Stochastic Finance 2004
Autumn School & International Conference

Upper Bounds for American Option Prices using Regression with Martingale Basis Functions

N. P. Firth*
OCIAM, Mathematical Institute, 24-29 St. Giles', Oxford, OX1 3LB UK
Email: firth@maths.ox.ac.uk
URL: http://www.maths.ox.ac.uk/∼firth

Abstract. High dimensional American options have no analytic solution and are difficult to price numerically. Progress has been made in using Monte Carlo simulation to give both lower and upper bounds on the price. Building on an idea of Glasserman and Yu we investigate the utility of martingale basis functions in regression based approximation methods. Regression methods are known to give lower bounds easily, however upper bounds are usually computationally expensive. Martingale basis functions enable fast calculation of upper bounds on the price. These algorithms are implemented in the open-source derivatives pricing library, QuantLib.

1. Introduction.

No closed form solutions have been found for American option prices[1]. Therefore much work has been done pricing American options numerically. Early examples include finite differences [12] and the binomial lattice [20]. Grid based methods work well for single asset options, and have been extended to higher dimensions [8, 10]. However, these methods suffer from the 'curse of dimensionality', as they become exponentially more expensive as the dimension increases, and cannot be used for options depending on more than three assets. Monte Carlo methods are better suited to high dimensional problems.

Using Monte Carlo to price European style financial derivatives was first suggested in [7]. Progress in using simulation methods to price American style options was stimulated by Tilley [42]. This and other early methods [4, 6, 14, 15, 24] are reviewed in Boyle et al. [9]. Since then the stochastic mesh method [15] has been made more efficient [2, 16] and has been modified to use low-discrepancy sequences [11]. Other methods include parameterization of the optimal exercise boundary [6, 24, 30], a quantization tree algorithm [3], wavelets [21, 22], irregular grid approximations [5], and sparse grid methods [37].

Regression methods include [18, 33, 43]. The regression method of Longstaff and Schwartz [33] has proved particularly popular, due to its accuracy and simplicity. The paper explains the method, but an introduction to the method is also given in Tavella [41]. Further analysis can be found in Stentoft [39]. Convergence results for regression methods are given in [19, 43]. The relationship between the number of basis functions and number of paths required is investigated by Glasserman and Yu [28]. For a comparison of simulation approaches see Fu et al. [25] or the recent book by Glasserman [27].

Glasserman and Yu [29] investigate the relative merits of 'regression now' versus 'regression later'. 'Regression now' involves using basis functions defined at the current time step and regressing discounted option values from the next period. 'Regression later' uses option values and basis functions defined one time step ahead. Glasserman and Yu [29] prove a theorem indicating that, under certain conditions, 'regression later' should produce more accurate estimates than 'regression now'. This involves using basis functions that are martingales in the regression. However, they do not suggest basis functions and do not implement their method.

Primal-Dual representations of the American option problem allow both an upper and lower bound to be calculated [1, 32, 34, 38]. However, the upper bound involves calculating an expectation, which has been done using another Monte Carlo simulation. This simulation on simulation is computationally expensive. Glasserman and Yu [29] suggest using basis functions that are martingales in the regression. This allows the immediate calculation of the required conditional expectation. Again, there is no numerical evaluation of this idea in Glasserman and Yu [29].

*The author would like to thank the EPSRC and Nomura International plc. for funding this research. The author would also like to thank Dr. W. T. Shaw, Dr. B. Hambly, A. Dickinson, and particularly Dr. R. A. Stalker Firth.

[1]analytical work includes [13, 17, 35]

The rest of this paper is structured as follows: We formulate the American option pricing problem, following [29], as an optimal stopping problem whose solution can be found by dynamic programming, in section 2. In section 3 we present algorithms for the implementation, as well as the theoretical statements on the approximation, of the dynamic programming problem. In section 4 we state Theorem 1 from [29], which suggests that 'regression later' may produce more accurate estimates than 'regression now'. We construct regression basis functions that are martingales under geometric Brownian motion in section 5 to enable us to test Theorem 1. In sections 6 and 7 we present the theory and algorithms for calculating low and high biased estimates for American option prices, respectively. The accuracy and speed of this method are compared with existing methods for single asset options. The numerical results are presented in section 8. In section 9 we evaluate the results and indicate possible future directions. Conclusions are offered in section 10.

2. Problem Formulation. We use the notation of Glasserman and Yu [29] and formulate the American option pricing problem as an optimal stopping problem. X_0 is the fixed initial financial information, $X_0, X_1, \ldots, X_m$ is a $\mathbb{R}^d$ valued Markov chain describing all relevant financial information. If exercised at time i, $i = 0, 1, \ldots, m$, the option pays out $h_i(X_i)$, each h_i being a function from $\mathbb{R}^d$ into $[0, \infty)$. $\mathcal{T}_i$ denotes the set of randomized stopping times taking values in $\{i, i+1, \ldots, m\}$. We define

$$V_i^*(x) = \sup_{\tau \in \mathcal{T}_i} \mathbb{E}\left[h_\tau(X_\tau) \,|\, X_i = x\right], \quad x \in \mathbb{R}^d, \quad i = 0, 1, \ldots, m, \tag{1}$$

where $V_i^*(x)$ is the value of the option at time i and in state x. All expectations are under the risk neutral measure. Note that V^* indicates the true option value. As noted in Glasserman and Yu [29] restricting τ to be an ordinary stopping time, such that each event $\{\tau = i\}$ be determined by $X_1, \ldots, X_i$, does not allow for stopping rules estimated through simulation. Therefore we allow randomized stopping times to depend on other random variables, independent of $X_{i+1}, \ldots, X_m$. We want to find $V_0^*(X_0)$. The option value can be written

$$V_m^*(x) = h_m(x) \tag{2}$$

$$V_i^*(x) = \max\left(h_i(x), \mathbb{E}\left[V_{i+1}^*(X_{i+1}) | X_i = x\right]\right), \tag{3}$$

$i = 0, 1, \ldots, m-1$. The dynamic programming equations can also be written in terms of the continuation value

$$C_i^*(x) = \mathbb{E}\left[V_{i+1}^*(X_{i+1}) | X_i = x\right], \quad i = 0, 1, \ldots, m-1,$$

as

$$C_m^*(x) = 0 \tag{4}$$

$$C_i^*(x) = \mathbb{E}\left[\max\left(h_{i+1}(X_{i+1}), C_{i+1}^*(X_{i+1})\right) | X_i = x\right], \tag{5}$$

$i = 0, 1, \ldots, m-1$. The option values satisfy

$$V_i^*(x) = \max\left(h_i(x), C_i^*(x)\right).$$

As in Glasserman and Yu [29], we absorb discount factors into the definitions of X_i and h_i. In algorithms we include discounting explicitly.

3. Approximate Dynamic Programming. As in Glasserman and Yu [29] consider approximations using time dependent basis functions $\psi_i(X_i)$ that are functions from $\mathbb{R}^d$ to $\mathbb{R}$. We approximate (2)–(3) and (4)–(5) as follows. Consider approximations of the form

$$V_i^*(x) \approx \sum_{k=0}^{K} \beta_{ik} \psi_{ik}(x)$$

and

$$C_i^*(x) \approx \sum_{k=0}^{K} \gamma_{ik} \psi_{ik}(x)$$

for some constants β_{ik} and γ_{ik}. We approximate these coefficients by projection onto the span of $\psi_{ik}(X_i)$, $k = 0, 1, \ldots, K$. We follow Glasserman and Yu [29] and define, for any square–integrable random variable Y, the projection

$$\Pi_i Y = \mathbb{E}\left[Y \psi_i\,(X_i)^T\right] \left(\mathbb{E}\left[\psi_i(X_i)\,\psi_i(X_i)^T\right]\right)^{-1}$$

as

$$\Pi_i Y = \sum_{k=0}^{K} a_k \psi_{ik}(X_i)\,, \tag{6}$$

where

$$(a_0, \ldots, a_K) = \mathbb{E}\left[Y \psi_i\,(X_i)^T\right] \left(\mathbb{E}\left[\psi_i(X_i)\,\psi_i(X_i)^T\right]\right)^{-1}\,, \tag{7}$$

and the residual $Y - \Pi_i Y$ is uncorrelated with the basis functions. We also write the function defined by the coefficients (7), as in Glasserman and Yu [29],

$$(\Pi_i Y)(x) = \sum_{k=0}^{K} a_k \psi_{ik}(x)\,.$$

Following Glasserman and Yu [29] we impose the condition
(**C1**). For each $i = 1, \ldots, m$, $\psi_{i0} \equiv 1$, $\mathbb{E}\left[\psi_{ik}(X_i)\right] = 0$, $k = 1, \ldots, K$, and

$$\mathbb{E}\left[\psi_i(X_i)\,\psi_i(X_i)^T\right] = \begin{pmatrix} 1 & & & & \\ & \sigma_{i1}^2 & & & \\ & & \sigma_{i2}^2 & & \\ & & & \ddots & \\ & & & & \sigma_{iK}^2 \end{pmatrix},$$

with $0 < \sigma_{il}^2 < \infty$ for all i, l.

This condition ensures that the matrix is finite and nonsingular. The basis functions must be linearly independent and have finite variance. The requirement that they also be uncorrelated can be achieved through a linear transformation.

3.1. Regression Now. Glasserman and Yu [29] interpret approximate methods as regressions in this framework see also [27]. In particular, they distinguish between 'regression now' and 'regression later'. In 'regression now' we perform the regression at the current time step. In 'regression later' we perform the regression at the next time step.

We define an approximation to (4)–(5) by

$$C_m(x) = 0 \tag{8}$$
$$C_i(x) = (\Pi_i \max\left(h_{i+1}(X_{i+1}), C_{i+1}(X_{i+1})\right))(x)\,. \tag{9}$$

We use (6) and write the continuation value as a linear combination of basis functions

$$C_i(x) = \sum_{k=0}^{K} \beta_{ik} \psi_{ik}(x)\,,$$

where the coefficients β_{ik} are defined as in (7) with Y replaced by

$$V_{i+1}(X_{i+1}) \equiv \max\left(h_{i+1}(X_{i+1}), C_{i+1}(X_{i+1})\right)\,.$$

Glasserman and Yu [29] define the residual ε_{i+1} by writing

$$V_{i+1}(X_{i+1}) = \sum_{k=0}^{K} \beta_{ik} \psi_{ik}(X_i) + \varepsilon_{i+1}\,,$$

and prove that under the condition
(**C2**). For all $i = 0, \ldots, m-1$,

$$\mathbb{E}\left[\varepsilon_{i+1} \,|\, X_i\right] = 0\,,$$

the approximate dynamic programming problem (8)–(9) is exact.

Approximation. If (C2) does not hold we can use a sample estimation of the projection through simulation. We denote the approximation to the option value V_i^* by $\hat{V}_i$, and the approximation to the continuation value C_i^* by $\hat{C}_i$. For $j = 1, \ldots, N$ let $(X_{1j}, \ldots, X_{mj})$ be independent replications of the underlying Markov chain. Set $\hat{C}_m = 0$ and

$$\hat{C}_i(x) = \sum_{k=0}^{K} \hat{\beta}_{ik} \psi_{ik}(x)\,,$$

where $\hat{\beta}_i^T = (\hat{\beta}_{i0}, \ldots, \hat{\beta}_{iK})$ is the vector of regression coefficients

$$(\hat{\beta}_{i0}, \ldots, \hat{\beta}_{iK}) = \left(\sum_{j=1}^{N} \hat{V}_{i+1}(X_{i+1,j})\, \psi_i(X_{ij})^T \right) \left(\sum_{j=1}^{N} \psi_i(X_{ij})\, \psi_i(X_{ij})^T \right)^{-1} .$$

Also

$$\hat{V}_{i+1} = \max(h_{i+1}, \hat{C}_{i+1})\,,$$

$i = 0, \ldots, m-1$. The initial state X_0 is fixed, we set

$$\hat{C}_0(X_0) = \frac{1}{N} \sum_{j=1}^{N} \hat{V}_1(X_{1j})$$

and $\hat{V}_0(X_0) = \max(h_0(X_0), \hat{C}_0(X_0))$. As this method uses future information, but also estimates a sub–optimal exercise strategy the estimate has mixed bias. Estimators with definite low bias and definite high bias will be given in sections 6 and 7, respectively. Convergence results are given in [19, 43].
Algorithm 1. We make the method concrete by describing the computation in algorithm 1. This describes the Least Squares Monte Carlo (LSMC) method of [33] in this setting, as well as the method developed in [43]. In Glasserman and Yu [29] discounting is not explicitly considered. We explicitly denote the discount factor between times i and $i+1$ by $D_{i,i+1}$. In the case of a continuously compounded constant short interest rate r the discount factor is given by $D_{i,i+1} = e^{-r(t_{i+1}-t_i)}$. The option exercise value, h_i, is denominated in time i dollars. The asset price tensor $x_{ij\lambda}$ is the price of asset λ at time i for path j.

Algorithm 1 Least Squares Monte Carlo — Regression Now

Choose the number of asset paths, N, to simulate, $j = 1, \ldots, N$
$x_{ij\lambda}$ generate the asset price tensor (e.g. under geometric Brownian motion) from initial data X_0
$\hat{V}_j \leftarrow h_m(x_{mj\lambda})$ evaluate the final time payoff for each path
while Loop backwards through time steps from $i = m-1$ to 1 **do**
  $\hat{V}_j \leftarrow D_{i,i+1}\hat{V}_j$ discount the estimated option value to the current time
  $A_{jk} \leftarrow \psi_{ik}(x_{ij\lambda})$ the design matrix
  $\hat{\beta}_k \leftarrow (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\hat{V}_j$ the regression
  $\hat{C}_j \leftarrow A_{jk}\hat{\beta}_k$ the estimated continuation value
  $h_j \leftarrow h_i(x_{ij\lambda})$ the value of immediate exercise (payoff) for each path
  if $h_j > \hat{C}_j$ **then**
    $\hat{V}_j \leftarrow h_j$
  end if
end while
$\hat{C}_0 \leftarrow \sum D_{0,1}\hat{V}_j/N$ final discount step and initial continuation value
$\hat{V}_0 \leftarrow \max(h_0(X_0), \hat{C}_0)$ the estimated option value

---

3.2. Regression Later. Glasserman and Yu [29] interpret the method of Broadie et al. [16] as a regression method, but with the regression taking place one time step ahead. They refer to this as 'regression later'. To do this Glasserman and Yu [29] introduce the idea of using basis functions in the regression which are martingales. This condition is expressed as
(**C3**). Martingale property :

$$\mathbb{E}\left[\psi_{i+1}(X_{i+1}) \,|\, X_i\right] = \psi_i(X_i)\,, \qquad i = 0, 1, \ldots, m-1\,.$$

We write the continuation value as a linear combination of basis functions

$$C_i^+(x) = \sum_{k=0}^{K} \gamma_{ik}\psi_{ik}(x) \quad i = 1, \ldots, m-1\,.$$

Also $V_{i+1}^+ = \max(h_{i+1}, C_{i+1}^+)$, $i = 0, \ldots, m-1$. Define the residual ε_{i+1}^+ through [29]

$$V_{i+1}^+(X_{i+1}) = \sum_{k=0}^{K} \gamma_{ik}\psi_{i+1,k}(X_{i+1}) + \varepsilon_{i+1}^+$$

with $\varepsilon_{i+1}^+ = V_{i+1}^+(X_{i+1}) - \Pi_{i+1}V_{i+1}^+(X_{i+1})$ uncorrelated with the components of $\psi_{i+1}(X_{i+1})$. Consider the conditions:
(**C4**). For all $i = 0, \ldots, m-1$ and $k = 0, \ldots, K$,

$$\mathbb{E}\left[\varepsilon_{i+1}^+\left(\psi_{i+1,k}(X_{i+1}) - \psi_{ik}(X_i)\right)\right] = 0\,.$$

(**C4′**). For all $i = 0, \ldots, m-1$,

$$\mathbb{E}\left[\varepsilon_{i+1}^+ \,|\, X_i\right] = 0\,.$$

Glasserman and Yu [29] prove that if (C3) and (C4) hold then $C_i^+ = C_i$ for all i. This means that the solution satisfies the approximate dynamic programming problem (8)–(9) exactly. They also prove that if (C3) and (C4') hold then in addition $C_i^+ = C_i^*$ for all i.
Approximation. We follow Glasserman and Yu [29] and write $\hat{V}^+(x)$ and $\hat{C}^+(x)$ to indicate that the estimates have been found using 'regression later'. As in the 'regression now' case above, for $j = 1, \ldots, N$ let $(X_{1j}, \ldots, X_{mj})$ be independent replications of the underlying Markov chain. Define $\hat{C}_m^+ = 0$. Under this assumption Glasserman and Yu [29] prove

$$\hat{C}_i^+(x) = \sum_{k=0}^{K} \hat{\gamma}_{ik}\psi_{ik}(x) \quad i = 1, \ldots, m-1\,,$$

converges to the true value, where $\hat{\gamma}_i^T = (\hat{\gamma}_{i0}, \ldots, \hat{\gamma}_{iK})$ is the vector of regression coefficients,

$$(\hat{\gamma}_{i0}, \ldots, \hat{\gamma}_{iK}) = \left(\sum_{j=1}^{b} \hat{V}_{i+1}^+(X_{i+1,j})\,\psi_{i+1}(X_{ij})^T\right)\left(\sum_{j=1}^{b} \psi_{i+1}(X_{ij})\,\psi_{i+1}(X_{ij})^T\right)^{-1}\,.$$

Note that the regression coefficients are estimated using later basis functions ψ_{i+1}, rather than ψ_i. Also $\hat{V}_{i+1}^+ = \max(h_{i+1}, \hat{C}_{i+1}^+)$, $i = 0, \ldots, m-1$.
Algorithm 2. Algorithm 2 details the computation. Note that the discounting of the estimated continuation value is done implicitly using the martingale condition (C3) on the basis functions. However, we have not yet defined any martingale basis functions. The implementation can be made more efficient by storing the A_{jk} from the previous time step for use in the next time step.

Algorithm 2 Least Squares Monte Carlo — Regression Later

Require: Basis functions $\psi_{ik}(x)$ are martingales satisfying condition (C3)
  Choose the number of asset paths, N, to simulate, $j = 1, \ldots, N$
  $x_{ij\lambda}$ generate the asset price tensor (e.g. under geometric Brownian motion) from initial data X_0
  $\hat{V}_j^+ \leftarrow h_m(x_{mj\lambda})$ evaluate the final time payoff for each path
  while Loop backwards through time steps from $i = m-1$ to 1 **do**
    $A_{jk} \leftarrow \psi_{i+1,k}(x_{i+1,j\lambda})$ the design matrix - later
    $\hat{\gamma}_k \leftarrow (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\hat{V}_{i+1,j}^+$ the regression
    $A_{jk} \leftarrow \psi_{ik}(x_{ij\lambda})$ the design matrix - now
    $\hat{C}_j^+ \leftarrow A_{jk}\hat{\gamma}_k$ the estimated continuation value
    $h_j \leftarrow h_i(x_{ij\lambda})$ the value of immediate exercise (payoff) for each path
    if $h_j > \hat{C}_j^+$ **then**
      $\hat{V}_j^+ \leftarrow h_j$
    else
      $\hat{V}_j^+ \leftarrow D_{i,i+1}\hat{V}_j^+$ discount the estimated option value to the current time
    end if
  end while
  $\hat{C}_0^+ \leftarrow \sum D_{0,1}\hat{V}_j^+/N$ final discount step and initial continuation value
  $\hat{V}_0^+ \leftarrow \max(h_0(X_0), \hat{C}_0^+)$ the estimated option value

4. Comparison.

To state Theorem 1 from [29] we need to define stronger conditions on the residuals:

(**C5a**).

$$\mathbb{E}\left[\varepsilon_{i+1}^+ \,|\, \psi_{i+1}(X_{i+1})\right] = 0 \quad \text{and} \quad \mathbb{E}\left[(\varepsilon_{i+1}^+)^2 \,|\, \psi_{i+1}(X_{i+1})\right] = \text{var}\left[\varepsilon_{i+1}^+\right]$$

(**C5b**).

$$\mathbb{E}\left[\varepsilon_{i+1} \,|\psi_i(X_i)\right] = 0 \quad \text{and} \quad \mathbb{E}\left[(\varepsilon_{i+1})^2 \,|\, \psi_i(X_i)\right] = \text{var}\left[\varepsilon_{i+1}\right]$$

We also define the coefficients of determination

$$\begin{aligned} R_\beta^2 &= \text{var}\left[\beta_i^T \psi_i(X_i)\right] / \text{var}\left[V_{i+1}(X_{i+1}\right] \\ R_\gamma^2 &= \text{var}\left[\gamma_i^T \psi_{i+1}(X_{i+1})\right] / \text{var}\left[V_{i+1}(X_{i+1}\right] . \end{aligned}$$

Informally we can think of the coefficient of determination as giving an indication of the percentage of the variation explained by the regression. Therefore a higher coefficient of determination indicates a better regression fit.

We write the covariance matrix of $\hat{\beta}$ as $\text{cov}[\hat{\beta}]$ and let $\Sigma_\beta = \lim_{N\to\infty} N\text{cov}[\hat{\beta}]$ whenever the limit exists. Similarly denote the limiting covariance matrix of $\hat{\gamma}$ by Σ_γ. The existence of these limits is implied by the following uniform integrability conditions:

(**C6**). As $N \to \infty$,

$$N\,\mathbb{E}\left[\Big(\sum_{j=1}^N \psi_i(X_{ij})\,\psi_i(X_{ij})^T\Big)^{-1}\right] \to \left(\mathbb{E}\left[\psi_i(X_i)\,\psi_i(X_i)^T\right]\right)^{-1}$$

and

$$N\,\mathbb{E}\left[\Big(\sum_{j=1}^N \psi_{i+1}(X_{i+1,j})\,\psi_{i+1}(X_{i+1,j})^T\Big)^{-1}\right] \to \left(\mathbb{E}\left[\psi_{i+1}(X_{i+1})\,\psi_{i+1}(X_{i+1})^T\right]\right)^{-1} .$$

We can now state Theorem 1 of Glasserman and Yu [29]

Theorem 1 *If (C1) and (C3)–(C4) hold, then $R_\beta^2 \leq R_\gamma^2$. If also (C5)–(C6) hold then $\Sigma_\gamma \leq \Sigma_\beta$.*

Note that this only holds over a single time step.

5. Martingale Basis Functions. Particular martingale basis functions satisfying (C3) are not given by Glasserman and Yu [29]. We suggest basis functions that are martingales under geometric Brownian motion.

5.1. One Dimensional Martingale Basis Functions. We want to find basis functions that satisfy the martingale property (C3). Let X_i be a scalar describing the evolution of a single asset. The simplest basis function is a constant, we see immediately this satisfies the condition (C3)

$$\mathbb{E}\left[1|X_i\right] = 1\,.$$

To consider what happens for other basis functions we need to specify a model for the asset price process. The risk–neutral asset price process follows the stochastic differential equation (SDE)

$$\frac{dX}{X} = r\,dt + \sigma\,dW$$

where r is the risk free interest rate, σ is the volatility of asset returns. W is a standard Brownian motion. The initial asset price is X_0. The asset does not pay any dividends. We can solve the SDE exactly [27] to find

$$X_t = X_0 e^{(r-\sigma^2/2)t+\sigma\sqrt{t}N(0,1)}\,,$$

where $N(0,1)$ is the normal distribution with mean 0 and standard deviation 1. We expect the monomial

$$\psi_i(X_i) = e^{-r(t_i-t_0)}X_i\,,$$

to satisfy the martingale condition (C3) under GBM. To verify this we calculate

$$\mathbb{E}\left[\psi_{i+1}(X_{i+1})|X_i\right] = \mathbb{E}\left[e^{-r(t_{i+1}-t_0)}X_{i+1}|X_i\right] \tag{10}$$

$$= e^{-r(t_{i+1}-t_0)}\mathbb{E}\left[X_{i+1}|X_i\right]\,, \tag{11}$$

as the discount factor is a constant. To evaluate the expectation we integrate the function against the normal density

$$\frac{1}{\sqrt{2\pi}}\,e^{-x^2/2}\,,$$

hence, setting $\delta = t_{i+1} - t_i$,

$$\mathbb{E}\left[X_{i+1}|X_i\right] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} X_i\,e^{(r-\sigma^2/2)\delta+\sigma\sqrt{\delta}\,x}e^{-x^2/2}\,dx = X_i e^{r\delta}\,.$$

Continuing from (11) we have

$$\begin{aligned}\mathbb{E}\left[\psi_{i+1}(X_{i+1})|X_i\right] &= e^{-r(t_{i+1}-t_0)}X_i\,e^{r(t_{i+1}-t_i)}\\ &= e^{-r(t_i-t_0)}X_i\\ &= \psi_i(X_i)\,.\end{aligned}$$

Hence, we have a martingale basis function for a single asset under geometric Brownian motion. We need more than two basis functions, so we look for a basis function with squares of the price process. First we investigate the expectation

$$\begin{aligned}\mathbb{E}\left[(X_{i+1})^2|X_i\right] &= \frac{1}{\sqrt{2\pi}}\int_{\infty}^{\infty}\left(X_i e^{(r-\sigma^2/2)\delta+\sigma\sqrt{\delta}\,x}\right)^2 e^{-x^2/2}\,dx\\ &= (X_i)^2\,e^{(2r+\sigma^2)\delta}\,.\end{aligned}$$

Therefore, we try the basis function

$$\psi_i(X_i) = (X_i)^2\,e^{-(2r+\sigma^2)(t_i-t_0)}\,.$$

We test the martingale condition (C3) as above,

$$\begin{aligned}\mathbb{E}\left[\psi_{i+1}(X_{i+1})\,|\,X_i\right] &= \mathbb{E}\left[(X_{i+1})^2\, e^{-(2r+\sigma^2)(t_{i+1}-t_0)}\,|\,X_i\right] \\ &= e^{-(2r+\sigma^2)(t_{i+1}-t_0)}\,\mathbb{E}\left[(X_{i+1})^2\,|\,X_i\right] \\ &= e^{-(2r+\sigma^2)(t_{i+1}-t_0)}\,(X_i)^2\, e^{(2r+\sigma^2)(t_{i+1}-t_i)} \\ &= (X_i)^2\, e^{-(2r+\sigma^2)(t_i-t_0)} \\ &= \psi_i(X_i)\,.\end{aligned}$$

Hence we have three martingale basis functions,

$$\begin{aligned}\psi_{i0}(X_i) &= 1 &(12)\\ \psi_{i1}(X_i) &= X_i\, e^{-r(t_i-t_0)} &(13)\\ \psi_{i2}(X_i) &= (X_i)^2\, e^{-(2r+\sigma^2)(t_i-t_0)}\,. &(14)\end{aligned}$$

General Form of Martingale Basis Functions. Now we find general monomial basis functions that are martingales under geometric Brownian motion. Consider monomial basis functions of the form

$$\psi_{ik}(X_i) = a_{ik}\,(X_i)^k\,,$$

where we consider each a_{ik} fixed. We investigate the condition (C3),

$$\mathbb{E}\left[\psi_{i+1,k}(X_{i+1})|X_i\right] = \mathbb{E}\left[a_{i+1,k}(X_{i+1})^k\,|\,X_i\right]\,. \tag{15}$$

In general we can write

$$X_{i+\delta} = X_i\epsilon_\delta\,,$$

where

$$\epsilon_\delta = e^{r\delta}\, e^{\sigma W_\delta-\sigma^2\delta/2}\,.$$

Hence, from (15),

$$\begin{aligned}\mathbb{E}\left[\psi_{i+1,k}(X_{i+1})|X_i\right] &= a_{i+1,k}\,\mathbb{E}\left[X_i^k\epsilon_\delta^k\,|\,X_i\right] \\ &= a_{i+1,k}\,(X_i)^k\,\mathbb{E}\left[\epsilon_\delta^k\,|\,X_i\right]\,, &(16)\end{aligned}$$

and we will have a martingale basis function satisfying the condition (C3) if

$$a_{ik} = a_{i+1,k}\,\mathbb{E}\left[\epsilon_\delta^k\,|\,X_i\right]\,. \tag{17}$$

Now in general we have

$$\begin{aligned}\mathbb{E}\left[\epsilon_\delta^k\right] &= e^{kr\delta}\,\mathbb{E}\left[(e^{\sigma W_\delta})^k\right]e^{-k\sigma^2\delta/2} \\ &= e^{kr\delta}\,\mathbb{E}\left[e^{k\sigma\sqrt{\delta}Z}\right]e^{-k\sigma^2\delta/2} \\ &= e^{kr\delta}\,e^{k^2\sigma^2\delta/2-k\sigma^2\delta/2} \\ &= e^{kr\delta}\,e^{k(k-1)\sigma^2\delta/2}\,.\end{aligned}$$

Substituting into (17) we have a martingale if

$$a_{ik} = a_{i+1,k}\,e^{kr\delta}\,e^{k(k-1)\sigma^2\delta/2}\,.$$

Hence, a general formula for martingale basis functions under GBM is

$$\psi_{ik}(X_i) = (X_i)^k\,e^{-\left(kr+k(k-1)\sigma^2/2\right)(t_i-t_0)}\,. \tag{18}$$

The first three basis functions generated using $k = 0, 1, 2$ in this formula coincide with (12)–(14).

European Options. Many studies have found that including the payoff function as a basis function improves the accuracy of regression methods. We look for a general way of constructing martingales from payoff functions that does not require solving the full option pricing problem. Therefore , we fix a payoff function $h(X_i)$ and consider the basis function

$$\psi_{ik}(X_i) = e^{-r(t_m - t_i)}\,\mathbb{E}\left[h(X_m)|X_i\right] ,$$

which is the value at time i of a European option paying $h(X_i)$ at time t_m. This can be used as a basis function satisfying (C3)

$$\begin{aligned}
\mathbb{E}\left[\psi_{i+1}(X_{i+1})|X_i\right] &= e^{-r(t_m - t_{i+1})}\,\mathbb{E}\left[\mathbb{E}\left[h(X_m)|X_{i+1}\right] \mid X_i\right] \\
&= e^{-r(t_m - t_{i+1})}\,e^{-r(t_{i+1} - t_i)}\,\mathbb{E}\left[h(X_m)|X_i\right] \\
&= e^{-r(t_m - t_i)}\,\mathbb{E}\left[h(X_m)|X_i\right] \\
&= \psi_{ik}(X_i)\,.
\end{aligned}$$

5.2. Two Dimensional Martingale Basis Functions. We consider geometric Brownian motions,

$$\frac{dX_\lambda}{X_\lambda} = r\,dt + \sigma_\lambda\,dW_\lambda \quad \lambda = 1, \ldots, d\,,$$

where r is the risk free interest rate, σ_λ is the volatility of returns for asset X_λ. Each W_λ is a standard Brownian motion, and W_λ and W_μ have correlation $\rho_{\lambda\mu}$. The initial asset price is X_0. The asset does not pay any dividends. We can solve each SDE exactly [27] to find the representation

$$X_\lambda(t) = X_\lambda(0)e^{(r-\sigma_\lambda^2/2)t+\sigma_\lambda W(t)}\,.$$

We now consider X_i to be two dimensional, with components x and y, evolving under correlated geometric Brownian motion. Expectations are found by integrating against the joint probability density

$$\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\,\exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right)\right] ,$$

where σ_1 is the volatility of the variable x, σ_2 is the volatility of the variable y, and ρ is the correlation between the two variables. We want to find basis functions that satisfy the martingale property (C3)

$$\mathbb{E}\left[\psi_{i+1}(X_{i+1})|X_i\right] = \psi_i(X_i)\,,$$

in two dimensions. As in the one dimensional case the simplest basis function is a constant. The one dimensional martingale monomials suggested in the previous section remain martingales in this two dimensional setting. We require martingale basis functions that involve products of the two underlying assets. We leave this as future work.

Other options. Various multi-dimensional options can be priced easily analytically and potentially used as martingale basis functions. European options on the maximum or minimum of two asset are priced in [40], and the results extended in [31]. A well known example is an option on the geometric average of a basket of stocks [27]. Some martingales under geometric Brownian motion are given in [26].

6. Low Biased Estimators.

The LSMC estimator has undetermined bias because it both generates a sub-optimal exercise strategy and uses future information (as described in [27]). If we fix a sub-optimal exercise policy generated by the LSMC algorithm and perform a Monte Carlo simulation with a new set of independent simulated paths we have a low biased estimator of the American option value. The low biased estimator for the option value is denoted by $\hat{V}_i^{\downarrow}(x)$, and the high biased estimator by $\hat{V}_i^{\uparrow}(x)$.

Algorithm 3 Low Biased Monte Carlo for American Options — Regression Now

Require: Basis function coefficients $\hat{\beta}_{ik}$ stored from Algorithm 1
  Choose the number of asset paths, N, to simulate, $j = 1, \ldots, N$
  $x_{ij\lambda}$ generate the asset price tensor (e.g. under geometric Brownian motion) from initial data X_0
  for j = 0 to N **do**
    for i = 0 to m **do**
      $\hat{C}_j \leftarrow \hat{\beta}_{ik}\psi_{ik}(x_{ij\lambda})$ the estimated continuation value using stored basis function coefficients
      $h_j \leftarrow h_i(x_{ij\lambda})$ the value of immediate exercise (payoff)
      if $h_j > \hat{C}_j$ **then**
        $\hat{V}_j^{\downarrow} = D_{0,i}\, h_j$ discount the payoff
        break from time loop
      end if
      if $i = m$ **then**
        $\hat{V}_j^{\downarrow} = D_{0,m}\, h_j$ exercise at the final time
      end if
    end for
  end for
  $\hat{V}_0^{\downarrow} \leftarrow \sum \hat{V}_j^{\downarrow}/N$ the low biased estimated option value

Algorithm 3. To produce an estimate for the American option value with definite low bias we use algorithm 1 but on each time step we store the vector of basis function coefficients $\hat{\beta}_{ik}$. We now have a matrix of coefficients β_{ik} for the coefficient of the k^{th} basis function at time step i. This matrix defines a function at each time step i for the estimated continuation value of the option. We define an exercise strategy by exercising the option if the value of immediate exercise is greater than the estimated continuation value.

7. High Biased Estimators.

Many methods have now been suggested to give high biased estimates of the American option value. The simplest is simulating each path and choosing the optimal exercise time. This simple estimate uses future information and is biased high. The stochastic mesh method [15] gives lower and upper bounds which both converge asymptotically to the true value. This method is computationally demanding, but the speed has been improved by using low–discrepancy sequences in [11]. Duality approaches have been suggested in [1, 32, 34, 38], however these methods tend to converge even more computationally expensive than the low biased estimators.

We view

$$\hat{C}_i^{+}(.) = \sum_{k=0}^{K} \hat{\gamma}_{ik}\, \psi_{ik}(.)\,, \qquad i = 1, 2, \ldots, m-1\,,$$

and

$$\tilde{V}_{i+1}^{+}(.) \triangleq \sum_{k=0}^{K} \hat{\gamma}_{ik}\, \psi_{i+1,k}(.)\,, \qquad i = 0, 1, \ldots, m-1\,,$$

as deterministic functions.

Define both $M_0 = 0$ and

$$M_n = \sum_{i=0}^{n-1} \left[\tilde{V}_{i+1}^{+}(X_{i+1}) - \hat{C}_i(X_i)\right]\,, \qquad n = 1, \ldots, m\,. \tag{19}$$

Each summand is

$$\tilde{V}_{i+1}^{+}(X_{i+1}) - \hat{C}_i(X_i) = \sum_{k=0}^{K} \hat{\gamma}_{ik}\, \left[\psi_{i+1,k}(X_{i+1}) - \psi_{ik}(X_i)\right]\,.$$

We now quote Theorem 2 from Glasserman and Yu [29].

Theorem 2 *If (C3) holds then*

$$V_0^{\downarrow}(X_0) = \mathbb{E}\left[h_{\hat{\tau}}(X_{\hat{\tau}})\right] \leq V_0^*(X_0) \leq \mathbb{E}\left[\max_{n=0,1,\ldots,m}(h_n(X_n) - M_n)\right] = V_0^{\uparrow}(X_0)\,.$$

As we assume (C3) and use "regression later" we have all the information we require to calculate (19). However, basis functions satisfying (C3) are not suggested by Glasserman and Yu [29].

Algorithm 4. To produce an estimate for the American option value with definite high bias we first use algorithm 2 and store the vector of martingale basis function coefficients $\hat{\gamma}_k$ at each time step i. This gives a matrix of coefficients $\hat{\gamma}_{ik}$. We define an exercise strategy by comparing the value of immediate exercise to the estimated continuation value. This gives a low biased estimate.

To obtain the high biased estimate we use martingales basis functions as proposed in Glasserman and Yu [29]. As the high biased estimator builds on the the previous algorithms we obtain a LSMC estimate, a low biased estimate, and a high biased estimate from the algorithm.

Algorithm 4 Low and High Biased Monte Carlo for American Options — Regression Later

---

Require: Basis functions $\psi_{ik}(x)$ are martingales satisfying condition (C3)
Require: Martingale basis function coefficients $\hat{\gamma}_{ik}$ stored from Algorithm 2.
  Choose the number of asset paths, N, to simulate, $j = 1, \ldots, N$
  $x_{ij\lambda}$ generate the asset price tensor (e.g. under geometric Brownian motion) from initial data X_0
  for j = 0 to N **do**
    $M \leftarrow 0$, $U \leftarrow 0$
    for i = 0 to m **do**
      $h_j \leftarrow h_i(x_{ij\lambda})$ the value of immediate exercise
      $\hat{C}_j^+ \leftarrow \hat{\gamma}_{ik}\,\psi_{ik}(x_{ij\lambda})$ the estimated continuation value
      $\tilde{V}_j^+ \leftarrow \hat{\gamma}_{ik}\,\psi_{i+1,k}(x_{i+1,j\lambda})$ the estimated future option value
      if $h_j > \hat{C}_j^+$ **then**
        $\hat{V}_j^{\downarrow +} \leftarrow h_j$
        do not repeat
      end if
      $M \leftarrow M + \tilde{V}_j^+ - \hat{C}_j^+$
      $U \leftarrow \max(h_j, M)$
    end for
    $\hat{V}_j^{\uparrow +} \leftarrow U$
  end for
  $V_0^{\downarrow +} \leftarrow \sum \hat{V}_j^{\downarrow +}/N$ low biased estimator
  $V_0^{\uparrow +} \leftarrow \sum \hat{V}_j^{\uparrow +}/N$ high biased estimator

---

8. Numerical Results.

We implemented algorithms 1–4 in C++ in the derivatives pricing library QuantLib, discussed below. A single processor x86 machine was used to run the simulations.

We ran simulations using algorithms 1 and 2 to test the practical impact of Theorem 1 in Glasserman and Yu [29]. This says that 'regression later' will produce a better fit and less variable estimates of coefficients than 'regression now', which should translate into a more accurate estimate for the option value.

To test the accuracy of using martingale basis functions to produce upper bounds to the American option price, as suggested in Glasserman and Yu [29], we used algorithm 4.

8.1. QuantLib. QuantLib is a derivatives pricing library implemented in C++ . As the project website [36] says,

X_0	σ	T	Finite difference American	Closed form European	Longstaff Schwartz estimates	(s.e)	Regression now	(s.e)	Regression later	(s.e)
36	0.2	1	4.478	3.844	4.472	(.010)	4.467	(.009)	4.466	(.009)
36	0.2	2	4.840	3.763	4.821	(.012)	4.835	(.011)	4.839	(.011)
36	0.4	1	7.101	6.711	7.091	(.020)	7.104	(.019)	7.100	(.019)
36	0.4	2	8.508	7.700	8.488	(.024)	8.495	(.023)	8.495	(.023)

Table 1: Comparison of 'regression now' and 'regression later' with the results for American style put option in Table 1 in Longstaff and Schwartz [33]. X_0 is the initial asset price, σ is the volatility of returns, and T is the number of years until the option expiry date. The continuously compounded short–term interest rate is 0.06, and the strike price of all the put options is 40. 'Regression now' uses a constant and the first three Laguerre polynomials. 'Regression later' uses a constant and the first three martingale basis functions under geometric Brownian motion. The estimates are made using 100,000 simulations (50,000 and 50,000 antithetic). (s.e) denotes the standard error of the value to the left. Only in-the-money paths were used in the regression.

> The QuantLib project is aimed at providing a comprehensive software framework for quantitative finance. QuantLib is a free/open–source library for modeling, trading, and risk management in real-life.

As code is available on the internet it will be verified, tested, and adopted more quickly than academic papers alone. Academics researching in computational areas such as mathematical finance can greatly increase the impact of their work if both their papers and software implementations are available on the internet. For a survey of open–source derivatives pricing libraries and the benefits of adopting QuantLib see [23]. For instructions on how to download, install, run and contribute see the project website [36].

8.2. Regression Precision. Under certain conditions on the residuals, 'regression later' gives an estimator with less variance than 'regression now' (Theorem 1 Glasserman and Yu [29]). In this section we test this theoretical result numerically.

First we check that 'regression later' provides good estimates for the option value. The results in table 1 compare 'regression now' with 'regression later', following the results in table 1 in Longstaff and Schwartz [33]. Only in-the-money paths are used in the regression [33]. Both regressions use the first three martingale basis functions described above. We find that 'regression later' does indeed provide a good estimate for the option value.

Next we compare the accuracy of 'regression now' to 'regression later'. We ran the simulations 50 times with 1000 paths for each simulation. Table 2 show the results. We find that the estimates of the option value using 'regression later' and on average three times more accurate than estimates using 'regression now'. This is very promising and we propose to test these results over a wider range of parameter values.

8.3. Lower and Upper Bounds. Preliminary results using algorithm 4 to obtain lower and upper bounds from Theorem 2 show that care has to be take choosing martingale basis functions. While it is relatively easy to obtain reasonable lower bounds we were unable to obtain enough repeatable results to make any conclusions about the speed or quality upper bounds.

9. Evaluation and Future Directions.

The numerical results show that regression estimates using 'regression later' are more accurate than existing methods using 'regression now', such as Longstaff and Schwartz [33]. However, the 'regression later' method depends, through condition (C3), on the availability of basis functions that are martingales. We have obtained results for single asset options on underlying assets driven by geometric Brownian motion. It may not be so easy to find martingale basis functions for other stochastic processes.

			Finite	Regression now			Regression later		
			difference		absolute	relative		absolute	relative
X_0	σ	T	American	value	error	error (%)	value	error	error (%)
36	0.2	1	4.478	4.525	.047	1.050	4.496	.018	.402
36	0.2	2	4.840	4.888	.048	0.992	4.860	.020	.413
36	0.4	1	7.101	7.240	.139	1.957	7.075	.026	.366
36	0.4	2	8.508	8.629	.121	1.422	8.579	.071	.835

Table 2: We use the same parameters as in table 1, but run 1000 samples (500 and 500 antithetic) 50 times and average the estimated option value to assess the accuracy. Both regressions use a constant and the first three martingale basis functions under geometric Brownian motion. X_0 is the initial asset price, σ is the volatility of returns, and T is the number of years until the option expiry date. The continuously compounded short–term interest rate is 0.06, and the strike price of all the put options is 40. The relative error is calculated by dividing the absolute error by the true option price. Only in-the-money paths were used in the regression.

9.1. Future Directions. It would be prudent to perform a more thorough test of the 'regression later' method using different martingale basis functions, including the payoff and European option values. We propose to extend the numerical implementation to test the multi–asset basis functions suggested.

It is also necessary to investigate basis functions that are martingales for assets driven by stochastic processes other than geometric Brownian motion.

A systematic comparison of the duality methods of Andersen and Broadie [1], Glasserman and Yu [29], Kogan and Haugh [32], Rogers [38], using QuantLib for the implementation, would allow an informed evaluation of their relative merits to be made.

Andersen and Broadie [1] produces an upper bound by generating an estimate for the duality gap, and adding that to an estimate of the lower bound. As their method attempts to estimate a value that should be small it may be possible to find an estimator with less variance than that in Theorem 2 and hence find a tighter upper bound.

Numerical tests on the utility of quasi Monte Carlo in the computation of upper bounds could be made relatively easily using the low–discrepancy sequences implemented in QuantLib.

10. Conclusions.

We have built on the contribution of Glasserman and Yu [29] by presenting martingale basis functions satisfying the condition (C3) under geometric Brownian motion for a single asset. We have also suggested specific martingale basis functions that could be used to extend this work numerically to options on more underlying assets.

Theorem 1 in Glasserman and Yu [29] suggests that 'regression later' should produce more accurate estimates than 'regression now'. This is is supported by our numerical results, in which 'regression later' is on average three times more accurate.

The use of martingale basis functions and duality in regression methods for pricing American options enables the fast calculation of upper bounds. However, we are not able to make any conclusions without further numerical investigation.

References

[1] L. Andersen and M. Broadie. A primal-dual simulation algorithm for pricing multi-dimensional American options. Technical report, Columbia University, 2001.

[2] A. N. Avramidis and P. Hyden. Efficiency improvements for pricing American options with a stochastic mesh. In P. A. Farrington, H. B. Nembhard, D. T. Sturrock, and G. W. Evans, editors, *Proceedings of the 1999 Winter Simulation Conference*, pages 344–350, 1999.

[3] V. Bally, G. Pagès, and J. Printems. A quantization tree method for pricing and hedging multi–dimensional American options. Technical report, INRIA, 2002.

[4] J. Barraquand and D. Martineau. Numerical valuation of high dimensional multivariate American securities. *Journal of Financial and Quantitative Analysis*, 30(3):383–405, 1995.

[5] S. Berridge and H. Schumacher. An irregular grid method for solving high-dimensional problems in finance. In Sloot et al, editor, *Computational Science - ICCS 2002, Part II*. Springer-Verlag, 2002.

[6] P. Bossaerts. Simulation estimators of optimal early exercise. Technical report, Carnegie–Mellon University, 1989.

[7] P. P. Boyle. Options: A Monte Carlo approach. *Journal of Financial Economics*, 4:323–338, 1977.

[8] P. P. Boyle. A lattice framework for option pricing with two state variables. *Journal of Financial and Quantitative Analysis*, 23:1–12, 1988.

[9] P. P. Boyle, M. Broadie, and P. Glasserman. Monte Carlo methods for security pricing. *Journal of Economic Dynamics and Control*, 21:1267–1321, 1997.

[10] P. P. Boyle, J. Evine, and S. Gibbs. Numerical evaluation of multivariate contingent claims. *Review of Financial Studies*, 2:241–250, 1989.

[11] P. P. Boyle, A. Kolkiewicz, and K. S. Tan. Pricing American–style options using low discrepancy mesh methods. Technical report, University of Waterloo, 2000.

[12] M. J. Brennan and E. S. Schwartz. The valuation of the American put option. *Journal of Finance*, 32:449–462, 1977.

[13] M. Broadie and J. Detemple. The valuation of American options on multiple assets. *Mathematical Finance*, 7(3):241–286, 1997.

[14] M. Broadie and P. Glasserman. Pricing American-style securities using simulation. *Journal of Economic Dynamics and Control*, 21:1323–1352, 1997.

[15] M. Broadie and P. Glasserman. A stochastic mesh method for pricing high-dimensional American options. Technical report, Columbia University, 1997.

[16] M. Broadie, P. Glasserman, and Z. Ha. Pricing American options by simulation using a stochastic mesh with optimized weights. In S. P. Uryasev, editor, *Probabilistic Constrained Optimization: Methodology and Applications*, pages 32–50. Kluwer, 2000.

[17] P. Carr, R. Jarrow, and R. Myneni. Alternative characterizations of American put options. *Journal of Mathematical Finance*, 2:87–106, 1992.

[18] J. Carriere. Valuation of early–exercise price of options using simulations and nonparametric regression. *Insurance: Mathematics and Economics*, 19:19–30, 1996.

[19] E. Clément, D. Lamberton, and P. Protter. An analysis of a least squares regression algorithm for American option pricing. *Finance and Stochastics*, 6(4):449–471, 2001.

[20] J. Cox, S. Ross, and M. Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7:229–263, 1979.

[21] M. A. H. Dempster, A. Eswaran, and D. G. Richards. Wavelet methods in PDE valuation of financial derivatives. In K. Leung, L. W. Chang, and H. Meng, editors, *Proceedings of the Second International Conference of Intelligent Data Engineering and Automated Learning*, pages 215–238, 2000.

[22] M. A. H. Dempster and E. Eswaran. Solution of PDEs by wavelet methods. Technical report, University of Cambridge, 2001.

[23] N. Firth. Why use QuantLib? URL `http://www.maths.ox.ac.uk/~firth/research/quantlib.pdf`. 2004.

[24] M. Fu and J. Q. Hu. Sensitivity analysis for Monte Carlo simulation of option pricing. *Probability in the Engineering and Information Sciences*, 49:417–446, 1995.

[25] M. Fu, S. Laprise, D. B. Madan, Y. Su, and R. Wu. Pricing American options: A comparison of Monte Carlo simulation approaches. *Journal of Computational Finance*, 4, 2001.

[26] H. U. Gerber and E. S. W. Shiu. Martingale approach to pricing perpetual American options on two stocks. *Mathematical Finance*, 6:303–322, 1996.

[27] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, 2004.

[28] P. Glasserman and B. Yu. Number of paths versus number of basis functions in American option pricing. *to appear in Annals of Applied Probability*, 2003.

[29] P. Glasserman and B. Yu. Pricing American options by simulation: Regression now or regression later? In H. Niederreiter, editor, *Monte Carlo and Quasi-Monte Carlo Methods 2002*, 2004.

[30] A. Ibáñez and F. Zapatero. Monte Carlo valuation of American options through computation of the optimal exercise frontier. Technical report, Marshall School of Business - USC, 2001.

[31] H. Johnson. Options on the maximum or the minimum of several assets. *Journal of Financial and Quantitative Analysis*, 22(3):277–283–, 1987.

[32] L. Kogan and M. B. Haugh. Pricing American options: A duality approach. *Operations Research*, 52(2), 2004.

[33] F. A. Longstaff and E. S. Schwartz. Valuing American options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 14(1):113–147, Spring 2001.

[34] N. Meinshausen and B. Hambly. Monte Carlo methods for multiple exercise problems. *to appear in Mathematical Finance*, 2004.

[35] R. Myneni. The pricing of the American option. *The Annals of Applied Probability*, 2(1):1–23, 1992.

[36] QuantLib. QuantLib website, 2004. URL `http://www.quantlib.org/`.

[37] C. Reisinger. *Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben*. PhD thesis, University of Heidelberg, Germany, 2004.

[38] L. C. G. Rogers. Monte Carlo valuation of American options. *Mathematical Finance*, 12(3), July 2002.

[39] L. Stentoft. Assessing the least squares Monte Carlo approach to American option valuation. Technical report, University of Aarhus, 2003.

[40] R. M. Stultz. Options on the minimum or the maximum of two risky assets. *Journal of Financial Economics*, 10:161–185, 1982.

[41] D. Tavella. *Quantitative Methods in Derivatives Pricing: An Introduction to Computational Finance*. J. Wiley, 2002.

[42] J. A. Tilley. Valuing American options in a path simulation model. *Transactions of the Society of Actuaries*, 45:83–104, 1993.

[43] J. N. Tsitsiklis and B. Van Roy. Regression methods for pricing complex American–style options. *IEEE Transactions on Neural Networks*, 12(4):694–703, 2001.