}^m, δ_{ij}^m), which is quick and easy to solve. Over the unbounded rectangles, we simply extrapolate the linear trend observed over the adjacent bounded rectangles, towards infinity. The piecewise-bilinear surface $\widehat{w}_m$ is thus an *interpolation* of the values of $\tilde{w}_m$ at the grid points. Being linear along each rectangle boundary, this function is continuous over the entire positive quadrant. At time t_{n-1}, we use the exact closed-form expression for the value function (since we know it) instead of a piecewise bilinear approximation.

5.2 Explicit Integration for Function Evaluation

We now examine how to compute the approximation $\tilde{w}_m(a_i, b_j)$ given an available piecewise bilinear approximation $\widehat{w}_{m+1}$ of w_{m+1}. Observe that w_m^h in (16) is expressed as an expectation with respect to a *single* random variable, τ_{m+1}, and we have chosen our change of variable $(s, \overline{s}) \to (s, \overline{s}')$ precisely to obtain this property. Moreover, the fact that our approximation of w_{m+1} is piecewise linear with respect to its first coordinate makes the integral very easy to compute *explicitly* when this approximation is put into (16). More specifically, the holding value $w_m^h(s, \overline{s}')$ is approximated by

$$\begin{aligned}
\tilde{w}_m^h(s, \overline{s}') &= \rho\, E_{m,s,\overline{s}}\left[\widehat{w}_{m+1}(s\tau_{m+1}, \overline{s})\right] \\
&= \rho \sum_{i=0}^{p} \sum_{j=0}^{q} \Big(\big(\alpha_{ij}^{m+1} + \gamma_{ij}^{m+1}\overline{s}\big)\, E_{m,s,\overline{s}}\left[I_{ij}(S(t_{m+1}), \overline{s})\right] \\
&\qquad + \big(\beta_{ij}^{m+1} + \delta_{ij}^{m+1}\overline{s}\big)\, s E_{m,s,\overline{s}}\left[I_{ij}(S(t_{m+1}), \overline{s})\tau_{m+1}\right]\Big) \\
&= \rho \sum_{i=0}^{p} \Big(\big(\alpha_{i\xi}^{m+1} + \gamma_{i\xi}^{m+1}\overline{s}\big)\, E_{m,s,\overline{s}}\left[I_{i\xi}(s\tau_{m+1}, \overline{s})\right] \\
&\qquad + \big(\beta_{i\xi}^{m+1} + \delta_{i\xi}^{m+1}\overline{s}\big)\, s E_{m,s,\overline{s}}\left[I_{i\xi}(s\tau_{m+1}, \overline{s})\tau_{m+1}\right]\Big),
\end{aligned}$$

where $I_{ij}(x, y) = I\{(x, y) \in R_{ij}\}$, I is the indicator function, and ξ is the value of k such that $\overline{s} \in [b_k, b_{k+1})$. The function $\tilde{w}_m$ is to be evaluated over the points of G. If we denote $c_{kl} = ((m-1)b_l + a_k)/m$ for $k = 0, \ldots, p$ and $l = 0, \ldots, q$, we obtain

$$\tilde{w}_m^h(a_k, b_l) = \rho \sum_{i=0}^{p} \Big(\big[\alpha_{i\xi}^{m+1} + \gamma_{i\xi}^{m+1} c_{kl}\big] P_{ik} + \big[\beta_{i\xi}^{m+1} + \delta_{i\xi}^{m+1} c_{kl}\big] a_k Q_{ik} \Big) \qquad (19)$$

where ξ is the index such that $c_{kl} \in [b_\xi,\, b_{\xi+1})$,

$$\begin{aligned} P_{ik} &= E\left[I\{a_i \le a_k\tau < a_{i+1}\}\right] \;=\; \Phi(x_{i+1,k}) - \Phi(x_{i,k}), \\ Q_{ik} &= E\left[\tau\, I\{a_i \le a_k\tau < a_{i+1}\}\right] \\ &= \left[\Phi(x_{i+1,k} - \sigma\sqrt{h}) - \Phi(x_{i,k} - \sigma\sqrt{h})\right] e^{(\mu+\sigma^2)h/2}, \end{aligned}$$

τ is a lognormal random variable with parameters μh and $\sigma\sqrt{h}$, Φ is the standard normal distribution, and

$$x_{j,k} = \frac{\ln(a_j/a_k) - \mu h}{\sigma\sqrt{h}}.$$

This yields the approximate value function

$$\tilde{w}_m(a_k, b_l) = \max\left(\tilde{w}_m^h\,(a_k, b_l)\,,\, (c_{kl} - K)^+\right). \tag{20}$$

These values at the grid points are then interpolated to obtain the function $\widehat{w}_m$ as explained previously. Integration and interpolation stages are repeated successively until $m = 0$, where an approximation of w_0 and of the option value v_0 is finally obtained. Note that v_0 depends only on the initial price $s = S(0)$, so it is approximated by a one-dimensional piecewise linear function. An important advantage of choosing the same grid G for all m is that the values of the expectations P_{ik} and Q_{ik} can be precomputed once for all. Evaluating $\tilde{w}_m^h$ at the grid points via (19) then becomes very fast.

It would also be possible to use an adaptive grid, where the grid points change with the observation dates. This could be motivated by the fact that the probability distribution of the state vector changes with time. In that case, the mathematical expectations P_{ik} and Q_{ik} would depend on m and would have to be recomputed at each observation date. This would significantly increase the overhead.

As it turns out, this procedure evaluates, with no extra cost, the option value and the optimal decision at all observation dates and in all states. This could be used for instance to estimate the sensitivity of the option value with respect to the initial price. Of course, Eurasian options can be evaluated via this procedure as well, because they are a special case.

Note that whenever $\tilde{w}_m^h(a_k, b_l) < w_m^e(a_k, b_l)$ at some point (a_k, b_l) for some $m \geq m^*$, i.e., if it is optimal to exercise at that point, we know that we are above the optimal exercise fonction, so that for all $j \geq l$, $\tilde{w}_m^h(a_k, b_j) < w_m^e(a_k, b_j)$ and there is no need to compute $\tilde{w}_m^h(a_k, b_j)$. Our implementation exploits this property to save computations for the pricing of Amerasian options. If $m < m^*$, we cannot avoid the computation of $\tilde{w}_m^h(a_k, b_j)$ because one is not allowed to exercise the option. For this reason, Amerasian options are somewhat faster to evaluate than Eurasian options (see the numerical results in Tables 1 and 4). We also save some computation time by exploiting the fact that whenever $\tilde{w}_m(a_k, b_l)$ is small enough to be negligible (less than some fixed threshold ϵ_2), there is no need to compute $\tilde{w}_m(a_k, b_j)$ for $j < l$; it can be replaced by 0. We took $\epsilon_2 = 10^{-5}$ for the numerical examples of Section 6.

With the approximation $\widehat{w}_m$ of the value function w_m in hand, at any given time step m, the (approximate) optimal decision in any state $(s, \overline{s})$ is easily found by comparing $\widehat{w}_m(s, \overline{s}')$ with $v_m^e(\overline{s}) = (\overline{s} - K)^+$. If one is willing to memorize the approximations $\hat{w}_m$, nothing else is needed. In fact, it suffices to memorize the piecewise-linear approximations over a set of rectangles that covers the optimal exercise frontier, for each m. If one does not want to store these approximations, one may compute and memorize an approximation of the function φ_m for each m. This involves additional work, but storing these one-dimensional functions requires less memory than storing the two-dimensional functions $\widehat{w}_m$. To approximate φ_m, one can first approximate $\varphi_m(s)$, on a grid of values of s, by the value $\overline{s}$ such that $(\overline{s} - K)^+ = \widehat{w}_m(s, \overline{s}')$. This value of $\overline{s}$, denoted $\hat{\varphi}_m(s)$, can be found by a bissection algorithm or by more refined methods. One can then fit a spline to the points $(s, \hat{\varphi}_m(s))$, by least squares or something of that flavor, and retain it as an approximation of the optimal exercise frontier at each exercise date.

The time complexity of the algorithm for computing the value function is $O(pq)$ to precompute the P_{ik}'s and Q_{ik}'s, plus $O(np^2q)$ to compute the sum (19) for each of the pq grid points at each of the n time steps, assuming that we compute *all* the terms of each sum, plus $O(npq)$ to solve the systems of linear equations giving the coefficients in (18). The overall time complexity is thus $O(np^2q)$.

However, when p is large, most of the work is for computing the terms of the sum (19) and time can be saved by observing that several of these terms are negligible. When h is small, the largest terms are typically for i near k, and the terms become negligible when i is far from k. For larger h, the most important terms are for i somewhat larger than k. In our implementation, we initially choose a small constant $\epsilon_1 > 0$ (we took $\epsilon_1 = 10^{-7}$ for our numerical examples). When computing (19), we first select a starting point j, we add up the terms for $i = j+1, j+2, \dots$ and stop as soon as a term is less than ϵ_1, then we add the terms for $i = j, j-1, \dots,$ and stop whenever a term is less than ϵ_1. To select j, if the term for $i = k$ exceeds ϵ_1, we take $j = k$, otherwise (this happens when the value function is practically equal to 0 in the rectangle $R_{k,\xi}$), we try $j = k + \lfloor (p-k)/2 \rfloor$. In the (rare) cases where the term of the sum for this latter j is still less than ϵ_1, then we sum all the terms for $i = k+1, \dots, p+1$, because in this case we don't know where the most important terms are. The number of terms that are actually computed, in the average, is typically much less than p. For large p and small h, the overall time complexity becomes "practically" $O(npq)$.

For comparison, the time complexity of the algorithm of Zvan, Forsyth, and Vetzal (1999) is $O(n'pq)$ where n' is the number of time steps in their time discretization. Unless the distance h between the observation dates is very small, these authors must use $n' \gg n$ to reduce the discretization error with their method. With our method, there is no need to discretize the time.

5.3 Grid Choice

The number of rectangles defined in (12) should be increased in the regions that are visited with high probability and where the value function tends to be less linear. In the experiments reported here, we took $a_1 = S(0) \exp\left(\mu t_{n-1} - 3\sigma\sqrt{t_{n-1}}\right)$, $a_{p-1} = S(0) \exp\left(\mu t_{n-1} + 3\sigma\sqrt{t_{n-1}}\right)$,

$a_p = S(0)\exp(\mu t_{n-1} + 4\sigma\sqrt{t_{n-1}})$, and for $2 \leq i \leq p-2$, a_i was the quantile of order $(i-1)/(p-2)$ of the lognormal distribution with parameters μt_{n-1} and $\sigma\sqrt{t_{n-1}}$. For $b_1, \ldots, b_q$, we partitioned the positive part of the vertical axis into the subintervals $I_0 = [0, b_1)$, $I_1 = [b_1, b_{q/4})$, $I_2 = [b_{q/4}, b_{3q/4})$, $I_3 = [b_{3q/4}, b_q)$, and $I_4 = [b_q, \infty)$, where $b_1 = S(0)\exp(\mu t_{n-1} - 2\sigma\sqrt{t_{n-1}})$, $b_{q/4} = ((n-1)\rho - 1)K/(n-2)$, $b_{3q/4} = nK/(n-2)$, and $b_q = S(0)\exp(\mu t_{n-1} + 3.9\sigma\sqrt{t_{n-1}})$. We then defined the other values so that the b_j's were equally spaced within each of the intervals I_1, I_2, and I_3. This choice is purely heuristic and certainly not optimal.

5.4 Convergence

A rigorous proof of convergence of the DP algorithm as the grid size becomes finer and finer is somewhat tricky, mainly because the state space is unbounded and the value function increases unboundedly when s or $\overline{s}$ goes to infinity. Here we sketch heuristic proof arguments. We first note that if $c = \min(a_p, b_q) \to \infty$, by Lemma 3 of Conze and Viswanathan (1991) and standard large deviations approximation for the normal distribution, we have

$$\begin{aligned}
Q\left[\max_{0\leq t\leq T} S(t)/S(0) > c\right] &= Q\left[\max_{0\leq t\leq T} \ln[S(t)/S(0)] > \ln c\right] \\
&= 1 - \Phi\left(\frac{\ln c - \mu t}{\sigma\sqrt{t}}\right) + \exp\left(\frac{2\mu\ln c}{\sigma^2}\right)\Phi\left(\frac{-\ln c - \mu t}{\sigma\sqrt{t}}\right) \\
&= O\left(\frac{1}{\ln c}\exp\left[-\frac{\ln^2 c}{2\sigma^2 t} + O(\ln c)\right]\right) \\
&= O\left(\frac{1}{\ln c}\, c^{-\ln c/(2\sigma^2 t)}\right).
\end{aligned}$$

Thus, the probability that the trajectory of $\{(S(t), \overline{S}(t)),\ 0 \leq t \leq T\}$ exits the box $B = (0, a_p] \times (0, b_q]$ decreases to 0 at rate faster than $O(1/p(c))$ for any polynomial p. On the other hand, the error on the value function can only increase linearly when s or $\overline{s}$ goes to infinity. For a large enough box B, we can therefore neglect the effect of the approximation error outside the box. We use this heuristic argument to justify the next proposition, which also says that our procedure *overestimates* the option value if we neglect the outside of the box. Define $\delta_a = \sup_{1\leq i\leq p}(a_i - a_{i-1})$ and $\delta_b = \sup_{1\leq j\leq q}(b_j - b_{j-1})$.

Proposition 4. *If $p \to \infty$, $q \to \infty$, $a_p \to \infty$, $b_q \to \infty$, $\delta_a \to 0$, and $\delta_b \to 0$, then for any constant $c > 0$,*

$$\sup_{0 \le m < n} \sup_{(s,\overline{s}) \in (0,c]^2} |\widehat{w}_m(s,\overline{s}) - w_m(s,\overline{s})| \to 0.$$

Moreover, $\widehat{w}_m(s,\overline{s}) \ge w_m(s,\overline{s})$ for all m and all $(s,\overline{s}) \in (0,c]^2$, when a_p and b_q are large enough.

Proof. (Sketch) For each m, w_m is a convex increasing function whose partial derivative with respect to each of its arguments never exceeds 1. Therefore

$$\sup_{1 \le k \le p,\, 1 \le \ell \le q} [w_m(a_{k+1}, b_{\ell+1}) - w_m(a_k, b_\ell)] \le \delta_a + \delta_b. \tag{21}$$

We use backward induction on m to show that for all $m \ge 0$ and for $(s,\overline{s}) \in (0,c]^2$,

$$0 \le \widehat{w}_m(s,\overline{s}) - w_m(s,\overline{s}) \le (\delta_a + \delta_b) \sum_{i=0}^{n-1-m} \rho^i \stackrel{\text{def}}{=} \epsilon_m. \tag{22}$$

This holds for $m = n$ with $\epsilon_n = 0$. For $m = n-1$, (22) holds because of (21) and the fact that $\widehat{w}_{n-1}$ is a bilinear interpolation of the increasing and convex function $\tilde{w}_{n-1} = w_{n-1}$. Now, if we assume that $0 \le \widehat{w}_{m+1}(s,\overline{s}) - w_{m+1}(s,\overline{s}) \le \epsilon_{m+1}$ and if we neglect the error on w_{m+1} outside the box B, we obtain from (16) and (17) that $0 \le \tilde{w}_m(s,\overline{s}) - w_m(s,\overline{s}) \le \rho\epsilon_{m+1}$. Then, because $\widehat{w}_m$ is a piecewise bilinear interpolation of $\tilde{w}_m \ge w_m$, where the latter is increasing and convex, we have $0 \le \widehat{w}_m(s,\overline{s}) - w_m(s,\overline{s}) \le \rho\epsilon_{m+1} + (\delta_a + \delta_b) = \epsilon_m$. Therefore, we obtain that

$$\sup_{0 < s \le a_p,\, 0 < \overline{s} \le b_q} |\widehat{w}_m(s,\overline{s}) - w_m(s,\overline{s})| \le \epsilon_m,$$

where ϵ_m converges to 0 under the assumptions of the proposition.

This argument is not rigorous because we cannot neglect the effect that the error on w_{m+1} outside the box B has on the error on w_m at points near the boundary of B. This is why the proposition's statement is in terms of a constant box $(0,c]^2$ instead of B. Because the distance from this box to the boundary of B increases towards infinity, the effect of the error outside B becomes negligible on the error in the box $(0,c]^2$ at earlier steps. ∎

Table 1: Approximations of the Amerasian call option prices

(K, T, σ)	$p \times q$					
	40×60	60×80	100×160	300×400	v_{eu}(DP)	v_{eu}(simul)
$(100, 0.25, 0.15)$	2.3474	2.3344	2.3246	2.3214	2.1653	2.165
$(100, 0.25, 0.25)$	3.6864	3.6680	3.6553	3.6507	3.3646	3.364
$(100, 0.50, 0.25)$	5.3779	5.3545	5.3386	5.3328	4.9278	4.927
$(105, 0.50, 0.25)$	3.0204	2.9931	2.9734	2.9667	2.8068	2.806
CPU (sec)	0.06	0.14	0.59	14	37	

6 Numerical Experiments and Examples

We now present the results of numerical experiments on the computation of the value of Amerasian options.

Example 1 For our first example, we take the parameter values $S(0) = 100$, $K = 100$, $T = 1/4$ (years), $\sigma = 0.15$, $r = 0.05$, $h = 1/52$, $m^* = 1$, and $n = 13$. We thus have a 13-week contract, with an exercise opportunity at each observation epoch, which is every week. We also consider 3 variants of this example: We first increase the volatility σ from 0.15 to 0.25, we then increase T from 1/4 to 1/2 (26 weeks) while keeping $n = 13$, and finally we increase K from 100 to 105, which yields an out-of-the-money option. In each case, we evaluate the Amerasian option with 4 grid sizes, as indicated in Table 1, where our approximation of $v_0(S(0))$ with each grid size can be found. The table also gives the value of the corresponding Eurasian option computed by DP with a 300×400 grid (denoted v_{eu}(DP)) and the same value estimated by the *efficient* Monte Carlo simulation scheme (using variance reduction) described by Lemieux and L'Ecuyer (1998) and L'Ecuyer and Lemieux (2000) (denoted v_{eu}(simul)). For the latter values, the sample size is always large enough so that the half-length of a 99% confidence interval on the true value is less than 0.0005, which means that all the reported digits are significant. We see no significant difference between the values obtained by the two methods. This certainly reassures us on the precision of the approximation in the DP algorithm.

The approximation of $v_0(S(0))$ seems to converge quite well as the grid size is refined. A grid of 100×160 appears sufficient for a precision of less than 1 penny, and the computing time for that grid size is small. The timings reported here are for a 500Mhz PC running the Linux operating system. The programs are written in FORTRAN and the compiler was "`f90`". The CPU times are approximately the same for each line of the table.

When we compare the value of the Amerasian option with its Eurasian counterpart, we see that the privilege of early exercise increases the value of the option, as expected. The contract becomes more expensive when the volatility or the maturity date are increased (because this gives more chance of achieving a large average), and becomes cheaper when the strike price is increased.

To quantify the impact of increasing the number of early exercise opportunities (and observation dates), we performed additional experiments with the same parameter sets as in Table 1, but with different values of n ranging from 1 to 52. For each of the 4 parameter sets in Table 2, the top and bottom lines give the value of the Amerasian call option computed via DP with a 300×400 grid, and the value of the corresponding Eurasian option computed via efficient simulation (again with 99% confidence interval half-width less than 0.0005), respectively. We see that increasing n *decreases* the option value. The explanation is that increasing the number of observation dates increases the stability of the average prices, and this offsets the advantage of having more exercise opportunities. Note that $n = 1$ corresponds to a standard European call. For $n = 2$, it is optimal to exercise at time t_1 only if $S(t_1) = \overline{S}_1 \geq 2K$ (see section 4.1), which is an extremely rare event with our choice of parameters. This is why the Amerasian and Eurasian options have practically the same value when $n = 2$.

Figures 1 and 2 show the optimal exercise frontier at times t_{n-2} and t_2, respectively, for this example, for the Amerasian option with parameters $(K, T, \sigma, n) = (100, 0.5, 0.25, 52)$. These figures illustrate the fact that the farther away from the time horizon we are, the

Table 2: Amerasian (top) and Eurasian (bottom) option values as a function of n

	n					
(K, T, σ)	1	2	4	13	26	52
$(100, 0.25, 0.15)$	3.635	2.842	2.513	2.321	2.290	2.278
	3.635	2.842	2.443	2.165	2.103	2.072
$(100, 0.25, 0.25)$	5.598	4.395	3.921	3.651	3.609	3.596
	5.598	4.395	3.788	3.364	3.269	3.222
$(100, 0.50, 0.25)$	8.260	6.463	5.745	5.333	5.268	5.247
	8.260	6.462	5.558	4.927	4.787	4.716
$(105, 0.50, 0.25)$	5.988	4.245	3.476	2.967	2.860	2.810
	5.988	4.245	3.389	2.806	2.678	2.614

higher is the exercise frontier: It makes sense to wait even if the current price is somewhat below the current average, because things have time to change. The function w_n (not shown here) depends almost only on $\overline{s}$ (very little on s) and is almost piecewise linear when we are near the time horizon, but the dependence on s and the nonlinearity increase when we move away (backwards) from the time horizon.

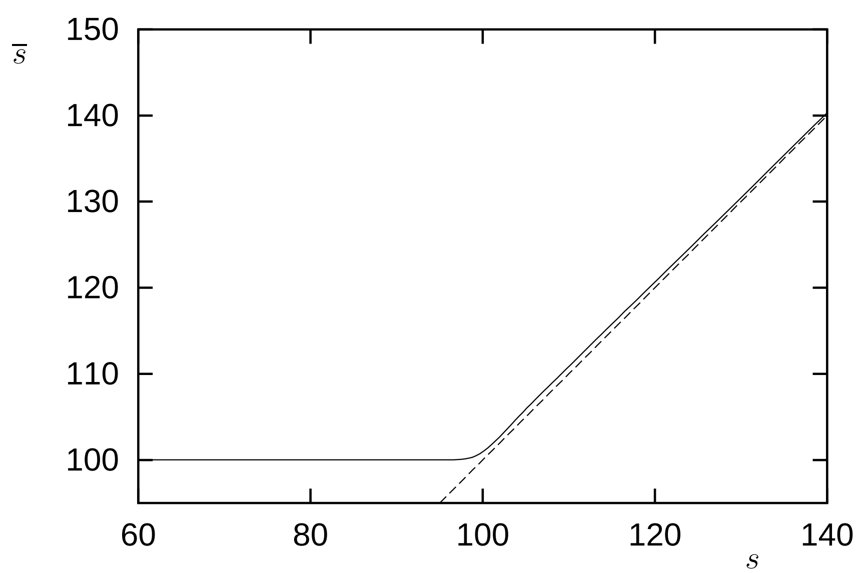

Figure 1: The optimal exercise frontier at time t_{n-2} for Example 1 (solid line). The dotted line is the diagonal $\overline{s} = s$.

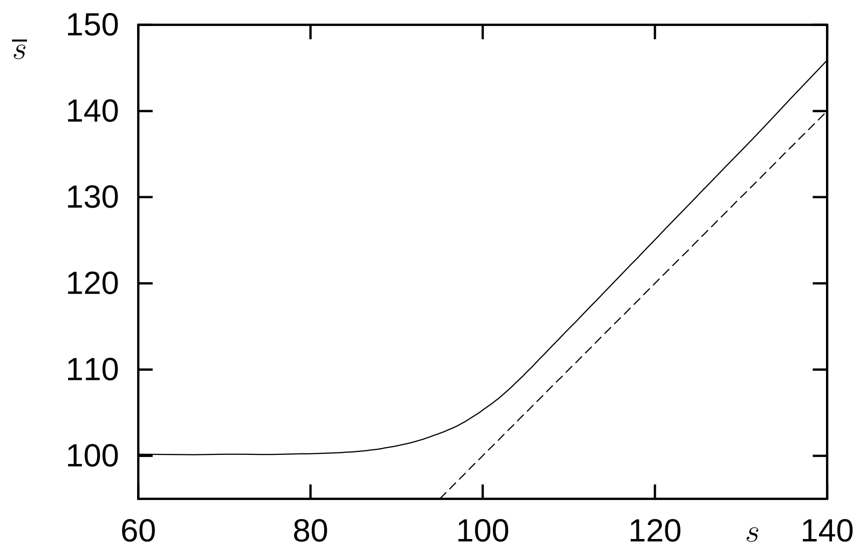

Figure 2: The optimal exercise frontier at time t_2 for Example 1 (solid line). The dotted line is the diagonal $\overline{s} = s$.

Example 2 Our second example is the one considered by Grant, Vora, and Weeks (1997). The time increment is fixed at $h = 1/365$ (one day), the first observation date is at $t_1 = 91/365$ (91 days), and the first exercise opportunity is at $t_{m^*} = 105/365$ (105 days). The other parameters are: $S(0) = 100$, $K = 100$, $T = 120/365$, $\sigma = 0.20$, and $r = 0.09$. Table 3 gives our approximation of $v_0(S(0))$ for the Amerasian option with different grid sizes, as in Table 1. The column labeled GVW gives the 95% confidence interval reported by Grant, Vora, and Weeks (1997) for the value of the option with the strategy obtained by their procedure. The difference from our values could be explained in part by the fact that their procedure systematically underestimates the values of Amerasian call options, because the exercise strategy found by their simulation procedure is suboptimal, so when they use this strategy in a simulation, their estimator of the optimal price has a negative bias. Further negative bias is introduced when they assume that the exercise frontier at each stage is determined by two straight lines. On the other hand, our piecewise-bilinear approximation method overestimates the exact value when the grid is too coarse. The last column reports the value of the corresponding Eurasian option, again with an error less than 0.0005 with

Table 3: Approximation of the option value for the GVW example

(K, σ)	$p \times q$					
	40×60	60×80	100×160	300×400	GVW	v_{eu}(simul)
(100, 0.2)	5.902	5.859	5.825	5.804	5.80 ± 0.02	5.543
(105, 0.2)	3.439	3.401	3.372	3.354	3.35 ± 0.02	3.189
(100, 0.3)	8.127	8.058	8.001	7.966	7.92 ± 0.02	7.652
(105, 0.3)	5.714	5.651	5.601	5.569	5.53 ± 0.02	5.269

Table 4: Approximation of $v_0(100)$ for Example 3

σ		$p \times q = 145 \times 145$		$p \times q = 289 \times 289$		v_{eu}(simul)
		ZFV	DP	ZFV	DP	
0.2	Amerasian	3.213 (168)	3.277 (13)	3.213 (714)	3.242 (103)	
0.2	Eurasian	2.929 (162)	2.973 (17)	2.929 (710)	2.953 (163)	2.953
0.4	Amerasian	5.823 (211)	5.977 (12)	5.825 (718)	5.891 (94)	
0.4	Eurasian	5.160 (161)	5.255 (16)	5.161 (712)	5.210 (161)	5.204

99% confidence.

Example 3 We tried our method with the example given in Tables 3 and 4 of Zvan, Forsyth, and Vetzal (1999). One has $n = 250$, $S(0) = K = 100$, $T = 0.25$, $r = 0.1$, and $\sigma = 0.2$ and 0.4. The results are in Table 4. Each table entry gives the approximate option value computed by the method, followed in parenthesis by the CPU time in seconds. DP refers to our method and ZFV refers to the numbers reported by Zvan, Forsyth, and Vetzal (1999). Taking into account that these authors used a 200Mhz PC whereas we used a 500Mhz PC, our method still appears faster for an equivalent grid size. In terms of precision, if we compare the approximations of the value of the Eurasian option to the value v_{eu}(simul) obtained by simulation (which can be considered as exact to at least the first 4 digits), we see that DP is closer to the true value than ZFV. A possible explanation might be that the ZFV values have significant time-discretization error, in contrast to DP. This advantage of DP over ZFV increases further when the observation dates become sparser.

7 Conclusion

We showed in this paper how to price an Amerasian option on a single asset, under the GBM model, via dynamic programming coupled with a piecewise-polynomial approximation of the value function after an appropriate change of variable. We also proved continuity, monotonicity, and convexity properties of the value function and of the optimal exercise function (which delimits the optimal region of exercise). These properties characterize the optimal exercise strategy for the option. One of our examples illustrates that increasing the number of exercise opportunities tends to decrease the value of the option when the average is taken over the dates where there is an exercise opportunity: The increase in the stability of the average price offsets the value for having more exercise opportunities. Our method compares advantageously with the PDE methods in terms of speed, precision, and simplicity of implementation. It beats the PDE methods especially when the observation dates are sparse.

The computational approach does not rely on the form of the exercise region and could be adapted for pricing other types of discretely-sampled American-style options for which the relevant information process can be modeled as a Markov process over a low-dimensional state space (for the case considered in this paper, the Markov process is $\{(S(t), \overline{S}(t)),\ 0 \leq t \leq T\}$). The GBM with constant volatility could be replaced by a more general CEV process (Cox 1996), or by other models for the underlying asset. For the CEV process, the quantities $P_{i,k}$ and $Q_{i,k}$ in (19) can also be computed in closed form. In general, however, the implementation may have to be reworked differently. A key ingredient is the ability to approximate the value function at each time step. Here we have used piecewise polynomials, with the pieces determined by a rectangular grid that remains the same at all steps. Adapting the grid to the shape of the value function at each step (with the same number of pieces) could provide a better approximation but would bring additional overhead, so it would not necessarily be an improvement. Perhaps a good compromise would be to readjust the grid every d steps (say), for some integer d, and reajust it only in the areas just below the optimal exercise frontier, where the value function is significantly nonlinear.

It may be useful to study, for each case of practical interest, how to exploit the structure of the problem to characterize the value function and the optimal exercise strategy, and to improve the efficiency of the numerical method, as we have done here. When the dimension of the state space is large, e.g., if the payoff depends on several underlying assets, approximating the value function becomes generally much more difficult (we hit the "curse of dimensionality") and pricing the option then remains a challenging problem.

Acknowledgments

This research has been supported by NSERC-Canada grants No. ODGP0110050 and FCAR Grant No. 00ER3218 to the third author. Part of this work was done while M. Breton was a visiting professor at ITAM, México, and while P. L'Ecuyer was on sabbatical at North Carolina State University, USA, in 1998. We thank Christiane Lemieux for letting us use her simulation programs for pricing the Eurasian options, Richard Simard who helped producing the figures, and Lawrence Kryzanowski for his helpful comments.

References

Alziary, B., J. P. Décamps, and P. F. Koehl, "A P.D.E. Approach to Asian Options: Analytical and Numerical Evidence," *Journal of Banking and Finance*, 21 (1997), 613–640.

Barraquand, J. and D. Martineau, "Numerical Valuation of High-Dimensional Multivariate American Securities," *Journal of Financial and Quantitative Analysis*, 30 (1995), 383–405.

Barraquand, J. and T. Pudet, "Pricing of American Path-Dependent Contingent Claims," *Mathematical Finance*, 6 (1996), 17–51.

Bertsekas, D. P., *Dynamic Programming: Deterministic and Stochastic Models*, Prentice-Hall, Englewood Cliffs, N.J., 1987.

Billingsley, P., *Probability and Measure*, second edition, Wiley, New York, 1986.

Black, F. and M. Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81 (1973), 637–654.

Bouaziz, L., E. Briys, and M. Crouhy, "The Pricing of Forward-Starting Asian Options," *Journal of Banking and Finance*, 18 (1994), 823–839.

Boyle, P., M. Broadie, and P. Glasserman, "Monte Carlo methods for Security Pricing," *Journal of Economic Dynamics and Control*, 21 (1997), 1267–1321.

Boyle, P. and Y. Tian, "Pricing Lookback and Barrier Options under the CEV Process," *Journal of Financial and Quantitative Analysis*, 34, 2 (1999), 241–264.

Broadie, M. and P. Glasserman, "Pricing American-Style Securities Using Simulation," *Journal of Economic Dynamics and Control*, 21, 8–9 (1997a), 1323–1352.

Broadie, M. and P. Glasserman, "A Stochastic Mesh Method for Pricing High-Dimensional Americal Options," 1997b. Manuscript.

Carverhill, A. P. and L. J. Clewlow, "Flexible Convolution," *Risk*, 3 (1990), 25–29.

Conze, A. and Viswanathan, "Path Dependent Options: The Case of Lookback Options," *Journal of Finance*, 46 (1991), 1893–1906.

Cox, J. C., "The Constant Elasticity of Variance Option Pricing Model," *Journal of Portfolio Management*, 22 (1996), 15–17.

Cox, J. C., S. A. Ross, and M. Rubinstein, "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7 (1979), 229–263.

Curran, M., "Valuing Asian and Portfolio Options by Conditioning on the Geometric Mean Price," *Management Science*, 40, 12 (1994), 1705–1711.

de Boor, C., *A Practical Guide to Splines*, number 27 in Applied Mathematical Sciences Series, Springer-Verlag, New York, 1978.

Geman, H. and M. Yor, "Bessel Processes, Asian Options, and Perpetuities," *Mathematical Finance*, 3 (1993), 349–375.

Glasserman, P., P. Heidelberger, and P. Shahabuddin, "Asymptotically Optimal Importance

Sampling and Stratification for Pricing Path Dependent Options," *Mathematical Finance*, 9, 2 (1999), 117–152.

Grant, D., G. Vora, and D. Weeks, "Path-Dependent Options: Extending the Monte Carlo Simulation Approach," *Management Science*, 43 (1997), 1589–1602.

Haurie, A. and P. L'Ecuyer, "Approximation and Bounds in Discrete Event Dynamic Programming," *IEEE Trans. on Automatic Control*, AC-31, 3 (1986), 227–235.

Hull, J., *Options, Futures, and Other Derivative Securities*, fourth edition, Prentice-Hall, Englewood-Cliff, N.J., 2000.

Hull, J. and A. White, "Efficient Procedures for Valuing European and American Path-Dependent Options," *Journal of Derivatives*, 1, Fall (1993), 21–31.

Karatzas, I. and S. E. Shreve, *Methods of Mathematical Finance*, Springer, New York, 1998.

Kemna, A. G. Z. and A. C. F. Vorst, "A Pricing Method for Options Based on Average Asset Values," *Journal of Banking and Finance*, 14 (1990), 113–129.

L'Ecuyer, P. and C. Lemieux, "Variance Reduction via Lattice Rules," *Management Science* (2000). To appear.

L'Ecuyer, P. and J. Malenfant, "Computing Optimal Checkpointing Strategies for Rollback and Recovery Systems," *IEEE Transactions on Computers*, 37, 4 (1988), 491–496.

Lemieux, C. and P. L'Ecuyer, "Efficiency Improvement by Lattice Rules for Pricing Asian Options," In *Proceedings of the 1998 Winter Simulation Conference*, ed. D. J. Medeiros, E. F. Watson, J. S. Carson, and M. S. Manivannan, IEEE Press, 1998, 579–586.

Lemieux, C. and P. L'Ecuyer, "Selection Criteria for Lattice Rules and Other Low-Discrepancy Point Sets," *Mathematics and Computers in Simulation* (2000). To appear.

Levy, E., "Pricing European Average Rate Currency Options," *Journal of International Money and Finance*, 11 (1992), 474–491.

Ritchken, P., L. Sankarasubramanian, and A. M. Vijh, "The Valuation of Path-Dependent Contracts on the Average," *Management Science*, 39 (1993), 1202–1213.

Rockafellar, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.

Rogers, L. C. G. and Z. Shi, "The Value of an Asian Option," *Journal of Applied Probability*, 32 (1995), 1077–1088.

Turnbull, S. M. and L. M. Wakeman, "A Quick Algorithm for Pricing European Average Options," *Journal of Financial and Quantitative Analysis*, 26 (1991), 377–389.

Wilmott, P., J. Dewynne, and J. Howison, *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press, Oxford, 1993.

Wu, R. and M. C. Fu, "Optimal Exercise Policies and Simulation-Based Valuation for American-Asian Options," 2000. manuscript.

Zhang, P. G., "Flexible Arithmetic Asian Options," *Journal of Derivatives*, 2 (1995), 53–63.

Zvan, R., P. A. Forsyth, and K. R. Vetzal, "Robust Numerical methods for PDE Models of Asian Options," *Journal of Computational Finance*, 1, 2 (1998), 39–78.

Zvan, R., P. A. Forsyth, and K. R. Vetzal, "Discrete Asian Barrier Options," *Journal of Computational Finance*, 3, 1 (1999), 41–67.