

An Irregular Grid Method for Solving High-Dimensional Problems in Finance

Steffan Berridge^{*} and Hans Schumacher

Department of Econometrics and Operations Research and
Center for Economic Research (CentER),
Tilburg University, Postbus 90153, 5000 LE Tilburg, The Netherlands.
`s.j.berridge@kub.nl, j.m.schumacher@kub.nl`

Abstract. We propose and test a new method for pricing American options in a high dimensional setting. The method is centred around the approximation of the associated variational inequality on an irregular grid. We approximate the partial differential operator on this grid by appealing to the SDE representation of the stock process and computing the logarithm of the transition probability matrix of an approximating Markov chain. The results of numerical tests in five dimensions are promising.

1 Introduction

The current mathematical approach to option pricing was first introduced by Black and Scholes [1] and Merton [10] in 1973. They showed that pricing a European option on a stock is equivalent to solving a certain parabolic initial boundary value problem, under some simplifying assumptions.

This pricing paradigm has revolutionised the financial world and much work has gone into extending this framework. In particular the pricing of American options, which involves the solution of a parabolic free-boundary problem under the Black-Scholes-Merton assumptions, has drawn much attention since most traded options are of this type. Given the nature of the problem, it is no surprise that numerical methods are nearly always used for pricing American options.

An example of such a product is a put option on a stock, giving the buyer the right to sell a certain stock for a fixed price K at some fixed future date T . In effect this gives the buyer a payoff of $\psi(x) = \max(K - x, 0)$ currency units at the time of exercise where x is the stock price at exercise date T and K is the so-called strike price. If exercise is allowed only at T , the option is called European and if it is also allowed at any time before T it is called American.

In practise we are often confronted with problems involving several state variables such as an option written on several underlying assets or a pricing problem in which we allow some of the model parameters to become stochastic. Pricing an American option in this case amounts to solving a free-boundary problem in a high-dimensional space. Numerical methods give some hope of

^{*} Research supported by Netherlands Organisation for Scientific Research (NWO)

finding approximate solutions in this case, but it is well known that the work involved in current grid-based methods grows exponentially with the number of dimensions; this is sometimes called the curse of dimensionality.

An alternative to grid-based methods is Monte Carlo simulation of the corresponding stochastic differential equation (SDE). The European pricing problem can be solved even in a high dimensional setting using this technique since it only involves a numerical integration of the payoff function with respect to the density of the state variables.

American problems in one dimension can often be solved under the Black-Scholes-Merton assumptions by using a method where the time and state variable axes are discretised into a regular grid and a complementarity problem solved at each time-step. Unfortunately this method suffers from the curse of dimensionality.

Apart from finding the option value for a given specification of the state variables, it is also of practical interest to find the sensitivity of the price to the values of the state variables, in particular to find approximations to the first and second derivatives. In specifying numerical schemes we would like to take this into account.

The literature on option pricing has been extensive since 1973. However there is only a limited number of papers dealing with high dimensional problems, and only a few of these concentrate on the American case.

Broadie and Glasserman [2] propose a method based on approximating the stock price dynamics with a stochastic tree in which they obtain two consistent estimators, one biased high and the other low. This leads naturally to a confidence interval method for pricing American options; however the number of nodes in the tree increases exponentially with the number of exercise opportunities. In a later paper [3] they suggest a method based on a stochastic mesh which alleviates this problem; however Fu et al. [6] find that this method has a substantial upward bias in their numerical tests.

Longstaff and Schwartz [9] propose and test a dynamic programming-like method based on estimating the value of continuation at each time step by projecting realised continuation values onto a linear combination of polynomial basis functions. They call this method least squares Monte Carlo (LSM). They provide few theoretical justifications for this procedure, although the results are quite reasonable for the examples they consider. Independently, Tsitsiklis and Van Roy [14] provide theoretical justification for such a method based on the projections onto the set of basis functions (features) being orthogonal, the orthogonality being with respect to a suitably chosen inner product which changes at each time step. The error in the method is bounded by a function of the error inherent in approximating the value function by the features, which is in practice difficult to assess since the value function is exactly what we want to determine.

Stentoft [13] and Moreno and Navas [11] perform further numerical tests on the pricing of American options using the LSM method, with promising results in simple cases. However results for more complex high dimensional examples

are mixed (see [11]). Fu et al. [6] perform extensive numerical tests on American option pricing using several alternative Monte Carlo-based methods.

It is clear that a regular grid approach cannot work in high dimensions. Instead, we approximate the diffusion on an irregular grid. This grid can in principle be arbitrary, but we think of it as having been generated with a pseudo or quasi Monte Carlo method. This has the advantage that the number of points in the grid can be directly controlled, and thus does not have to grow exponentially with the dimension.

We then construct a Markov chain by defining transition probabilities between grid points in a continuous time setting in such a way that the solution to the Markov chain converges to the solution to the PDE as the number of points in the grid increases. This gives us an approximation to the partial differential operator which we use to solve the PDE in the European case and the variational inequality in the American case.

2 Formulation

There are two main paradigms which allow us to formulate and compute values for options.

In the first place we have the SDE paradigm in which the value of the option is obtained as the discounted expected payoff at expiry T . This is the most natural for financial applications since we directly use the financial processes that are specified in our model.

For a European option with payoff function ψ this leads to the pricing formula

$$v_t = e^{-r(T-t)} \mathbb{E}(\psi(X_T)) \quad (1)$$

at time t where r is the (constant) risk free rate and the expectation is taken under the risk neutral dynamics of the system. This naturally leads to a numerical method using Monte Carlo trials

$$\hat{v}_t = e^{-r(T-t)} \frac{1}{N} \sum_{i=1}^N \psi(X_T(\omega))$$

where $X_T(\omega)$ are random variables with the risk neutral densities implied by the SDE. In the American option case we also have to optimise with respect to the exercise date. In this case the pricing problem becomes an optimal stopping problem and

$$v_t = \sup_{\tau \in \mathcal{T}} \mathbb{E} \left(e^{-r(T-\tau)} \psi(X_\tau) \right) \quad (2)$$

where \mathcal{T} is the set of all stopping times with respect to the natural filtration of the underlying process. This American formulation does not lead in such a natural fashion to a pricing method.

In the second place we have the PDE paradigm in which the problem of pricing European options becomes equivalent to solving a linear parabolic initial

boundary value problem, backwards in time. Here the initial value is the option payoff at expiry and the coefficients of the PDE are inferred from those of the SDE. If we restrict our region of interest to a subset $\Omega \subset \mathbb{R}^n$ then we want to find $v_0(x)$ such that

$$\begin{aligned} v(x, t)|_{t=T} &= \psi(x) \quad \text{for } x \in \Omega \\ v(x, t) &= \phi(x, t) \quad \text{for } x \in \partial\Omega, t \in [0, T] \\ \frac{\partial v}{\partial t} + \mathcal{L}v &= 0 \quad \text{for } (x, t) \in (\Omega \setminus \partial\Omega) \times [0, T] \end{aligned} \quad (3)$$

where \mathcal{L} is a second degree elliptic operator and $\phi(x, t)$ specifies appropriate boundary conditions. Note that it is not formally necessary to restrict the problem to $\Omega \subset \mathbb{R}^n$ but computationally it is convenient. Such problems possess a unique solution given suitable regularity conditions. This formulation for the price of a European option leads naturally to finite difference and finite element methods in which the most time consuming operation is solving a system of linear equations at successive time points.

For American options the early exercise property manifests itself as a free boundary in the PDE. The problem becomes to find the solution $v(x, t)$ to the variational inequality

$$\begin{cases} \frac{\partial v}{\partial t} + \mathcal{L}v \geq 0 \\ v - \psi \geq 0 \\ \left(\frac{\partial v}{\partial t} + \mathcal{L}v\right)(v - \psi) = 0 \end{cases} \quad \text{for } (x, t) \in (\Omega \setminus \partial\Omega) \times [0, T] \quad (4)$$

with initial conditions $v(x, t)|_{t=T} = \psi(x)$ for $x \in \Omega$ and boundary conditions $v(x, t) = \phi(x, t)$ for $x \in \partial\Omega, t \in [0, T]$.

Again some regularity conditions are required for the problem to possess a unique solution (see Jaillet et al. [7]). The most popular numerical method to solve the problem in up to two dimensions is formulated by adapting a finite difference method using projected SOR (PSOR) so that the extra constraint is satisfied at each time-step. The discretised system can be treated as a linear complementarity problem. The existence and uniqueness of solutions at each time-step can be proved when the matrix M_L (see Section 3) multiplying the vector of approximate values $\mathbf{v}(t)$ is of type P (see Cottle, Pang and Stone [4]). The convergence of PSOR for real symmetric positive definite M_L is proved by Cryer [5] and the convergence of the overall computed solution at $t = 0$ (allowing for numerical errors at each time-step) is proved for certain classes of payoff functions ψ by Jaillet et al. [7]. The European and American pricing problems in one dimension are treated in detail in this paradigm by Wilmott, Dewynne and Howison [15] for example.

3 An Irregular Grid Method

The method we propose basically follows the second of the two paradigms mentioned in Section 2, but we also make use of the SDE paradigm and Monte Carlo

methodology. We approximate the value function $v(x, t)$ on a grid just as in the PDE method, but, to avoid an exponential growth of the number of grid points, we use an irregular grid. We first approximate the partial differential operator in the space direction via a semidiscrete Markov chain approximation of the SDE. Then we discretise in the second argument, using a time-stepping method which gives rise to a system of linear equations (or a complementarity problem in the American case) at each time step.

It is important to specify our approximation A in a consistent manner, in the sense that the solution obtained via the numerical method converges to the true solution as the number of points in the grid (and the number of time steps) goes to infinity. The construction of such an approximation is the main consideration of Section 4.

In this analysis we assume homogeneous Neumann conditions at the boundary (where the derivative of $v_{\cdot, t}$ in the direction normal to $\partial\Omega$ is zero). We plan to extend the analysis to other types of boundary conditions in future work.

Suppose that $\mathcal{X} = \{x_1, \dots, x_N\}$ is a representative set of points (states) in Ω on which our process can evolve for $t \in [0, T]$. For the moment we can think of \mathcal{X} as being a generic set of low discrepancy points in the sense of Niederreiter [12]. The structure of \mathcal{X} will be specified in more detail later.

Let A be our discrete approximation to the diffusion operator \mathcal{L} on \mathcal{X} , let $v_i(t)$ denote the approximated value of $v(x_i, t)$, at points x_i in \mathcal{X} , in continuous time and let $\mathbf{v}(t) = (v_1(t), \dots, v_N(t))'$ and $\psi = (\psi(x_1), \dots, \psi(x_N))'$. Thus in the European case we now wish to find the solution $\mathbf{v}(0)$ to the system of ODEs

$$\frac{d\mathbf{v}}{dt}(t) + A\mathbf{v}(t) = 0 \quad \text{for } t \in [0, T] \quad (5)$$

with initial condition $\mathbf{v}(T) = \psi$.

We now discretise the time axis so that the problem can be solved numerically at intermediate time steps. For a small time step $\delta t = T/K$ a simple approach is to discretise this system using a θ -method, where $\theta = 0, \frac{1}{2}, 1$ corresponds to explicit, Crank-Nicolson and implicit discretisation respectively. Thus for $k = K - 1, \dots, 0$ we have

$$\frac{\mathbf{v}((k+1)\delta t) - \mathbf{v}(k\delta t)}{\delta t} + \theta A\mathbf{v}(k\delta t) + (1-\theta)A\mathbf{v}((k+1)\delta t) = 0 \quad (6)$$

and the approximate solution is obtained by solving

$$M_L \mathbf{v}^k = M_R \mathbf{v}^{k+1} \quad k = K - 1, \dots, 0 \quad (7)$$

with initial condition $\mathbf{v}^K = \psi$ where \mathbf{v}^k is the approximation to $\mathbf{v}(k\delta t)$, $K = T/\delta t$ is the number of time steps and the matrices M_L and M_R are equal to $(I - \theta A\delta t)$ and $(I + (1 - \theta)A\delta t)$ respectively. Numerically we must solve a system of linear equations at each step, and thus if N is large, we would like the matrix A to be sparse.

In the American case we must include the American constraint $\mathbf{v}^k \geq \psi$, so the complementarity problem to solve is

$$\begin{cases} M_L \mathbf{v}^k - M_R \mathbf{v}^{(k+1)} \geq 0 \\ \mathbf{v}^k - \psi \geq 0 \\ (M_L \mathbf{v}^k - M_R \mathbf{v}^{(k+1)}) (\mathbf{v}^k - \psi) = 0 \end{cases} \quad k = K-1, \dots, 0 \quad (8)$$

with initial condition $\mathbf{v}^K = \psi$ and where the inequalities are componentwise. Numerically we must solve a complementarity problem at each step and again the sparsity of A is of great interest when it comes to solving such a system for large N . The complementarity problem can be solved using PSOR or linear programming, depending on the nature of the matrices involved.

4 Approximating the Partial Differential Operator

Here we provide a method for approximating the partial differential operator \mathcal{L} on our grid \mathcal{X} . To do this we appeal to the SDE representation of the problem.

We assume that the density of the SDE $f_{x_0,t}(x)$ is available for arbitrary initial points x_0 and time horizons t . We first choose a time horizon T_0 at which the density can be approximated suitably well on our grid \mathcal{X} for initial points in \mathcal{X} . Before continuing we must specify the nature of the grid \mathcal{X} . We suppose that \mathcal{X} consists of points generated from pseudo or quasi Monte Carlo trials with (importance sampling) density $g(x)$ where $g(x) > 0$ whenever $f(x) > 0$ and $x \in \Omega$.

We now let the transition probability between x_i and x_j for time horizon T_0 be approximated by the continuous density $f_{x_i,T_0}(x_j)$ weighted by the importance sampling or empirical density $\tilde{g}(x)$. That is, our transition probability matrix is $P_{T_0} = (p_{ij})$ where $p_{ij} = \mathbb{P}(X_{T_0} = x_j | X_0 = x_i)$ and we set

$$p_{ij} = \frac{1}{\sum_{k=1}^N \tilde{f}_{x_i,T_0}(x_k)} \cdot \tilde{f}_{x_i,T_0}(x_j) \quad (9)$$

and $\tilde{f}_{x_i,T_0} = \frac{f_{x_i,T_0}}{\tilde{g}}$.

As mentioned earlier, we assume homogeneous Neumann boundary conditions; in this case the value function is relatively flat in a region of the boundary, so that the distortion of probabilities in the Markov chain near the boundary will not affect the consistency of the solution.

In the European case, with no time discretisation and choosing $T_0 = T$, we see that solving the option pricing problem using these transition probabilities amounts to Monte Carlo integration with importance sampling,

$$\begin{aligned} \mathbf{v}(0) &= \int_{\mathbb{R}^n} \psi(x) f_{x_0,t}(x) dx \approx \int_{\Omega} \psi(x) \left[\frac{f}{g} \right] (x) [g(x) dx] \\ &\approx \sum_{j=1}^N \psi(X_j) \frac{\tilde{f}_{x_0,T}(X_j)}{\sum_{k=1}^N \tilde{f}_{x_0,T}(X_k)} = \sum_{j=1}^N \psi(X_j) p_{ij} \end{aligned}$$

where X_j are iid random variables with density $g(x)$ and the importance sampling function may also incorporate the empirical density of the grid. The approximation $\hat{\mathbf{v}}(0) = P_T \psi$ is good for points x_i in the centre of our grid where the transition probabilities approximate the density well. The American problem is more complicated in that a time discretisation is called for.

The transition probability matrix P now gives us access to an approximation to the partial differential operator on our grid. We note that the evolution of state probabilities in the semidiscrete setting is given by $p(t) = e^{A^t} p(0)$ where $p(t)$ is the discrete probability distribution over our grid at time t . Thus we obtain an approximation to \mathcal{L} through finding the following matrix

$$A = \frac{1}{T_0} \log(P_{T_0}) \quad (10)$$

In effect $A = (a_{ij})$ where $a_{ij} = \lim_{\delta t \downarrow 0} \frac{1}{\delta t} [\mathbb{P}(X_{t+\delta t} = x_j | X_t = x_i) - \delta_{ij}]$.

We could also find a transition probability matrix for small time steps δt by taking a root of P_{T_0} , $P_{\delta t} = e^{A\delta t} = (P_{T_0})^{\delta t/T_0}$. This could in principle be useful for directly specifying the matrices M_L and M_R as used in Section 3, and has been found to be a faster operation than computing the logarithm in Matlab.

The more naive approach of calculating the transition probability matrix directly through (9) with $T_0 = \delta t$ has been found to lead to inaccurate results. This suggests that the transition probabilities do not reflect the density of the process over longer horizons when they are calculated in this way.

5 Experimental Results

We conduct experiments to test our proposed method in a Matlab environment, and using no special techniques to accelerate the speed of computation. In particular no attention is paid to the approximately sparse nature of the transition probability matrix when calculating the logarithm. This meant that the maximum feasible grid size was 1500, which is a relatively low number in terms of Monte Carlo integration.

Our experiments are based on results from Stentoft [13], who obtains approximate prices for options written on three and five assets using the LSM method. He considers stock processes driven by correlated Brownian motions.

Specifically our stock prices $\mathbf{S} = (S_1, \dots, S_n)$ (with Itô correction term included) are given by $\mathbf{S}(t) = \mathbf{S}(0) \exp\left\{(r\mathbf{1} - \frac{1}{2}\text{diag}(\Sigma))t + R\mathbf{W}(t)\right\}$ where $\mathbf{W}(t) = (W_1(t), \dots, W_n(t))'$ is a vector of independent Brownian motions, $\mathbf{1}$ is a vector of ones, $\text{diag}(\Sigma)$ is a vector of the diagonal entries of Σ and R is a Cholesky factor of the covariance matrix $\Sigma = RR'$, the latter giving the covariances of the stock processes in the log domain.

We are given initial stock prices $S_i(0) = 40$ for each i , the correlations between log stock prices are $\rho_{ij} = 0.25$, $i \neq j$, and volatilities¹ are $\sigma_i^2 = 0.04$ for all i , the risk-free interest rate is fixed at $r = 0.06$ and the expiry is $T = 1$.

¹ In Stentoft's paper the volatilities are incorrectly specified as $\sigma_i^2 = 0.2$.

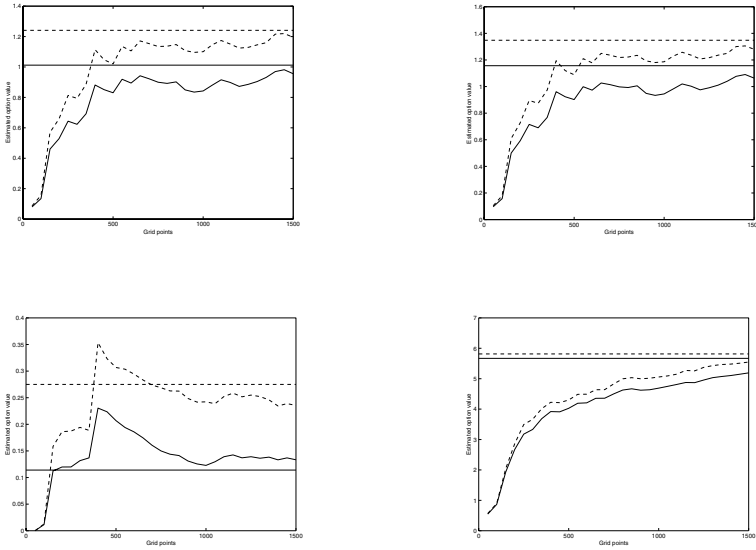


Fig. 1. QMC grid valuation of European (*solid lines*) and American (*dotted lines*) put options on the arithmetic average, geometric average, maximum and minimum respectively over five assets. Stentoft's solutions are drawn as horizontal lines

Furthermore we set $T_0 = T = 1$ in the approximation of the partial differential operator and use implicitness $\theta = \frac{1}{2}$ and $K = 20$ time-steps (in the European case we found that the difference between the $(K = 20)$ time-discretised and single time-step calculations was negligible).

Our grid is generated using Sobol points with uniform density $g(\mathbf{x}) \equiv (\int_{\Omega} d\mathbf{x})^{-1}$. The first grid point is set to be the vector of log stock prices at $t = 0$; we only consider the estimated value at this point in our convergence analysis. The region of interest Ω is set to be the rectangular region (L, U) where $L_i = E(S_i(T)) - 3\sigma_i\sqrt{T}$ and $U_i = E(S_i(T)) + 3\sigma_i\sqrt{T}$.

The payoff functions considered correspond to put options on the maximum, minimum, arithmetic mean and geometric mean which have respective payoff functions $\psi_1(\mathbf{x}) = (K - \max_i(x_i))^+$, $\psi_2(\mathbf{x}) = (K - \min_i(x_i))^+$, $\psi_3(\mathbf{x}) = (K - \frac{1}{n} \sum x_i)^+$ and $\psi_4(\mathbf{x}) = (K - (\prod x_i)^{1/n})^+$.

Figure 1 shows the QMC grid approximation of the option values for increasing numbers of grid points. For most cases we see that the approximation for

1500 grid points is within 10% of Stentoft's estimate; however it appears that more than 1500 points are required to obtain a satisfactory convergence. This is not surprising if we consider the number of trials required to obtain accurate results in numerical integration.

The similarity between American and European convergence patterns is encouraging if we consider that the numerical scheme in the European case is essentially Monte Carlo numerical integration, which we know is a tractable numerical scheme with respect to dimensionality.

6 Conclusions and Further Research

We have proposed a new method for finding the value of American options in a high-dimensional setting. Central to this method is the use of an irregular grid over the state space and an approximation of the partial differential operator on this grid.

In our analysis we allow any grid which is generated using Monte Carlo trials with respect to a known density function. We also suggest a method to correct for the empirical density of these points in order to give a more accurate Markov chain approximation to the SDE on our region of interest, although this was not tested experimentally.

Once the Markov chain approximation has been obtained, we use the transition probability matrix to form a semidiscrete approximation to the partial differential operator corresponding to this Markov chain. This is done through taking a logarithm or root of the transition probability matrix.

Our initial experiments suggest that this method has promising convergence properties. Unfortunately we have not yet been able to use more than 1500 points in our grid since we have not implemented an efficient method for computing the matrix logarithm or root. Such a method could take account of the fact that the logarithm and root of the transition probability matrix would have the same approximately sparse structure as the transition probability matrix itself.

In terms of future research possibilities there are a number of areas which can be explored, in terms of both the numerical and theoretical properties of the algorithm.

To consolidate the numerical results contained in this paper, we would like to increase the number of points in the grid. This relies crucially on being able to determine the matrix logarithm or root in a more efficient manner. Two ways to approach this are to use the approximately sparse nature of the transition probability matrix or to use Krylov subspace methods. Such a procedure may add a further approximation to the algorithm, but the effects of this are expected to be easily outweighed by the efficiency gains.

In the experiments a uniform grid was used, for convenience; however the evolution of the SDE would be better approximated on a grid of radially decreasing density, such as a normally distributed grid. We would also like to explore the possibility of making use of the empirical density of the points to further adjust the transition probabilities. This is expected to improve the accuracy of the

transition probability matrix especially when, for example, two points in the grid are very close together.

Other techniques which may prove useful are randomised QMC, which could help us in analysing the standard error of solutions, and various variance reduction techniques such as the use of control variates.

Extra information can also be obtained using our method. In particular we could manipulate the grid in order to be able to easily estimate derivatives of the option value, the so-called Greeks. This would be done by concentrating more points around the initial asset values, either by directly placing extra points or by modifying the importance sampling density.

Further work is also needed to obtain theoretical justification of the proposed method. Such a justification may follow the analysis of Jaillet et al. [7] in the American case and the analysis of Kushner and Dupuis [8] for the more general dynamic programming problem, but would also have to take into account the pseudo or quasi Monte Carlo nature of the solution method.

References

1. Black, F., Scholes, M.: The pricing of options and corporate liabilities. *Journal of Political Economy* **81** (1973) 637–654
2. Broadie, M., Glasserman, P.: Pricing American-style securities using simulation. *Journal of Economic Dynamics and Control* **21** (1997) 1323–1352
3. Broadie, M., Glasserman, P.: A stochastic mesh method for pricing high-dimensional American options. Working paper (1997)
4. Cottle, R.W., Pang, J., Stone, R.E.: *The Linear Complementarity Problem*. Academic Press (1992)
5. Cryer, C.W.: The solution of a quadratic programming problem using systematic overrelaxation. *SIAM Journal of Control* **9** (1971) 385–392
6. Fu, M.C., Laprise, S.B., Madan, D.B., Su, Y., Wu, R.: Pricing American options: a comparison of Monte Carlo simulation approaches. *Journal of Computational Finance* **4** (2001) 39–88
7. Jaillet, P., Lamberton, D., Lapeyre, B.: Variational inequalities and the pricing of American options. *Acta Applicandae Mathematicae* **21** (1990) 263–289
8. Kushner, H.J., Dupuis, P.G.: *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag (1992)
9. Longstaff, F.A., Schwartz, E.S.: Valuing American options by simulation: a simple least squares approach. *Review of Financial Studies* **14** (2001) 113–147
10. Merton, R.C.: Theory of rational option pricing. *Bell Journal of Economics and Management Science* **4** (1973) 141–183
11. Moreno, M., Navas, J.F.: On the robustness of least-squares Monte Carlo (LSM) for pricing American options. Working paper (2001)
12. Niederreiter, H.: *Random Number Generation and Quasi-Monte Carlo Methods*. SIAM (1992)
13. Stentoft, L.: Assessing the least squares Monte-Carlo approach to American option valuation. CAF Working paper No. 90 (2000)
14. Tsitsiklis, J.N., Roy, B. Van: Regression methods for pricing complex American-style options. Working paper (2000)
15. Wilmott, P., Dewynne, J., Howison, S.: *Option Pricing: Mathematical Models and Computation*. Oxford Financial Press (1993)