39th Actuarial Research Conference

On the existence of an optimal regression complexity in the Least-Squares Monte Carlo (LSM) framework for options pricing

Yu Zhou
Department of Statistics and Actuarial Science,
University of Waterloo,
Waterloo, Ontario N2L 3G1, Canada.

Key words: Least-Squares Monte Carlo (LSM); Simulation; Smoothing spline.

Abstract:

In this paper, we illustrate how to value American-style options using the Least-Squares Monte Carlo (LSM) approach proposed by Longstaff and Schwartz (2001) and investigate whether there exists an optimal regression complexity in the LSM framework for options pricing. In particular, we use the smoothing spline in the regression step, which allows us to control the regression complexity on a continuous scale with just one tuning parameter. Numerical results on American put options indicate that we need to use more than a linear regression, but as the regression becomes more complex, the accuracy of the LSM method quickly deteriorates.

1. Introduction

In modern financial markets, one of the most challenging and difficult problems is the valuation of American-style derivatives. Numerical PDE method and lattice method are widely used to price American-style derivatives, see Tavella (2002, Chapter 6). These two methods work backwards. However, in order to apply numerical PDE method, we have to restrict our attention to those derivatives with a small number of dimensions (less than four state variables). For multi-factor and path-dependent problems, Monte Carlo simulation is a nature tool, but it must work in a forward fashion.

Longstaff and Schwartz (2001) proposed a simple and powerful method, known as Least Squares Monte Carlo (LSM), to value American-style derivatives. By introducing a regression step, LSM allows us to use Monte Carlo to solve problems where backward induction cannot possibly be avoided. The LSM method is simple, promising and powerful. Only simple least-squares is required. The two authors claim that the method can be applied to most general stochastic processes.

This paper is organized as follows: Chapter 2 reviews the LSM method in detail. Chapter 3 examines the optimal complexity for the regression step in the LSM algorithm. We use nonparametric techniques as a tool to identify the optimal regression complexity. Chapter 4 is a short conclusion.

2. Least Squares Monte Carlo Simulation

Monte Carlo (MC) simulation is an alternative to the numerical PDE method. Boyle (1977) is the first researcher to introduce Monte Carlo simulation into finance. The method itself is simple and easy to implement. We can simulate as many sample paths as desired according to the underlying stochastic differential equation that describes the stock process. For each sample path, the option value is determined and the average from all paths is the estimated option price. The variance of our estimate is $O(1/\sqrt{N})$, which is independent of the number of stochastic dimensions.

For the European option, the MC method works well. In fact, we even have an analytical solution, e.g., using the Black-Scholes formula. More importantly, the value is determined only by the terminal stock price if one assumes constant interest rate and volatility. It is easy to see that Monte Carlo simulation must work in a forward fashion. Therefore, even though it is simple and capable of handling multi-factor problems, once we have to solve a problem backwards, Monte Carlo simulation becomes awkward to implement.

There are basically two ways to value American-style options. The first is to approximate the early exercise boundary so that we can have the boundary before we run simulations. Then, for each sample path, the simulation runs forward until the stock price hits the exercise boundary. At the end of the simulation, the averaging process is exactly the same as that of the European option. Bossaerts (1989) first proposed this method and solved for the optimal strategy by maximizing the simulated value of the option. Tilley (1993), Barraquand and Martineau (1995), Broadie and Glasserman (1997), and Carr (1998) are other examples of this approach, while the list is by no means complete. These authors applied different techniques to approximate the transitional density function or the early exercise boundary.

The other approach is similar to what will be presented in this paper. Instead of determining the exercise boundary before simulation, this approach focuses on the conditional expectation function; see e.g., Carriere (1996), Tsitsiklis and Roy (1999). Longstaff and Schwartz (2001) proposed the Least-Squares Monte Carlo (LSM) method, an easy way to implement this approach, which will be main focus on this paper. Clement, Lamberton and Protter (2001) studied related convergence issues. Tian and Burrage (2002) discussed the accuracy of the LSM method. Moreno and Navas (2003) further discussed the robustness of LSM with regard to the choice of the basis functions.

2.1 The Valuation Algorithm

Longstaff and Schwartz (2001) introduce the use of Monte Carlo simulation and least squares to value American options. At each exercise time point, option holders compare the payoff for immediate exercise with the expected payoff for continuation. If the payoff for immediate exercise is higher, then they exercise the options. Otherwise, they will leave the options alive. The expected payoff for continuation is conditional on the information available at that time point. The authors propose that the conditional expectation can be estimated using simulated cross-sectional data by least squares. To find out the conditional expectation function, we regress the realized payoffs from continuation on a set of basis functions in the underlying asset prices. The fitted values are chosen as the expected continuation values. We simply compare these continuation values with the immediate exercise values and make the optimal exercise decisions. We recursively use this algorithm and discount the optimal payoffs to time zero. That is the option price. For details of the algorithm, see Longstaff and Schwartz (2001).

Consider an American option, which can be exercised at any time point $t \in [0,T]$; we have to use discretization to approximate the continuous exercise feature. Suppose we use m time points $0 < t_1 \le t_2 \le ... \le t_m = T$. At maturity, the exercise strategy is the same as the European counterpart. If the option is in the money, the investor should exercise it. Otherwise, let it expire. Before maturity, the option holder must choose between exercising the option and holding it to the next exercisable time. We adopt the notations from Longstaff and Schwartz's paper. Define a probability space (Ω, F, P) and an equivalent martingale measure **Q**. Denote $C(\omega, s; t, T)$ as cash flows at time s generated by the option for the sample path ω, conditional on the option not being exercised at or prior to time t, and on the option holder following the optimal stopping rule for all $s, t < s \leq T$. Then the value of continuation at time $t_k, V\left(\omega; t_k\right)$ can be expressed as

$$V\left(\omega, t_k\right) = E^Q\left[\sum_{j=k+1}^{m} \exp\left(-\int_{t_k}^{t_j} r\left(\omega, s\right) ds\right) C\left(\omega, t_j; t_k, T\right) \middle| F_{t_k}\right],$$

where $r\left(\omega, t\right)$ is the riskless interest rate. Within this framework, the problem of pricing an American option reduces to comparing the immediate exercise value with this conditional expectation. The immediate exercise value is obtained directly from the payoff function. The essential task is to determine the conditional expectation. The LSM method uses least square regression to find the conditional expectation function at $t_{m-1}, t_{m-2}, ..., t_1$. We assume that the unknown function $V(\omega; t_k)$ can be represented as a linear combination of a countable set of basis functions. The simplest basis function is the polynomial function:

$$L_n(X) = X^n.$$

More complicated choices of basis functions can be the Laguerre polynomials:

$$L_n(X) = \exp(-\frac{X}{2}) \frac{e^X}{n!} \frac{d^n}{dX^n} (X^n e^{-X}).$$

Other types of basis functions include the Hermite, Legendre, Chebysheve, etc. $V(\omega; t_k)$ can be expressed as $V(\omega; t_k) = \sum_{i=1}^{\infty} b_i L_i(X)$.

Suppose we have only one state variable, e.g., in the case of an American option. We estimate the conditional expectation $V(\omega; t_{k-1})$ by using the first $M < \infty$ basis functions and, borrowing the notation from Longstaff and Schwartz again, denote the estimate $V_M(\omega; t_{k-1})$. It is obtained by regressing the discounted values of $C(\omega, t_{k-1})$ onto the basis functions of the state variable for paths that are in-the-money. The fitted value $\hat{V}_M(\omega; t_{k-1})$ is the continuation value of the derivative. The immediate exercise value $V_M^*(\omega; t_{k-1})$ is computed directly from the payoff function. By comparing the two values, we can determine whether it is optimal to exercise the option at t_{k-1} or to hold it until t_k. Once the decision at t_{k-1} is made, we are ready to approximate the cash flow $C(\omega, s; t_{k-2}, T)$. Then, we move backward recursively to get $C(\omega, s; t, T)$ for each time point $t = 0, t_1, t_2, ..., t_m = T$ and decide the optimal stopping time along the particular path considered.

2.2 Valuing American Put Options

Assume that the underlying asset is a common stock and the dynamics of the stock price under the risk-neutral measure is given by $dS_t = rS_t dt + \sigma S_t dW_t$. Here r and σ are constants, $\{W_t\}$ is a standard Brownian motion under the measure **Q**. The analytical solution is available:

$$S_t = S_0 \exp\left(\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right)$$

The Monte Carlo approach focuses on the Brownian motion and simulates the increment of the Brownian motion, which is normally distributed. As for the basis function, we choose a constant and the first four terms of the polynomial, or the first four terms of the Laguerre polynomials.

2.3 Least Squares Monte Carlo for valuing American put options

Following Longstaff and Schwartz (2001), we compare numerical PDE and simulation values for the early exercise option in an American-style put option on a share of non-dividend stock. The option can be exercised 50 times per year. The strike price of the put is 40; the short-term interest rate is 0.06; the underlying stock price S, the volatility of return σ, and the number of years until the final expiration of the option T are as listed. We simulate 100,000 (50,000 plus 50,000 antithetic) paths for the stock price process. The standard errors of the simulation estimates are reported as well.

Table 1 below shows that we can obtain fairly accurate results using the LSM method. The difference between PDE and LSM is very small. The standard errors are within 1 cent due to the use of the variance reduction technique, antithetic variates.

Table 1: Least Squares Monte Carlo for valuing American put options

S	σ	T	PDE	Simulated American	$(s.e.)$	Longstaff Paper
36	0.2	1	4.478	4.4789	0.0060	4.472
36	0.2	2	4.840	4.8376	0.0069	4.821
36	0.4	1	7.101	7.0997	0.0084	7.091
36	0.4	2	8.508	8.5003	0.0105	8.488
38	0.2	1	3.250	3.2494	0.0047	3.244
38	0.2	2	3.745	3.7389	0.0061	3.735
38	0.4	1	6.148	6.1475	0.0080	6.139
38	0.4	2	7.670	7.6564	0.0098	7.669
40	0.2	1	2.314	2.3157	0.0054	2.313

3. Examining Optimal Regression Complexity Using Nonparametric Techniques

In this chapter, we analyze the effect of the number of basis functions on the valuation accuracy and test whether there exists an optimal level of complexity for the regression step in the LSM framework. Throughout Chapter 3, we use four American put options as our primary examples (Table 2).

Table 2: Examples of American put options

Examples	S	σ	T	PDE
Option 1	36	0.2	1	4.478
Option 2	36	0.4	1	7.101
Option 3	38	0.2	1	3.250
Option 4	38	0.4	1	6.148

All of these options can be exercised 50 times per year. All of them have a strike price of 40. The short-term interest rate is 0.06. We simulate 100,000 (50,000 plus 50,000 antithetic) paths for each stock-price process. Standard errors are listed in parentheses. Our primary measure of valuation accuracy is the percent error:

$$\left(\frac{|LSMvalue - PDEvalue|}{PDEvalue}\right).$$

The numerical results with up to 10 polynomial basis functions are listed in Table 3 for the four American put options. Longstaff and Schwartz (2001) gave a convergence criterion for determining the number of basis functions needed to obtain an accurate approximation: simply add more basis functions until the value implied by the LSM algorithm no longer increases. Moreno and Navas (2003) find that, in some cases, the option prices do not increase monotonically with the number of basis functions, which means Longstaff and Schwartz's original convergence criterion is difficult to apply. Figure 1 plots the percent error vs. the number of basis functions; these plots suggest that Moreno and Navas (2003) are right.

In addition, we can observe empirically that using more basis functions can actually make LSM less accurate! In order to further test the observation that using more basis functions can degrade the performance of LSM, we now conduct an experiment in which we replace the ordinary least-squares with smoothing splines.

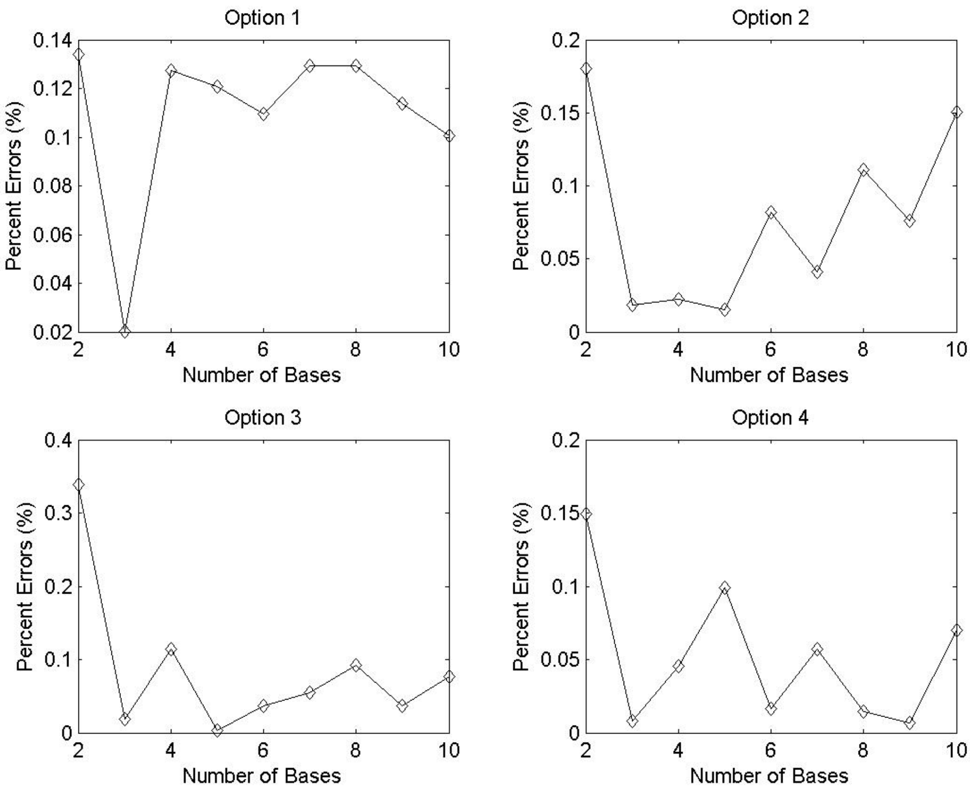

Figure 1: Plots of number of bases vs. percent errors (%) for option 1, 2, 3 and 4.

Table 3: Effects of the number of basis functions for the four American put options

Number of polynomial bases	Values of option 1	Percent Errors	Values of option 2	Percent Errors	Values of option 3	Percent Errors	Values of option 4	Percent Errors
2	4.4720(0.0064)	0.1340%	7.0882(0.0087)	0.1803%	3.2390(0.0050)	0.3385%	6.1388(0.0081)	0.1496%
3	4.4789(0.0060)	0.0201%	7.0997(0.0084)	0.0183%	3.2494(0.0047)	0.0185%	6.1475(0.0080)	0.0081%
4	4.4837(0.0060)	0.1273%	7.0994(0.0083)	0.0225%	3.2537(0.0047)	0.1138%	6.1452(0.0080)	0.0455%
5	4.4838(0.0060)	0.1206%	7.1021(0.0083)	0.0155%	3.2499(0.0046)	0.0031%	6.1419(0.0079)	0.0992%
6	4.4829(0.0060)	0.1094%	7.1068(0.0083)	0.0817%	3.2512(0.0046)	0.0369%	6.1470(0.0080)	0.0163%
7	4.4838(0.0060)	0.1295%	7.1039(0.0083)	0.0408%	3.2518(0.0046)	0.0554%	6.1445(0.0080)	0.0569%
8	4.4838(0.0059)	0.1295%	7.1089(0.0083)	0.1113%	3.2530(0.0046)	0.0923%	6.1471(0.0080)	0.0146%
9	4.4831(0.0059)	0.1139%	7.1064(0.0083)	0.0760%	3.2512(0.0046)	0.0369%	6.1484(0.0080)	0.0065%
10	4.4825(0.0059)	0.1005%	7.1117(0.0083)	0.1507%	3.2525(0.0046)	0.0769%	6.1523(0.0081)	0.0699%

3.1 Smoothing Splines

Polynomials are the approximating functions of choice when a smooth function is to be approximated locally. But if a function is to be approximated on a wider interval, the degree, n, of the approximating polynomial may have to be chosen unacceptably large. The alternative way is to subdivide the interval [a,b] of approximation into sufficiently small intervals $[\xi_j .. \xi_{j+1}]$, with $a = \xi_1 < ... < \xi_{l+1} = b$, so that, on each such interval, a polynomial p_j of relatively low degree can provide a good approximation to f. This can even be done in such a way that the polynomial pieces blend smoothly, i.e., so that the resulting patched or composite function $s(x) = p_j(x)$ for $x \in [\xi_j .. \xi_{j+1}]$, all j, has several continuous derivatives. Any such smooth piecewise polynomial function is called a spline. The points $\xi_1, \xi_1, ..., \xi_{l+1}$ are called "knots." For details about splines, see Hastie, Tibshirani and Friedman (2001, Chapter 5).

Here we focus on a spline basis method called "*smoothing spline*" which avoids having to select the knots *a priori* by using a maximal set of knots. The complexity of the fit is controlled by regularization. Consider the following problem: among all functions $f(x)$ with two continuous derivatives, find one that minimizes the penalized residual sum of squares

$$RSS(f, \lambda) = \sum_{i=1}^{N} \left\{ y_i - f(x_i) \right\}^2 + \lambda \int \left\{ f''(t) \right\}^2 dt,$$

where λ is a fixed smoothing parameter. The first term measures the goodness of fit while the second term penalizes the curvature in the function. The parameter λ establishes a tradeoff between the two. Two extreme cases are:

$\lambda = 0$: f can be any function that interpolates the data;

$\lambda = \infty$: f must be linear function of x since no second derivative can be tolerated.

These vary from the very rough to the very smooth, and the hope is that $\lambda \in (0, \infty)$ indexes an interesting class of functions in between. It can be shown (see Hastie, Tibshirani and Friedman (2001, Chapter 5)) that the solution to the above penalized minimization problem is a nature cubic spline with knots at all unique values of x_i.

3.2 Valuing American Put Options

Our experiments are conducted using the same four American put options (Table 2). We simulate 2,000 (1,000 plus 1,000 antithetic) paths for the stock-price process and use different values of the smoothing parameter **spar**. The smoothing parameter spar here has the same effect as the parameter λ above. When spar is 0, the smoothing spline produces the least-squares straight-line fit to the data. When spar is 1, it is the 'natural' cubic spline interpolant. For details about the smoothing spline, see the help file of Matlab. The results are in Table 4 and Figure 2.

Table 4: Valuing American put options using nonparametric regression

	S =36 σ =0.2	Percent Errors	S =36 σ =0.4	Percent Errors	S =38 σ =0.2	Percent Errors	S =38 σ =0.4	Percent Errors
PDE	4.478		7.101		3.250		6.148	
$spar = 0.01$	4.400	1.7418%	6.9973	1.4604%	3.1794	2.1723%	6.0557	1.5013%
$spar = 0.1$	4.3898	1.9696%	6.9814	1.6843%	3.1964	1.6492%	6.0457	1.6640%
$spar = 0.2$	4.3936	1.8848%	6.983	1.6617%	3.1778	2.2215%	6.0577	1.4688%
$spar = 0.3$	4.4224	1.2416%	6.9953	1.4885%	3.1906	1.8277%	6.0852	1.0215%
$spar = 0.4$	4.4272	1.1344%	7.005	1.3519%	3.1945	1.7077%	6.0823	1.0686%
$spar = 0.5$	4.4112	1.4917%	7.0337	0.9478%	3.1969	1.6338%	6.0943	0.8735%
$spar = 0.6$	4.4124	1.4649%	7.029	1.0139%	3.1932	1.7477%	6.088	0.9759%
$spar = 0.7$	4.4221	1.25%	7.0965	0.06%	3.1955	1.68%	6.1316	0.27%
$spar = 0.8$	4.4376	0.90%	7.0701	0.44%	3.2242	0.79%	6.1554	0.12%
$spar = 0.9$	4.4495	0.6364%	7.1152	0.2000%	3.22	0.9231%	6.1724	0.3969%
$spar = 0.99$	4.5052	0.6074%	7.2157	1.6153%	3.2768	0.8246%	6.2613	1.8429%

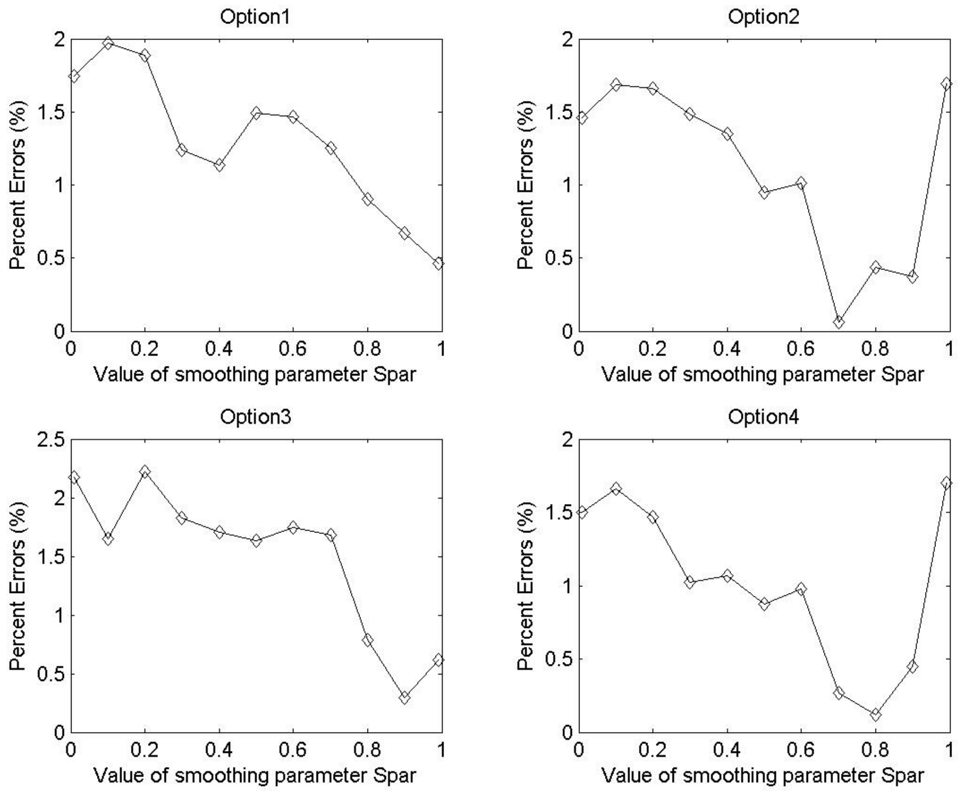

Figure 2: Plots of values of smoothing parameter spar vs. percent errors (%).

From Figure 2, we can observe empirically that the optimal smoothing parameter is around 0.8 on average. This means, in general, a linear fit (spar=1) is not adequate. However, the accuracy degrades quickly as we further decrease spar down to 0. This means we cannot use a very complex regression model either.

4. Conclusion

In this paper, we reviewed the use of the Least-Squares Monte Carlo (LSM) method to value American-style options and focused on the problem of optimal regression complexity in the LSM algorithm. Our main conclusion is that using more basis functions does not always improve the pricing accuracy; it can actually degrade the accuracy. Other researchers have similarly found that things get worse when you go overboard on regressors. For valuing American options, we found that the regression cannot be too complex, but that we do need more than a simple linear fit. The tools we used for experimentation are novel. We used a single tuning parameter in smoothing splines to control the regression complexity on a continuous scale. This is particularly convenient for our purpose, which is to investigate the effect of regression complexity on the accuracy of the LSM algorithm.

Acknowledgements

I would like to thank my advisor, Professor Mu Zhu. I have received much help and guidance from him. Thanks also to Professor Phelim Boyle, Professor Ken Seng Tan and Professor Weidong Tian for their valuable comments, suggestions and course materials. I am particularly grateful to Professor Francis A. Longstaff, who made specific suggestions for improvements. Lin Yuan, a math finance student at the University of Waterloo, also made useful suggestions for improving the paper.

Bibliography

[1] Abramowitz, M., and Stegun, I.A. (1972). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables,* Dover Publications, Inc., New York.

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

[3] Bossaerts, P. (1989). Simulation Estimators of Optimal Early Exercise, Working paper, Carnegie-Mellon University.

[4] Boyle, P. (1977). Options: A Monte Carlo Approach, *Journal of Financial Economics*, 4, 323-338.

[5] Boyle, P., Broadie, M. and Glasserman, P. (1997). Monte Carlo Methods for Security Pricing, *Journal of Economic Dynamics and Control*, 21, 1267-1321.

[6] Brandimarte, P. (2001). *Numerical Methods in Finance: a MATLAB-Based Introduction,* 1st edition, John Wiley & Sons, Inc., New York.

[7] Broadie, M., and Glasserman, P. (1997a). Monte Carlo Methods for Pricing High-Dimensional American Options: An Overview, Working paper, Columbia University.

[8] Broadie, M. and Glasserman, P. (1997b). Pricing American-Style Securities Using Simulation, *Journal of Economic Dynamics and Control*, 21, 1323-1352.

[9] Carriere, J. F. (1996). Valuation of the Early-Exercise Price for Options Using Simulations and Nonparametric Regression, *Insurance*, *Mathematics and Economics*, 19, 19-30.

[10] Carr, P. (1998). Randomization and the American Put, *Review of Financial Studies*, 11, 597-626.

[11] Clément, E., Lamberton, D. and Protter, P. (2001). An Analysis of a Least Squares Regression Method for American Option Pricing, *Finance and Stochastics*, 449-471.

[12] Garcia, D. (1999). A Monte Carlo Method for Pricing American Options, Working paper, University of California, Berkeley.

[13] Hastie, T., Tibshirani, R. and Friedman, J. H. (2001). *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, Springer Verlag.

[14] Hull, J. C. (2002). *Options, Futures and Other Derivatives*, 5th edition, Prentice Hall, Upper Saddle River, NJ.

[15] Joy, C., Boyle, P. P., and Tan, K. S. (1996). Quasi-Monte Carlo Methods in Numerical Finance, *Management Science*, 42, 926-938.

[16] Longstaff, F.A. and Schwartz, E.S. (2001). Valuing American Options by Simulations: A Simple Least-Squares Approach, *Review of Financial Studies*, 14, 1, 113-147.

[17] Moreno, M. and Navas, J. F. (2003). On the Robustness of Least-Squares Monte Carlo (LSM) for Pricing American Derivatives, Working paper, Universitat Pompeu Fabra.

[18] Pizzi, C. and Pellizzari, P. (2002). Monte Carlo Pricing of American Options Using Nonparametric Regression, Finance / Economics working paper archive at WUSTL.

[19] Raymar, S. and Zwecher, M. (1997). A Monte Carlo Valuation of American Call Options on the Maximum of Several Stocks, *Journal of Derivatives*, 5, 7-23.

[20] Rogers, L.C.G. (2002). Monte Carlo Valuation of American Options, *Mathematical Finance*, 12, 271-286.

[21] Tavella, D. (2002). *Quantitative Methods in Derivatives Pricing*, John Wiley & Sons, Inc., New York.

[22] Tian, T. and Burrage, K. (2003). Accuracy Issues of Monte-Carlo Methods for Valuing American Options, *Austral. Mathematical Soc.*2003.

[23] Tilley, J. A. (1993). Valuing American Options in a Path Simulation Model, *Transaction of the Society of Actuaries*, 45, 83-104.
[24] Tsitsiklis, J., and Roy, B. V. (1999). Optimal Stopping of Markov Processes: Hilbert Space Theory, Approximation Algorithms, and an Application to Pricing High-Dimensional Financial Derivatives, *IEEE Transactions on Automatic Control*, 44, 1840-1851.