

COMPARING OPTIMAL CONVERGENCE RATE OF STOCHASTIC MESH AND LEAST SQUARES METHOD FOR BERMUDAN OPTION PRICING

Ankush Agarwal
Sandeep Juneja

Tata Institute of Fundamental Research
Homi Bhabha Road, Colaba
Mumbai, MA 400005, INDIA

ABSTRACT

We analyze the stochastic mesh method (SMM) as well as the least squares method (LSM) commonly used for pricing Bermudan options using the standard two phase methodology. For both the methods, we determine the decay rate of mean square error of the estimator as a function of the computational budget allocated to the two phases and ascertain the order of the optimal allocation in these phases. We conclude that with increasing computational budget, while SMM estimator converges at a slower rate compared to LSM estimator, it converges to the true option value whereas LSM estimator, with fixed number of basis functions, usually converges to a biased value.

1 INTRODUCTION

The pricing of Bermudan options is a well studied problem in computational finance and several lines of research have been proposed to tackle this problem. Andersen (1999) parameterized stopping rules or exercise regions to reduce the optimal stopping problem to a much more tractable finite dimension optimization problem. Broadie and Glasserman (1997a) introduced nested simulation where they proposed upper and lower biased estimators to obtain valid confidence intervals for Bermudan option prices. Another technique of Bermudan option pricing is through the use of dual formulation of optimal stopping problem. Andersen and Broadie (2004), Haugh and Kogan (2004) and Rogers (2002) used this formulation to obtain both upper and lower bounds for Bermudan option price.

Two of the most commonly used approaches in practice are the stochastic mesh method (SMM) proposed by Broadie and Glasserman (1997b) and the least squares method (LSM) proposed by Carriere (1996). In the latter method, value function is modeled as a linear combination of chosen basis functions and the problem is reduced to finding optimal coefficients of the linear combination. Tsitsiklis and Van Roy (1999), Tsitsiklis and Van Roy (2001) and Longstaff and Schwartz (2001) further developed the theory and computational enhancements to this method.

A popular implementation for both these methods involves two phases: Phase 1 involves generation of, say, M independent sample paths of the underlying process and a specific backward recursion using these paths to ascertain an approximately optimal exercise policy over the state space of the underlying process. In Phase 2, N independent sample paths are generated and the policy determined in Phase 1 of the algorithm is evaluated using standard Monte Carlo. As is well known, the resulting estimator is lower biased as the price of a Bermudan option is the maximum over all possible exercise policies while Phase 1 provides one such policy.

In this paper, we analyze the optimal convergence rate of the resulting estimator from these two phases both for SMM and LSM. Letting Γ denote the computational budget available, we study these methods as $\Gamma \rightarrow \infty$. For SMM, the asymptotic bias is given as $\Theta(\frac{1}{M})$ and the asymptotic variance as $\Theta(\frac{1}{N})$ (for

given functions $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ and $g : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, we write $f(x) = \Theta(g(x))$ if there exists $c_1 > 0, c_2 > 0$ and x_1 large enough such that $c_1 g(x) \leq f(x) \leq c_2 g(x)$ for all $x > x_1$. We observe that to minimize the mean square error of the resultant estimator, we need to set $M = \Theta(\Gamma^{1/3})$ and $N = \Theta(\Gamma^{2/3})$. This optimal allocation results in mean square error decaying at the rate $\Theta(\Gamma^{-2/3})$. Through a numerical example, we illustrate that selecting M and N optimally can result in significant improvement in mean square error of the estimator for a given computational budget.

For LSM, we observe that once we fix the number of basis functions and assume that the true continuation value function is given as a linear combination of these fixed basis functions, the asymptotic bias of the option price estimator is given as $\Theta(\frac{1}{M})$ and the asymptotic variance as $\Theta(\frac{1}{N})$. In order to minimize the mean square error of the resulting estimator, we need $M = \Theta(\Gamma^{2/3})$ and $N = \Theta(\Gamma)$. This results in mean square error decaying at the rate $\Theta(\Gamma^{-1})$. Thus, while the mean square error in this case decays at a substantially faster rate compared to the mean square error for SMM, the drawback of the LSM approach with a fixed number of basis functions is that the resulting bias converges to a fixed negative value even as $\Gamma \rightarrow \infty$. We demonstrate this numerically via a simple example.

Thus, our contribution is two fold: We provide a nuanced comparison of the popular SMM and LSM methods illustrating the trade-offs in selecting one over the other. We also discuss how the computational budget should be optimally allocated in a standard two phase implementation of these methods.

The remainder of this paper is organized as follows. Section 2 provides a brief introduction to the Bermudan option pricing problem and the stochastic mesh method used to calculate lower biased estimator. Section 3 undertakes the analysis of convergence rate of asymptotic bias and variance of stochastic mesh estimator in two time-period Bermudan option case. Section 4 discusses results for the multiple time-period case and Section 5 presents optimal mean square error analysis. In section 6, we show similar results for the least squares method and perform the optimal mean square error analysis. The numerical examples conducted to validate our theoretical results are presented in Section 7. We offer conclusions in Section 8.

Due to the space constraints, the analysis in a few sections of this paper is at a heuristic level. A rigorous version in a more general framework will appear separately.

2 PROBLEM FORMULATION

We consider an \mathbb{R}^d -valued Markov process $\{X_t : t \in \{0, \dots, T\}\}$ which is defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \{0, \dots, T\}}, P)$ where $\mathcal{F} = \{\mathcal{F}_t, 0 \leq t \leq T\}$ is the natural filtration augmented by the \mathbb{P} -null sets of \mathcal{F} . It is assumed that X is adapted to $(\mathcal{F}_t)_{t \in \{0, \dots, T\}}$. In this setting, we consider a Bermudan option which pays $g_t(X_t)$ for known positive functions g_0, \dots, g_T when exercised at $t = 0, \dots, T$. Under the existence of risk-neutral probability measure, the price of this option is given as the solution to the following optimal stopping problem

$$V_t(x) = \sup_{\tau \in \mathcal{T}_t} \mathbb{E}[g_\tau(X_\tau) | X_t = x], x \in \mathbb{R}^d$$

where \mathcal{T}_t is the set of stopping times taking values in $\{t, \dots, T\}$ for $t = 0, \dots, T$.

According to the principle of dynamic programming, the option price $(V_t)_{t \in \{0, \dots, T\}}$ satisfies the following backward recursion

$$V_T(x) = g_T(x) \tag{2.1}$$

$$V_t(x) = \max\{g_t(x), C_t(x)\}, t = 0, \dots, T - 1. \tag{2.2}$$

where $C_t(x) := \mathbb{E}[V_{t+1}(X_{t+1}) | X_t = x]$ is known as the continuation value of the option at time t and the optimal stopping time can be shown to equal $\tau^* := \min\{t \in \{0, 1, \dots, T\} : g_t(X_t) \geq C_t(X_t)\}$.

In multidimensional optimal stopping problems, exact calculation of (2.1)-(2.2) is typically infeasible due to the difficulty of determining analytical form of conditional expectations. Hence, approximate dynamic programming methods are used which involve estimating the continuation value function at each exercise opportunity.

2.1 Stochastic Mesh Method

Broadie and Glasserman (1997b) proposed stochastic mesh method to perform the task of estimating continuation value function. Here, we explain the two phase algorithm to calculate the lower biased estimator.

In Phase 1, we generate M independent paths of the underlying Markov process X starting from x_0 at time 0, “forget” the path to which each node at time t belongs and interconnect all nodes at consecutive time steps. Under the risk-neutral measure P , conditional on $X_t = x$, density of X_{t+1} is denoted by $f_t(x, \cdot)$ and let $f_t(\cdot)$ denote the marginal density of X_{t+1} (with X_0 fixed). Then, the estimate of the continuation value function $\hat{C}_{t,M}(x)$ at exercise opportunity t is defined as

$$\hat{C}_{t,M}(x) := \frac{1}{M} \sum_{j=1}^M w_t^{(j)} \hat{V}_{t+1}(X_{t+1}^{(j)}),$$

where $w_t^{(j)} = \frac{f_t(x, X_{t+1}^{(j)})}{q_{t+1}(X_{t+1}^{(j)})}$, with $q_{t+1}(X_{t+1}^{(j)}) = \frac{1}{M} \sum_{i=1}^M f_t(X_t^{(i)}, X_{t+1}^{(j)})$ and $\hat{V}_t(X_t^{(j)}) = \max\{g_t(X_t^{(j)}), \hat{C}_t(X_t^{(j)})\}$. This choice of likelihood ratio not only provides variance reduction, as the ratio is always bounded by M , but at the same time allows correct pricing moving backwards through the mesh if the value function is known exactly. Then, for any path of the underlying process X , the estimated suboptimal stopping policy is given as $\hat{\tau} := \min\{1 \leq k \leq T : g_k(X_k) \geq \hat{C}_{k,M}(X_k)\}$.

In Phase 2, we generate N independent sample paths of the underlying Markov process X and evaluate the approximately optimal stopping policy on each of these sample paths to obtain the lower biased estimator. The option price estimator is given as $\hat{V}_{0,M}(x_0) = \frac{1}{N} \sum_{i=1}^N g_{\hat{\tau}_i}(X_{\hat{\tau}_i}^{(i)})$ where the suboptimal stopping policy on each Phase 2 sample path is defined as $\hat{\tau}_i := \min\{1 \leq k \leq T : g_k(X_k^{(i)}) \geq \hat{C}_{k,M}(X_k^{(i)})\}$. Before proving the results on lower biased option price estimator, we make the following assumptions which will be in force throughout. Let us assume:

Assumption 1 There exist constants $L_1, L_2 > 0$ such that

$$L_1 \delta \leq P(0 < |C_t(X_t) - g_t(X_t)| \leq \delta) \leq L_2 \delta, \quad \text{for } t \in \{0, \dots, T-1\}.$$

The above assumption is satisfied if the random variable $C_t(X_t) - g_t(X_t)$ has a bounded and positive density in a neighborhood that includes zero.

Assumption 2 There exists a constant $D > 0$ such that $0 < \frac{f_t(X_t, X_{t+1}^{(l)})}{f_t(X_t^{(l)}, X_{t+1}^{(l)})} \leq D < \infty$, $l = 1, 2$, for $t \in \{1, \dots, T\}$.

The above assumption may seem restrictive. But, we can find a compact set such that the underlying diffusion process when restricted to the set, is “close” to the unrestricted process. By continuity of the probability density functions we can then conclude that the ratio remains bounded over the compact set. More discussion on the compact set argument can be found in (Belomestny 2011).

Assumption 3 There exists constant $B_0 > 0$ such that for all $x \in \mathbb{R}^d$, $g_t(x) \leq B_0$ for $t \in \{0, 1, \dots, T\}$.

Various Bermudan-style securities, such as the vanilla put and Bermudan max call option, have payoff function which are continuous. If the underlying process is restricted to a compact set, Assumption 3 is satisfied. For the sake of presentation simplicity, we use the above form of condition on the payoff functions.

Assumption 4 For any continuous function $h : \mathbb{R}^d \rightarrow \mathbb{R}_+$ and independent normal random variable Z , there exist constants U_1, U_2 such that for any $\delta > 0$

$$U_1 \delta \leq \mathbb{E} \left[|C_t(X_t) - g_t(X_t)| \mathbb{1}_{\{0 < |C_t(X_t) - g_t(X_t)| \leq \delta h(X_t) Z\}} \right] \leq U_2 \delta$$

for $t \in \{0, \dots, T - 1\}$.

Given $h(\cdot)$ is a continuous function, it remains bounded when the underlying process X_t is restricted to a compact set. Then, the above assumption is satisfied if the random variable $C_t(X_t) - g_t(X_t)$ has a bounded positive density in the neighborhood which includes zero.

3 TWO TIME-PERIOD PROBLEM

The two time-period horizon case corresponds to exercise opportunities at $t \in \{0, 1, 2\}$. We assume without loss of generality that it is not optimal to exercise the Bermudan option at $t = 0$. Therefore, we need to evaluate the continuation value estimate only at $t = 1$. From our knowledge of the dynamic programming principle, at the expiry, $\hat{V}_2(x) = g_2(x)$ for all $x \in \mathbb{R}^d$. Then, we have

$$\hat{C}_{1,M}(x) = \frac{1}{M} \sum_{j=1}^M \frac{f_1(x, X_2^{(j)})}{q_2(X_2^{(j)})} g_2(X_2^{(j)}).$$

For a sample path (X_0, X_1, X_2) , let us denote the optimal stopping time by τ and the suboptimal stopping policy based on stochastic mesh continuation value estimates by $\hat{\tau}$. Then, bias of the stochastic mesh estimator $\hat{V}_{0,M}(x_0)$ is given as

$$\begin{aligned} & V_0(x_0) - \mathbb{E}[\hat{V}_{0,M}(x_0)] \\ &= \mathbb{E}[g_1(X_1)1_{\{\tau=1\}} + g_2(X_2)1_{\{\tau=2\}}] - \mathbb{E}[g_1(X_1)1_{\{\hat{\tau}=1\}} + g_2(X_2)1_{\{\hat{\tau}=2\}}] \\ &= \mathbb{E}[(g_2(X_2) - g_1(X_1))1_{\{\tau=2, \hat{\tau}=1\}} + (g_1(X_1) - g_2(X_2))1_{\{\tau=1, \hat{\tau}=2\}}] \\ &= \mathbb{E}[(C_1(X_1) - g_1(X_1))1_{\{\hat{C}_{1,M}(X_1) \leq g_1(X_1) < C_1(X_1)\}}] + \mathbb{E}[(g_1(X_1) - C_1(X_1))1_{\{C_1(X_1) \leq g_1(X_1) < \hat{C}_{1,M}(X_1)\}}]. \end{aligned} \tag{3.1}$$

From (3.1) we see that the key to the analysis of convergence of the bias of $\hat{V}_{0,M}(x_0)$ lies in the bounds of error probability of the continuation value estimator $\hat{C}_{1,M}(\cdot)$. We now state the result for asymptotic bias.

Proposition 1 Suppose that Assumption 1 - 4 hold. If $M \rightarrow \infty$, then there exist constants γ_1, γ_2 such that

$$\frac{\gamma_1}{M} \leq V_0(x_0) - \mathbb{E}[\hat{V}_{0,M}(x_0)] \leq \frac{\gamma_2}{M}.$$

This result shows that in the two time-period case, bias of the option price estimator decays as $\Theta(\frac{1}{M})$. The proof of Proposition 1 uses the following lemma.

Lemma 1 Suppose that Assumption 2 - 3 hold. Then,

$$\frac{1}{\sqrt{M}} \sum_{j=1}^M \left(\frac{f_1(x, X_2^{(j)})}{\frac{1}{M} \sum_{i=1}^M f_1(X_1^{(i)}, X_2^{(j)})} g_2(X_2^{(j)}) - C_1(x) \right) \xrightarrow{\mathcal{D}} \sigma(x) \mathcal{N}(0, 1) \text{ as } M \rightarrow \infty,$$

where $\sigma(x) = \mathbb{E} \left[\left(\frac{f_1(x, X_2^{(1)})}{f_1(X_2^{(1)})} g_2(X_2^{(1)}) - C_1(x) \right)^2 \right]$ and $\mathcal{N}(0, 1)$ denotes the standard normal distribution.

Proof. We note the following decomposition:

$$\begin{aligned} & \frac{1}{\sqrt{M}} \sum_{j=1}^M \left(\frac{f_1(x, X_2^{(j)})}{\frac{1}{M} \sum_{i=1}^M f_1(X_1^{(i)}, X_2^{(j)})} g_2(X_2^{(j)}) - C_1(x) \right) \\ &= \frac{1}{\sqrt{M}} \sum_{j=1}^M \left(\frac{f_1(x, X_2^{(j)})}{f_1(X_2^{(j)})} g_2(X_2^{(j)}) - C_1(x) \right) \\ &+ \frac{1}{\sqrt{M}} \sum_{j=1}^M \left(\left(\frac{f_1(x, X_2^{(j)})}{\frac{1}{M} \sum_{i=1}^M f_1(X_1^{(i)}, X_2^{(j)})} - \frac{f_1(x, X_2^{(j)})}{f_1(X_2^{(j)})} \right) g_2(X_2^{(j)}) \right). \end{aligned}$$

From the classical Central Limit Theorem, we know that $\frac{1}{\sqrt{M}} \sum_{j=1}^M \left(\frac{f_1(x, X_2^{(j)})}{f_1(X_2^{(j)})} g_2(X_2^{(j)}) - C_1(x) \right)$ converges in distribution to a Gaussian random variable with mean 0 and variance $\sigma(x)$. Moreover, we have by Strong Law of Large Numbers for all $y \in \mathbb{R}^d$

$$\frac{1}{M} \sum_{i=1}^M f_1(X_1^{(i)}, y) \rightarrow f_1(y) \text{ a.s.} \tag{3.2}$$

where recall $f_1(\cdot)$ denotes the marginal density function of X_2 (with X_0 fixed). From Assumption 3 and (3.2), we can show that

$$\frac{1}{\sqrt{M}} \sum_{j=1}^M \left(\left(\frac{f_1(x, X_2^{(j)})}{\frac{1}{M} \sum_{i=1}^M f_1(X_1^{(i)}, X_2^{(j)})} - \frac{f_1(x, X_2^{(j)})}{f_1(X_2^{(j)})} \right) g_2(X_2^{(j)}) \right)$$

converges to 0 in L^2 . Next, appealing to Slutsky's theorem (pg. 268 (Resnick 1999)), which states that if $X_n \Rightarrow X$ and $\xi_n \xrightarrow{P} 0$, then $X_n + \xi_n \Rightarrow X$, completes the proof. \square

Proof of Proposition 1. We calculate the bias of estimator $\hat{V}_{0,M}(x_0)$ based on (3.1). Note the decomposition:

$$\begin{aligned} & \mathbb{E}[(C_1(X_1) - g_1(X_1)) \mathbf{1}_{\{\hat{C}_{1,M}(X_1) \leq g_1(X_1) < C_1(X_1)\}}] \\ &= \mathbb{E}[(C_1(X_1) - g_1(X_1)) \mathbf{1}_{\{C_1(X_1) - g_1(X_1) > 0\}} \mathbf{1}_{\{C_1(X_1) - \hat{C}_{1,M}(X_1) \geq C_1(X_1) - g_1(X_1)\}}] \\ &= \mathbb{E}[(C_1(X_1) - g_1(X_1)) \mathbf{1}_{\{C_1(X_1) - g_1(X_1) > 0\}} \\ &\times \left(\mathbb{P}\left(\sqrt{M}(C_1(X_1) - \hat{C}_{1,M}(X_1)) \geq \sqrt{M}(C_1(X_1) - g_1(X_1)) \mid X_1\right) - \mathbb{P}\left(Y\sigma(X_1) \geq \sqrt{M}(C_1(X_1) - g_1(X_1)) \mid X_1\right) \right)] \\ &+ \mathbb{E}[(C_1(X_1) - g_1(X_1)) \mathbf{1}_{\{C_1(X_1) - g_1(X_1) > 0\}} \mathbb{P}\left(Y\sigma(X_1) \geq \sqrt{M}(C_1(X_1) - g_1(X_1)) \mid X_1\right)] \end{aligned} \tag{3.3}$$

where Y is an independent standard normal random variable and $\sigma(x) = \mathbb{E} \left[\left(\frac{f_1(x, X_2^{(1)})}{f_1(X_2^{(1)})} g_2(X_2^{(1)}) - C_1(x) \right)^2 \right]$ is assumed to be a continuous function. For the second term in the right hand side of (3.3), we note that:

$$\begin{aligned} & \mathbb{E} \left[(C_1(X_1) - g_1(X_1)) \mathbf{1}_{\{C_1(X_1) - g_1(X_1) > 0\}} \mathbb{P}\left(Y\sigma(X_1) \geq \sqrt{M}(C_1(X_1) - g_1(X_1)) \mid X_1\right) \right] \\ &= \mathbb{E} \left[(C_1(X_1) - g_1(X_1)) \mathbf{1}_{\left\{0 < \frac{C_1(X_1) - g_1(X_1)}{\sigma(X_1)} \leq \frac{Y}{\sqrt{M}}\right\}} \mathbb{P}\left(0 < \frac{C_1(X_1) - g_1(X_1)}{\sigma(X_1)} \leq \frac{Y}{\sqrt{M}}\right) \right]. \end{aligned}$$

Then, from Assumption 1 it follows that

$$\frac{L_1}{\sqrt{M}} \leq \mathbb{P}\left(0 < \frac{C_1(X_1) - g_1(X_1)}{\sigma(X_1)} \leq \frac{Y}{\sqrt{M}}\right) \leq \frac{L_2}{\sqrt{M}}$$

for some constants L_1, L_2 . Further, from Assumption 4, we get

$$\frac{U_1}{\sqrt{M}} \leq \mathbb{E} \left[(C_1(X_1) - g_1(X_1)) \mathbf{1}_{\left\{0 < \frac{C_1(X_1) - g_1(X_1)}{\sigma(X_1)} \leq \frac{Y}{\sqrt{M}}\right\}} \right] \leq \frac{U_2}{\sqrt{M}}$$

for some constants U_1, U_2 .

To handle the first term in the right hand side of (3.3), we note from Assumption 2 - 3 that $\mathbb{E} \left[\left(\frac{f_1(x, X_2^{(1)})}{\frac{1}{M} \sum_{i=1}^M f_1(X_1^{(i)}, X_2^{(1)})} g_2(X_2^{(1)}) \right)^3 \right] < \infty$. We know for i.i.d. random variables X_1, \dots, X_n with $EX_1 =$

$0, \text{Var}(X_1) = 1, \gamma = EX_1^3$, the distribution of the standardized sum $S_n = \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j$ admits a first-order Edgeworth expansion (pg. 541 (Feller 2008)) uniformly for all $x \in \mathbb{R}$, which gives

$$P(S_n \leq x) = \Phi(x) - \phi(x) \left(\frac{\gamma(x^2 - 1)}{6\sqrt{n}} + O\left(\frac{1}{n}\right) \right)$$

where $\phi(\cdot)$ denotes the standard normal density function and $\Phi(\cdot)$ denotes the standard normal distribution function. Therefore, from Lemma 1 it implies we can apply the Edgeworth expansion to show

$$\begin{aligned} &M \left(\mathbb{P} \left(\sqrt{M}(C_1(X_1) - \hat{C}_{1,M}(X_1)) \geq \sqrt{M}(C_1(X_1) - g_1(X_1)) \mid X_1 \right) \right. \\ &\left. - \mathbb{P} \left(Y \sigma(X_1) \geq \sqrt{M}(C_1(X_1) - g_1(X_1)) \mid X_1 \right) \right) \rightarrow 0 \end{aligned}$$

uniformly for $X_1 \in \mathbb{R}^d$. From Assumption 3, it follows that appealing to the Dominated Convergence theorem establishes the order relationship for the first term of (3.1). We can repeat the same arguments for the other term in the right hand side of (3.1) to complete the proof. \square

The following result holds for asymptotic variance of the stochastic mesh estimator.

Proposition 2 Suppose that Assumption 1 - 4 hold. Then, there exists constants $\bar{K}, \sigma > 0$ such that for large enough M and N ,

$$\frac{\sigma}{N} \leq \text{Var}(\hat{V}_{0,M}(x_0)) \leq \frac{\sigma}{N} + \frac{\bar{K}}{M^2}.$$

The variance of estimator $\text{Var}(\hat{V}_{0,M}(x_0))$ can be expressed as

$$\text{Var}(\hat{V}_{0,M}(x_0)) = \frac{1}{N} \text{Var}(g_{\hat{\tau}_1}(X_{\hat{\tau}_1}^{(1)})) + \left(1 - \frac{1}{N}\right) \text{Cov}(g_{\hat{\tau}_i}(X_i^{(i)}), g_{\hat{\tau}_j}(X_{\hat{\tau}_j}^{(j)})). \quad (3.4)$$

It can be shown using Dominated Convergence theorem that $\text{Var}(g_{\hat{\tau}}(X_{\hat{\tau}})) \rightarrow \text{Var}(g_{\tau}(X_{\tau}))$. Hence, the first term in the right hand side of (3.4) contributes $\Theta\left(\frac{1}{N}\right)$ and other smaller order terms to overall variance of the estimator. To handle the covariance term $\text{Cov}(g_{\hat{\tau}_i}(X_i^{(i)}), g_{\hat{\tau}_j}(X_{\hat{\tau}_j}^{(j)}))$, we show that it can be written as the sum of a term which is upper bounded by squared-bias and another term given as $\mathbb{E}[(g_{\tau_i}(X_{\tau_i}^{(i)}) - g_{\hat{\tau}_i}(X_{\hat{\tau}_i}^{(i)}))(g_{\tau_j}(X_{\tau_j}^{(j)}) - g_{\hat{\tau}_j}(X_{\hat{\tau}_j}^{(j)}))]$ where recall τ_i and $\hat{\tau}_i$ are the optimal and suboptimal stopping policy for i^{th} Phase 2 path respectively. Using the results proved for bias of estimator $\text{Var}(\hat{V}_0(x_0))$, this term can also be shown to contribute $O\left(\frac{1}{M^2}\right)$ to the overall variance. The details of the proof methodology can be found in (Agarwal and Juneja 2012).

4 MULTIPLE TIME-PERIOD PROBLEM

We calculate the asymptotic bias and variance of stochastic mesh estimator for multiple time-period Bermudan option. Let us denote by \mathcal{F}_M , the σ -algebra generated by Phase 1 paths $(X^{(i)})_{1 \leq i \leq M}$.

4.1 Asymptotic Bias

We observe that bias of the estimator can be split as the sum of a term which corresponds to single mistake on the suboptimal exercise policy and a term which corresponds to more than one mistake. We make the idea precise as follows:

$$V_0(x_0) - \mathbb{E}[\hat{V}_{0,M}(x_0)] = \sum_{j=1}^{T-1} \mathbb{E}[(g_{\tau}(X_{\tau}) - g_j(X_j)) 1_{\{\tau > j, \hat{\tau} = j\}}] + \sum_{j=1}^{T-1} \mathbb{E}[(g_j(X_j) - g_{\hat{\tau}}(X_{\hat{\tau}})) 1_{\{\tau = j, \hat{\tau} > j\}}]. \quad (4.1)$$

It is easy to see that first term in the right hand side of (4.1) can be written as

$$\sum_{j=1}^{T-1} \mathbb{E}[(g_{\tau}(X_{\tau}) - g_j(X_j)) 1_{\{\tau > j, \hat{\tau} = j\}}] = \sum_{j=1}^{T-1} \mathbb{E}[(C_j(X_j) - g_j(X_j)) 1_{\{\tau > j, \hat{\tau} = j\}}]. \quad (4.2)$$

For the second term in the right hand side of (4.1), we first note that for all j , the set $\{\hat{\tau} > j\}$ is measurable with respect to the product σ -algebra $\mathcal{F}_j \otimes \mathcal{F}_M$. We use this observation to write

$$\mathbb{E}[(g_j(X_j) - g_{\hat{\tau}}(X_{\hat{\tau}})) 1_{\{\tau = j, \hat{\tau} > j\}}] = \mathbb{E}[(g_j(X_j) - \mathbb{E}[g_{\hat{\tau}_{j+1}}(X_{\hat{\tau}_{j+1}}) | \mathcal{F}_j \otimes \mathcal{F}_M]) 1_{\{\tau = j, \hat{\tau} > j\}}]. \quad (4.3)$$

The conditional expectation term in the right hand side of (4.3) can be seen as the expected payoff received when we use the fixed suboptimal stopping policy starting from fixed X_j . We compare the suboptimal exercise decision with the optimal stopping policy to get

$$\begin{aligned} \mathbb{E}[g_{\hat{\tau}_{j+1}}(X_{\hat{\tau}_{j+1}}) | \mathcal{F}_j \otimes \mathcal{F}_M] &= \mathbb{E}[g_{\hat{\tau}_{j+1}}(X_{\hat{\tau}_{j+1}}) 1_{\{\hat{\tau}_{j+1} = \tau_{j+1}\}} | \mathcal{F}_j \otimes \mathcal{F}_M] + \mathbb{E}[g_{\hat{\tau}_{j+1}}(X_{\hat{\tau}_{j+1}}) 1_{\{\hat{\tau}_{j+1} \neq \tau_{j+1}\}} | \mathcal{F}_j \otimes \mathcal{F}_M] \\ &= C_j(X_j) - \mathbb{E}[(g_{\tau_{j+1}}(X_{\tau_{j+1}}) - g_{\hat{\tau}_{j+1}}(X_{\hat{\tau}_{j+1}})) 1_{\{\hat{\tau}_{j+1} \neq \tau_{j+1}\}} | \mathcal{F}_j \otimes \mathcal{F}_M]. \end{aligned} \quad (4.4)$$

It can then be seen from (4.2)-(4.4), that the term in the right hand side of (4.2) simplifies to

$$\begin{aligned} V_0(x_0) - \mathbb{E}[\hat{V}_{0,M}(x_0)] &= \sum_{j=1}^{T-1} \mathbb{E}[(C_j(X_j) - g_j(X_j)) 1_{\{\tau > j, \hat{\tau} = j\}}] + \sum_{j=1}^{T-1} \mathbb{E}[(g_j(X_j) - C_j(X_j)) 1_{\{\tau = j, \hat{\tau} > j\}}] \\ &\quad + \sum_{j=1}^{T-1} \mathbb{E}[\mathbb{E}[(g_{\tau_{j+1}}(X_{\tau_{j+1}}) - g_{\hat{\tau}_{j+1}}(X_{\hat{\tau}_{j+1}})) 1_{\{\hat{\tau}_{j+1} \neq \tau_{j+1}\}} | \mathcal{F}_j \otimes \mathcal{F}_M] 1_{\{\tau = j, \hat{\tau} > j\}}]. \end{aligned} \quad (4.5)$$

In the right hand side of (4.5), the first two sums correspond to paths where a single exercise error is committed using the suboptimal policy whereas the third sum corresponds to paths where more than one exercise error is committed using the suboptimal policy. We can perform analysis same as the two time-period problem to show that for all $1 \leq j \leq T - 1$,

$$\mathbb{E}[(C_j(X_j) - g_j(X_j)) 1_{\{\tau > j, \hat{\tau} = j\}}] = \Theta\left(\frac{1}{M}\right), \mathbb{E}[(g_j(X_j) - C_j(X_j)) 1_{\{\tau = j, \hat{\tau} > j\}}] = \Theta\left(\frac{1}{M}\right). \quad (4.6)$$

In the third sum, we can show using Assumption 3 and other results in Section 3, that for all $1 \leq j \leq T - 1$,

$$\mathbb{E}[(g_{\tau_{j+1}}(X_{\tau_{j+1}}) - g_{\hat{\tau}_{j+1}}(X_{\hat{\tau}_{j+1}})) 1_{\{\hat{\tau}_{j+1} \neq \tau_{j+1}\}} | \mathcal{F}_j \otimes \mathcal{F}_M]$$

remains upper bounded by the probability of two exercise errors on a sample path. Hence,

$$\mathbb{E}[\mathbb{E}[(g_{\tau_{j+1}}(X_{\tau_{j+1}}) - g_{\hat{\tau}_{j+1}}(X_{\hat{\tau}_{j+1}})) 1_{\{\hat{\tau}_{j+1} \neq \tau_{j+1}\}} | \mathcal{F}_j \otimes \mathcal{F}_M] 1_{\{\tau = j, \hat{\tau} > j\}}] = \Theta\left(\frac{1}{M}\right). \quad (4.7)$$

Therefore, from (4.6) and (4.7), we conclude that the asymptotic bias in (4.5) is given as

$$V_0(x_0) - \mathbb{E}[\hat{V}_{0,M}(x_0)] = \Theta\left(\frac{1}{M}\right).$$

4.2 Asymptotic Variance

We handle the variance of multiple time-period stochastic mesh estimator in the same way as for the two time-period problem. We can similarly show using Dominated Convergence theorem that variance of the individual paths will contribute $\Theta\left(\frac{1}{N}\right)$ and other smaller order terms. In the covariance term, we can use the expansion developed for bias in (4.5) and use the same results to show that the contribution of covariance to the overall variance of estimator $\hat{V}_{0,M}(x_0)$ remains upper bounded by $O\left(\frac{1}{M^2}\right)$. Hence, the variance can be shown to satisfy

$$\frac{\beta_0}{N} \leq \text{Var}(\hat{V}_{0,M}(x_0)) \leq \frac{\beta_0}{N} + \frac{\beta_1}{M^2}$$

for some constants β_0, β_1 .

5 MEAN SQUARE ERROR ANALYSIS

Typically, to perform any numerical study we have a fixed computational budget. In this section, we analyze how a fixed and large computational budget may be allocated to Phase 1 and Phase 2 to minimize the mean square error of the stochastic mesh estimator. Let M_Γ and N_Γ be the number of Phase 1 and Phase 2 independent sample paths generated for a fixed computational budget Γ . To this end, let c_1 denote the average amount of time required to generate a sample path X and c_2 be the average amount of time required to evaluate $\frac{f_t(X_t^{(1)}, X_{t+1}^{(1)})}{q_{t+1}(X_{t+1}^{(1)})} \hat{V}_{t+1}(X_{t+1}^{(1)})$. To estimate the continuation value function, on every sample

path (X_0, X_1, \dots, X_T) we calculate the average $\frac{1}{M} \sum_{j=1}^M \frac{f_t(X_t, X_{t+1}^{(j)})}{q_{t+1}(X_{t+1}^{(j)})} \hat{V}_{t+1}(X_{t+1}^{(j)})$ which involves M calculations.

Then, the aggregate effort required to compute $\hat{V}_{0,M}(x_0)$ is equal to $c_1(M + N) + c_2(M^2 + NM)T$. It then follows that

$$c_1(M_\Gamma + N_\Gamma) + c_2(M_\Gamma^2 + N_\Gamma M_\Gamma)T = \Gamma.$$

It is evident that for convergence of $\hat{V}_{0,M}(x_0)$ to the true value, we need $M_\Gamma \rightarrow \infty$ and $N_\Gamma \rightarrow \infty$ as $\Gamma \rightarrow \infty$. Consequently, $c_2T(M_\Gamma^2 + N_\Gamma M_\Gamma) \approx \Gamma$ for sufficiently large Γ . Hence, we need to minimize

$$\text{MSE}(\hat{V}_{0,M}(x_0)) = \text{Var}(\hat{V}_{0,M}(x_0)) + (V_0 - \mathbb{E}[\hat{V}_{0,M}(x_0)])^2$$

subject to the constraint $c_2T(M_\Gamma^2 + N_\Gamma M_\Gamma) \approx \Gamma$. From our results for two time-period problem in Section 3 and the subsequent discussion for multiple time-period problem in Section 4, we get for large enough M, N

$$\frac{\bar{K}_0}{N} + \frac{\bar{K}_1}{M^2} \leq \text{MSE}(\hat{V}_{0,M}(x_0)) \leq \frac{\bar{K}_0}{N} + \frac{\bar{K}_2}{M^2}$$

where \bar{K}_0, \bar{K}_1 and \bar{K}_2 are some constants. In order to minimize $\text{MSE}(\hat{V}_{0,M}(x_0))$, we can check that for fixed computational budget Γ , the optimal M_Γ^* and N_Γ^* are given as

$$M_\Gamma^* = \Theta(\Gamma^{1/3}), \quad N_\Gamma^* = \Theta(\Gamma^{2/3}),$$

and the optimal $\text{MSE}(\hat{V}_{0,M}(x_0))$ decays as $\Theta(\Gamma^{-2/3})$.

6 LEAST SQUARES METHOD

In this method, the continuation value function of Bermudan option, is parameterized as a linear combination of pre-specified basis functions. Although very popular and efficient, the essential drawback of least square method (LSM) algorithm is that the choice of basis functions is specific to the option pricing problem. Also, in general only a complete (infinite) set of basis functions results in continuation value function estimators that are consistent for the true option value. In practice, a finite set of basis functions is used which introduces an approximation error. This approximation error can propagate backwards through the exercise opportunities and produce estimators that do not converge to the true option value as shown in Section 7.

Clément, Lamberton, and Protter (2002) addressed the convergence issues and analyzed the asymptotic convergence rate of ‘interleaving’ LSM option price estimator. We combine their methodology with our analysis approach to derive the asymptotics for bias and variance of lower biased estimator based on the least squares method. Once again, we use the two phase algorithm to calculate the option price estimator. In Phase 1, we generate M independent sample paths of the underlying process X and use LSM algorithm to estimate the continuation value function. We fix the number of basis functions to $m \geq 1$ and denote by $e(x)$ the basis function vector $(e_1(x), \dots, e_m(x))$. The continuation value estimator $\hat{C}_t(\cdot)$ based on m basis functions is then given as $\hat{C}_t(X_t) := \alpha_{t,M} \cdot e(X_t)$ where $\alpha_{t,M} = (A_{t,M})^{-1} \frac{1}{M} \sum_{i=1}^M g_{\hat{c}_{t+1}^{(i)}}(X_{t+1}^{(i)}) e(X_t^{(i)})$, $A_{t,M}$ is an

$m \times m$ matrix, with coefficients given by $(A_{t,M})_{1 \leq j,k \leq m} = \frac{1}{M} \sum_{i=1}^M e_j(X_t^{(i)}) e_k(X_t^{(i)})$, and $\hat{\tau}_t^{(i)}$ is the suboptimal stopping time based on the LSM algorithm defined as follows:

$$\hat{\tau}_T^{(i)} = T, \hat{\tau}_t^{(i)} = t 1_{\{g_t(X_t^{(i)}) \geq \alpha_{t,M} \cdot e(X_t^{(i)})\}} + \hat{\tau}_{t+1}^{(i)} 1_{\{g_t(X_t^{(i)}) < \alpha_{t,M} \cdot e(X_t^{(i)})\}}, \quad 1 \leq t \leq T - 1.$$

In Phase 2, we generate a new set of N independent sample paths of X and evaluate the payoff values on each of them using the suboptimal stopping policy based on the estimated continuation value function from Phase 1. Then, the final estimator is given as $\hat{V}_{0,M}^m(x_0) = \frac{1}{N} \sum_{i=1}^N g_{\hat{\tau}_t^{(i)}}(X_t^{(i)})$.

Let $C_t^m(\cdot)$ denote the continuation value function which is exactly equal to the linear combination of m basis functions. If $V_0^m(\cdot)$ is the option value function when the continuation value function is given as $C_t^m(\cdot)$, we can show that the bias of the estimator for large enough M satisfies,

$$\frac{\alpha_1}{M} \leq (V_0^m(x_0) - \mathbb{E}[\hat{V}_{0,M}^m(x_0)]) \leq \frac{\alpha_2}{M}$$

for some constants α_1, α_2 . The idea of the proof is to use the Central Limit theorem shown for continuation value estimator in (Clément, Lamberton, and Protter 2002) and argue on the same lines as the stochastic mesh estimator to show that the result holds. We can also show by similar analysis that for large enough M, N and for some constants β_0, β_1 , variance of the LSM based option price estimator $\hat{V}_{0,M}^m(x_0)$ satisfies

$$\frac{\beta_0}{N} \leq \text{Var}(\hat{V}_{0,M}^m(x_0)) \leq \frac{\beta_0}{N} + \frac{\beta_1}{M^2}$$

Next, we analyze the convergence rate of mean square error of LSM based lower biased estimator. Let M_Γ and N_Γ be the number of Phase 1 and Phase 2 independent sample paths generated for a fixed computational budget Γ . The average cost of generating M independent sample paths is $c_1 M$. The coefficient in the continuation value approximation is calculated using matrix inversion which contributes c'_2 to the overall cost. Thus, repeating the procedure over T exercise opportunities gives the total cost of coefficient evaluation as $c'_2 T$. Next, on each sample path we evaluate the continuation value approximation to compare with the immediate exercise value. For m basis functions, this contributes $c'_3 m M$ at every exercise opportunity. Hence, the total Phase 1 computational burden is given as $c_1 M + c'_2 T + c'_3 m M T$. We can immediately see that in the second phase, computational cost is given as $c_1 N + c'_3 m N T$. Hence, for fixed values of m and T , the total computational cost denoted by Γ equals $k_1 (M + N)$ for some constant k_1 . Therefore, we need to minimize $\text{MSE}(\hat{V}_{0,M}^m(x_0))$ subject to the constraint $k_1 (M + N) \approx \Gamma$. From the discussion above, we have for some constants \bar{K}_0, \bar{K}_1 and \bar{K}_2 ,

$$\frac{\bar{K}_0}{N} + \frac{\bar{K}_1}{M^2} \leq \text{MSE}(\hat{V}_{0,M}^m(x_0)) \leq \frac{\bar{K}_0}{N} + \frac{\bar{K}_2}{M^2}.$$

By simple calculations, we can see that the optimal M_Γ^* and N_Γ^* are given as

$$M_\Gamma^* = \Theta(\Gamma^{2/3}), \quad N_\Gamma^* = \Theta(\Gamma),$$

for some constant \bar{K} and the optimal $\text{MSE}(\hat{V}_{0,M}^m(x_0))$ decays as $\Theta(\Gamma^{-1})$.

7 NUMERICAL EXAMPLES

In this section, we illustrate our theoretical results with the help of numerical examples. We consider a Bermudan call option on the maximum of five assets with the payoff function as $g_t(X_t) = (\max(X_t^{(1)}, \dots, X_t^{(5)}) - K)^+$ where K is the strike price. Under the risk-neutral measure, asset prices are assumed to follow correlated geometric Brownian motion processes, i.e. $dX_t^{(i)} = rX_t^{(i)} dt + \sigma_i X_t^{(i)} dW_t^i$, where W_t^i is a standard Brownian

motion process and the instantaneous correlation of W^i and W^j is ρ_{ij} . For simplicity, we take $\sigma_i = \sigma$ and $\rho_{ij} = 0$ for all $i, j = 1, \dots, 5$ and $i \neq j$. The exercise opportunities are assumed to occur at n equidistant time intervals $t_j = jT/(n-1), j = 0, 1, \dots, n-1$ where T is the expiry of the option. From the analysis in Section 3 - Section 5, we know that for sufficiently large values of M and N , the number of Phase 1 and Phase 2 paths respectively, computational budget Γ satisfies the following relationship: $\Gamma = c_1M^2 + c_2MN$ where Γ is measured in computer run time. In order to plot mean square error for different values of (M, N) given fixed Γ , we first need to estimate the coefficients c_1, c_2 . We perform this task by tabulating the running time for calculation of the estimator $\hat{V}_{0,M}$ using different values of (M, N) and then regressing the computational budget over (M, N) . We used 10 different values of (M, N) to estimate $\hat{c}_1 = 2.02 \times 10^{-3}$ and $\hat{c}_2 = 9.81 \times 10^{-4}$. Next, for different computational budget, we plot the MSE of lower biased estimator $\hat{V}_{0,M}$ with respect to M in Figure 1. We observe in the case where computational budget for a single iteration is 16 mins, a choice of $M = 12,000$ and $N = 57,000$ achieves an MSE of 0.049 whereas the optimal choice of $M = 16,100$ and $N = 28,000$ achieves an MSE of 0.038. Hence, the optimal choice of parameters provides 20% improvement over the naive choice. In this example, we set $X_0^{(1)} = X_0^{(2)} = \dots = X_0^{(5)} = 90$,

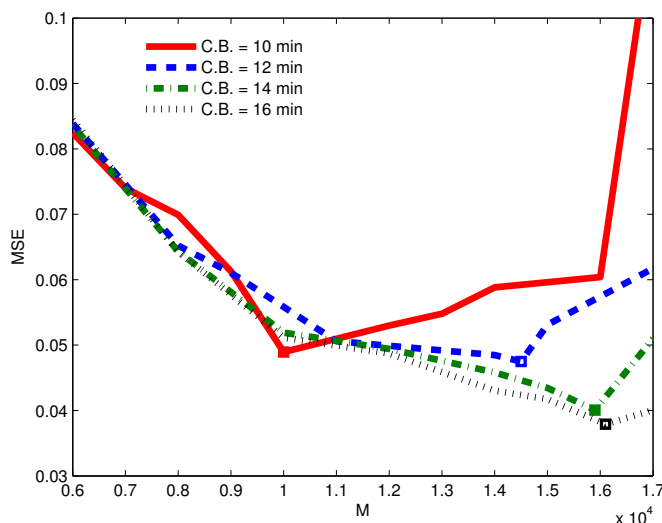


Figure 1: Mean square error of the lower biased estimator calculated using 50 independent samples is plotted with respect to the number of Phase 1 paths, M . The computational budget (C.B.) indicated is for single iteration of the algorithm.

$r = 5\%, \sigma = 20\%, \rho = 0, K = 100, n = 4$ and $T = 3$ years. True value of this option is estimated in Broadie and Glasserman (2004) to be 16.06.

Next, to validate the theoretical rate of convergence of $MSE(\hat{V}_{0,M})$, we plot the optimal mean square error for different values of the computational budget in Figure 2 with log scale on both the axes. The plot illustrates that for a given computational budget, logarithm of optimal MSE varies as $-0.6691 \times \log(\Gamma) + 0.575$ which is in close agreement with the theoretical rate of convergence, i.e. optimal $MSE = -\frac{2}{3} \times \log(\Gamma) + c$ for some constant c .

We demonstrate the limitation of least squares method (LSM) for estimating the true option value as follows. In the setting of previous example, we consider a strangle spread payoff function $g_t(\cdot)$ with strikes 50, 90, 110 and 150 as illustrated in the numerical example of Kohler, Krzyżak, and Todorovic (2010) based on the average of underlying stock prices $X_t^{(1)}, \dots, X_t^{(5)}$. Since we calculate the lower biased estimator, higher estimate values are desired. We calculate the LSM estimator with 100 million Phase 1 and Phase 2 sample paths and the stochastic mesh estimator with 10,000 Phase 1 and 40,000 Phase 2 sample paths.

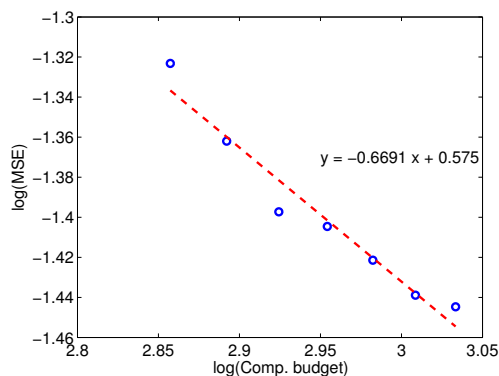


Figure 2: Relationship of optimal mean square error of the lower biased estimator for different computational budget is plotted. The calculations are based on 50 independent samples.

We note in Table 1 that with the same computational budget, LSM estimator fails to converge to the true option value whereas stochastic mesh estimator provides a better lower bound.

Table 1: Performance of lower biased estimator using LSM and stochastic mesh method (SMM). The calculations are based on 50 independent samples and the mentioned computational budget (Γ) is for single iteration.

Estimator type	Estimated value (Std. Dev.)	Γ (in secs)
LSM with up to 2 nd degree polynomials on $1, X_t^{(1)}, \dots, X_t^{(5)}, g_t$	24.21 (0.0017)	620
SMM	24.58 (0.0614)	550

8 CONCLUSION

We compared the optimal convergence rate of lower biased estimator based on stochastic mesh and least squares method. We showed how to optimally allocate the computational budget in Phase 1 and Phase 2 of the algorithm such that the mean square error of the estimator is minimized. We further conducted numerical experiments to validate the theoretical results established in this paper.

ACKNOWLEDGMENTS

The first author is supported by Tata Consultancy Services PhD scholarship.

REFERENCES

- Agarwal, A., and S. Juneja. 2012. “Nearest neighbor method based estimation technique for pricing Bermudan options”. *Preprint*.
- Andersen, L. 1999. “A simple approach to the pricing of Bermudan swaptions in the multi-factor Libor Market Model”. *Available at SSRN 155208*.
- Andersen, L., and M. Broadie. 2004. “Primal-dual simulation algorithm for pricing multidimensional American options”. *Management Science*:1222–1234.
- Belomestny, D. 2011. “Pricing Bermudan options by nonparametric regression: optimal rates of convergence for lower estimates”. *Finance and Stochastics* 15 (4): 655.
- Broadie, M., and P. Glasserman. 1997a. “Pricing American-style securities using simulation”. *Journal of Economic Dynamics and Control* 21 (8-9): 1323–1352.

- Broadie, M., and P. Glasserman. 1997b. “A stochastic mesh method for pricing high-dimensional American options”. Technical report.
- Broadie, M., and P. Glasserman. 2004. “A stochastic mesh method for pricing high-dimensional American options”. *Journal of Computational Finance* 7:35–72.
- Carriere, J. 1996. “Valuation of the early-exercise price for options using simulations and nonparametric regression”. *Insurance: Mathematics and Economics* 19 (1): 19–30.
- Clément, E., D. Lamberton, and P. Protter. 2002. “An analysis of a least squares regression method for American option pricing”. *Finance and Stochastics* 6 (4): 449–471.
- Feller, W. 2008. *An introduction to probability theory and its applications*, Volume 2. John Wiley & Sons.
- Haugh, M., and L. Kogan. 2004. “Pricing American options: a duality approach”. *Operations Research*:258–270.
- Kohler, M., A. Krzyżak, and N. Todorovic. 2010. “Pricing of high-dimensional American options by neural networks”. *Mathematical Finance* 20 (3): 383–410.
- Longstaff, F., and E. Schwartz. 2001. “Valuing American options by simulation: A simple least-squares approach”. *Review of Financial Studies* 14 (1): 113–147.
- Resnick, S. 1999. *A Probability Path*. Birkhäuser.
- Rogers, L. C. 2002. “Monte Carlo valuation of American options”. *Mathematical Finance* 12 (3): 271–286.
- Tsitsiklis, J., and B. Van Roy. 2001. “Regression methods for pricing complex American-style options”. *Neural Networks, IEEE Transactions on* 12 (4): 694–703.
- Tsitsiklis, J. N., and B. Van Roy. 1999. “Optimal stopping of Markov processes: Hilbert space theory, approximation algorithms, and an application to pricing high-dimensional financial derivatives”. *Automatic Control, IEEE Transactions on* 44 (10): 1840–1851.

AUTHOR BIOGRAPHIES

ANKUSH AGARWAL is a PhD student in the School of Technology and Computer Science at Tata Institute of Fundamental Research, Mumbai, India. He received a Bachelor of Technology and Master of Technology from Indian Institute of Technology, Mumbai. His research focuses on application of simulation methods to computational finance. His email address is ankush@tcs.tifr.res.in.

SANDEEP JUNEJA is a Professor in the School of Technology and Computer Science at Tata Institute of Fundamental Research. His Ph.D. was from Department of Operations Research at Stanford University (1993). His research interests include applied probability, computational finance, simulation and rare-event analysis. He serves on the editorial board on Mathematics of Operations Research. His email address is juneja@tifr.res.in and his web page is <http://www.tcs.tifr.res.in/~sandeepj>.