Enhancing the Accuracy of Pricing American and Bermudan Options

PETER W. DUCK, DAVID P. NEWTON, MARTIN WIDDICKS, AND YAN LEUNG

PETER W. DUCK
is a professor of mathematics at the University of Manchester in Manchester, UK.
duck@ma.man.ac.uk

DAVID P. NEWTON
is a professor of finance at Nottingham University Business School in Nottingham, UK.
david.newton@nottingham.ac.uk

MARTIN WIDDICKS
is a senior lecturer at Manchester Business School in Manchester.
m.widdicks@mbs.ac.uk

YAN LEUNG
is a graduate student in mathematics at the University of Manchester.
yleung@maths.man.ac.uk

The basic Monte Carlo procedure for option pricing is appealing and simple to implement, and thus a popular tool for practitioners. Unmodified, though, it suffers from two serious shortcomings: a difficulty in treating early exercise, and only modestly accurate convergence characteristics.

The Monte Carlo-based method of handling early exercise features can be improved by generating monotonically varying data that are then amenable to extrapolation procedures, thereby enhancing the overall accuracy and reliability of the scheme. Results for options on one and two underlyings compare extremely favorably with exact results for early exercise options. In a final test of the technique, results are computed for an option on five underlyings.

The simple Monte Carlo technique to price financial derivatives can be extended to handle quite complex forms of options such as Asian or lookback options. An especially attractive feature is its ability to cope with multidimensional problems without the exponential increase in computation load experienced with lattice or finite-difference techniques.

Yet Monte Carlo schemes have two severe drawbacks. The first is a difficulty in treating early exercise. As Wilmott [1998] says, applying Monte Carlo methods to American-style options "is very, very hard." More recently Glasserman [2004] notes that "any general method for pricing American options by simulation requires substantial computational effort."

The second difficulty with Monte Carlo methods is poor accuracy and convergence characteristics. While it is possible to obtain crude results with ease, refinement is difficult and accuracy hard to assess; generally errors diminish slowly, at a rate inversely proportional to the square root of the number of simulations, which can be a serious practical concern if more precise derivative values are required.

We use quite simple means to address the dual questions of early exercise and accuracy enhancement.

There has been considerable research into developing Monte Carlo schemes that can deal with early exercise features. A theoretically robust technique is the duality approach in which tight upper and lower bounds are calculated simultaneously (see Rogers [2002], Andersen and Broadie [2004], and Haugh and Kogan [2004]). The most popular methods from a practitioner's point of view are generally based on regression techniques (Ibanez and Zapatero [2004]). As Glasserman [2004] points out, the regression approach is reasonably fast and broadly applicable.

The use of regression to estimate continuation values from simulated paths in order to price American options via simulation has been considered by Carrière [1996], Tsitsiklis and Van Roy [1999, 2001], and Longstaff and Schwartz [2001], who summarize the previous work in this area (see also Broadie and Detemple [2004]). Its accuracy depends on the functions chosen for the regression, but this flexibility in choice allows for some exploitation of features particular to a pricing problem.

The principal idea behind the Longstaff and Schwartz [2001] method is to generate (forward in time) a set of random paths, each with a number of exercise times. Working backward from expiration, at each observation time, a least squares fit (based on a selected set of polynomials) is used to fit the stochastic payoff data, and it is then decided whether it is advantageous to exercise the option.

Convergence of the Longstaff and Schwartz method has been investigated by Clément, Lamberton, and Protter [2002] and Rogers [2002], who prove convergence to the true price, given that the expression for continuation value as the sum of a set of basis functions holds exactly. Moreno and Navas [2003] provide a practical assessment of the method and address issues with regard to the basis functions.

Although the Longstaff and Schwartz [2001] method has proven successful and is accordingly popular, it does suffer from the usual Monte Carlo slow convergence characteristics. We aim to improve the performance of the scheme, thereby making the procedure a more convenient and useful tool for practitioners.

I. METHODOLOGY

Consider N underlying assets, where a change in the i-th asset satisfies the stochastic equation:

$$dS_i = \mu_i S_i dt + \sigma_i S_i dX_i, \tag{1}$$

where μ_i and σ_i are the drift and volatility of this i-th asset, and dX_i is the increment of a Wiener process (drawn from a normal distribution with mean zero and standard deviation $\sqrt{dt}$). The i-th and j-th random processes are correlated as follows

$$E\left[dX_i dX_j\right] = \rho_{ij} dt \tag{2}$$

where ρ_{ij} is the correlation factor between the i-th and j-th random walks.

This produces a (symmetric) correlation matrix. In the case of five underlying assets ($N = 5$), it takes the form:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} & \rho_{15} \\ \rho_{21} & 1 & \rho_{23} & \rho_{24} & \rho_{25} \\ \rho_{31} & \rho_{32} & 1 & \rho_{34} & \rho_{35} \\ \rho_{41} & \rho_{42} & \rho_{43} & 1 & \rho_{45} \\ \rho_{51} & \rho_{52} & \rho_{53} & \rho_{54} & 1 \end{pmatrix} \tag{3}$$

where $\rho_{ij} = \rho_{ji}$, and $\rho_{ii} = 1$. It is then possible to write:

$$dX_i = \phi_i \sqrt{dt} \tag{4}$$

where ϕ_i is a random variable drawn from a normalized normal distribution. Using (4) and risk neutrality, (1) can be integrated exactly over a time step δt to yield:

$$S_i(t + \delta t) = S_i(t) \exp\left((r - \tfrac{1}{2}\sigma_i^2)\delta t + \sigma_i \phi_i \sqrt{\delta t}\right) \tag{5}$$

where $S_i(t)$ is the value of the asset i after time t, and r is the riskless interest rate (assumed constant). From (2), it is then clear that

$$E\left[\phi_i \phi_j\right] = \rho_{ij} = \rho_{ji} \tag{6}$$

It is routine to generate the correlated random variables $\phi = (\phi_1, \phi_2, \ldots, \phi_N)'$ from a set of uncorrelated random variables $\epsilon = (\phi_i, \ldots, \epsilon_1, \epsilon_2, \ldots, \epsilon_i, \ldots, \epsilon_N)'$ by writing:

$$\boldsymbol{\phi} = \mathbf{L}\boldsymbol{\epsilon} \tag{7}$$

where

$$\mathbf{L}\mathbf{L}' = \boldsymbol{\Sigma} \tag{8}$$

That is, $\mathbf{L}$ is the Cholesky factorization of $\boldsymbol{\Sigma}$ (see Wilmott [1998]) and a lower triangular matrix, and the prime denotes the transpose.

Consider now a put option at time $t = 0$ that can be exercised on M dates $t_1, t_2, \ldots, t_k, \ldots, t_M$, where expiration time is t_M. We assume that the exercise dates are equally spaced, so $t_k = k\delta t$, although it is a trivial matter to relax this condition. This type of option is generally referred to as Bermudan (or mid-Atlantic), and as $M \to \infty$ the option becomes American. General experience suggests that M need not be excessively large to give a good approximation of an American option (see Widdicks

[2002]). In any case, convergence to the true American price is at the rate of $1/M$, so extrapolation is possible, although preferably from smooth data.

The core of our method is based on extensions of the Longstaff and Schwartz [2001] method, which can be separated into a number of steps (and substeps), assuming a payoff of the form $\max(0, X_1 - S_1, \ldots, X_N - S_N)$, where X_i is the exercise price of the i-th asset:

1. The Cholesky factorization ($\mathbf{L}$) of the correlation matrix ($\mathbf{\Sigma}$) is evaluated.
2. At each time step, starting at $t = \delta t$:
   - N_{paths} sets of uncorrelated normally distributed variables ϵ are generated (this involves the generation of $N \times N_{\text{paths}}$ random numbers).
   - N_{paths} sets of correlated variables ϕ are generated using the above ϵ in (7); the p-th path set of the i-th underlying is denoted by ϕ_i^p, where $p = 1, 2, \ldots, N_{\text{paths}}$.
   - Then (5) is used (for each underlying, for each path) to extend the solution to the next value of t; this can be written in the form:

$$S_i^p(t+\delta t) = S_i^p(t)\exp\left((r - \tfrac{1}{2}\sigma_i^2)\delta t + \sigma_i \phi_i^p \sqrt{\delta t}\right) \quad (9)$$

3. At expiration ($t = t_M$), the payoffs of all paths ($p = 1, 2, \ldots, N_{\text{paths}}$) are evaluated as follows:

$$\text{Payoff}^p = \max(X_1 - S_1^p, X_2 - S_2^p, \ldots, X_i - S_i^p, \ldots, X_N - S_N^p, 0) \quad (10)$$

   The cash flow (at expiration) is then written as Cash Flow$^p(t_M)$ = Payoff^p.
4. The cash flow from the payoff for each of the paths is discounted back δt in time as follows:

$$\text{DCashflow}^p(t - \delta t) = \text{Cashflow}^p(t)e^{-r\delta t} \quad (11)$$

   By its very nature, DCash Flow^p $(t - \delta t)$ will be stochastic. Following Longstaff and Schwartz [2001], this is approximated by a least squares fit to the data points that are in the money, denoted by N^+_{paths} that will vary at each time observation.
   - Some experimentation involving various classes of polynomials suggests that Chebyshev polynomials T_{m_i} (corresponding to the i-th underlying) are an excellent choice for this purpose. This can be confirmed theoretically (see Atkinson [1989]). As a consequence, (11) is approximated as follows:

$$\text{DCashflow}^p(t-\delta t) \approx \mathcal{D}^p =$$

$$\sum_{m_1=0}^{J_1}\sum_{m_2=0}^{J_2}\cdots\sum_{m_i=0}^{J_i}\cdots\sum_{m_N=0}^{J_N}\kappa(m_1, m_2, \ldots, m_i, \ldots, m_N)\times$$

$$a_{m_1,m_2,\ldots,m_i,\ldots,m_N} T_{m_1}(s_1^p)T_{m_2}(s_2^p)\ldots T_{m_i}(s_i^p)\ldots T_{m_N}(s_N^p) \quad (12)$$

   where J_i is the order of the i-th Chebyshev polynomial, the a_{m_1, m_2, m_i, m_N} are coefficients to be determined, and

$$s_i^p = \frac{2S_i^p - S_i^{\min} - S_i^{\max}}{S_i^{\max} - S_i^{\min}} \quad (13)$$

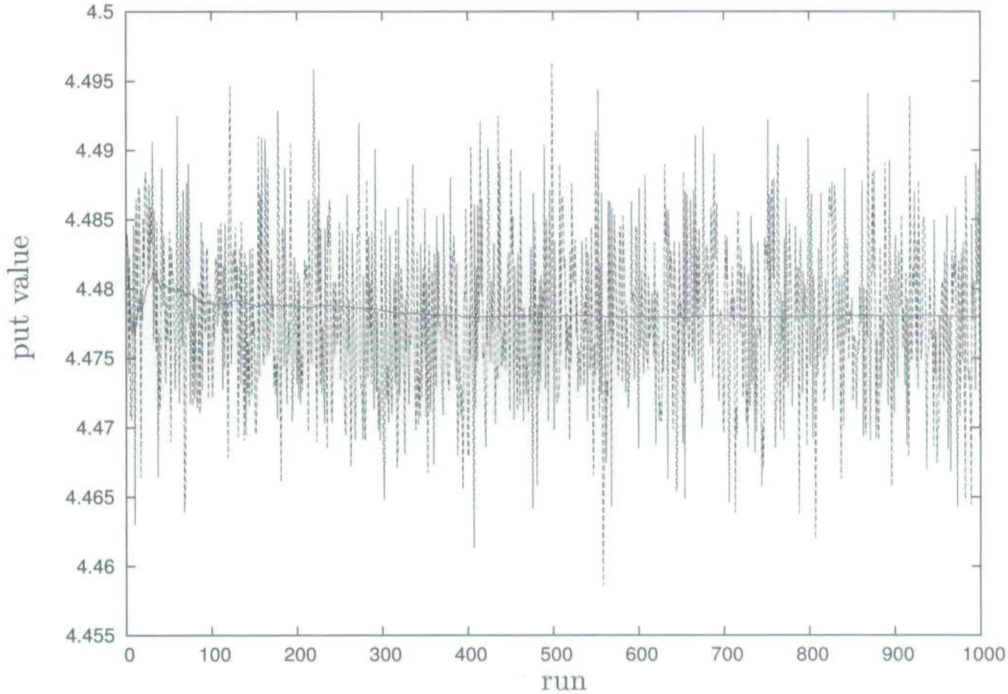
   where S_i^{min} and S_i^{max} are the minimum/maximum (respectively) values of the i-th underlying of all the in-the-money paths at time $t - \delta t$. This rescaling is necessary on account of the standard restrictions to the argument of Chebyshev polynomials (see, for example, Abramowitz and Stegun [1972]). Generally the choice $J_1 = J_2 = \ldots = J_i = \ldots = J_N = J$ is made, together with a further truncation procedure that eliminates unnecessary computational tasks involving high-order terms, namely:

$$\kappa(m_1, m_2, \ldots, m_i, \ldots, m_N) = 1 \quad \text{if} \quad \sum_{i=1}^{i=N} m_i \le \mathcal{F} \quad (14)$$

$$\kappa(m_1, m_2, \ldots, m_i, \ldots, m_N) = 0 \quad \text{if} \quad \sum_{i=1}^{i=N} m_i > \mathcal{F} \quad (15)$$

   where $\mathcal{F}$ is a prescribed trimming parameter.
   - A least squares fit is then used to determine the $a_{m_1, m_2, \ldots, m_i, \ldots, m_N}$
5. For each path that is notionally in the money, comparison is made to determine if it is best to exercise [i.e., if $X_i - S_i^p > \mathcal{D}^p$ for the p-th path, in which case the cash flow for this path becomes Cash Flow$^p(t - \delta t) = \max(X_i - S_i^p)$] or to wait, in which case the cash flow for the p-th path is naturally Cash Flow$^p(t - \delta t)$ = Dcashflow$^p(t - \delta t)$.
6. The procedure is repeated from step 4 until the first exercise time (δt) is reached, and then the value at $t = 0$ may be easily evaluated.

EXHIBIT 1
Computed and Averaged Values of Put Option



$N_{paths} = 10^5$, $J = 6$, $X = 40$, $S = 36$, $t_M = 1$, $M = 50$ time observations..

Steps 2 through 6 constitute one run or simulation that involves N_{paths}.

As an example of the implementation of this basic procedure (and to illustrate the difficulties encountered), we consider a one-dimensional ($N = 1$) example: $S_1 = 36$, $X_1 = 40$, $\sigma = 0.20$, $r = 0.06$, and $t_M = 1$, with $M = 50$ exercise dates. This corresponds to one case given in Table 1 of Longstaff and Schwartz [2001]. The highly accurate scheme of Andricopoulos et al. [2003], based on a quadrature method yields a put value for this Bermudan class of option of 4.477811109.

Exhibit 1 shows the results of the basic procedure, obtained using 100,000 paths (i.e., $N_{\text{paths}} = 10^5$) and $J = 6$, for 1,000 independent runs. The stochastic nature of each run is clearly observed, but the averaged value of the option, evaluated over the current and *all* preceding runs, also strongly suggests it asymptotes to a fixed value as the number of runs increases; indeed, this averaged value gives a reasonable estimate of the option value.

Exhibit 2 shows the put values averaged over the current and preceding runs, for $N_{\text{paths}} = 2{,}000$, 4,000, and 8,000. It should be clear from these results that the apparent (asymptotic) value, as the number of runs increases, does depend strongly on the number of paths, N_{paths}.

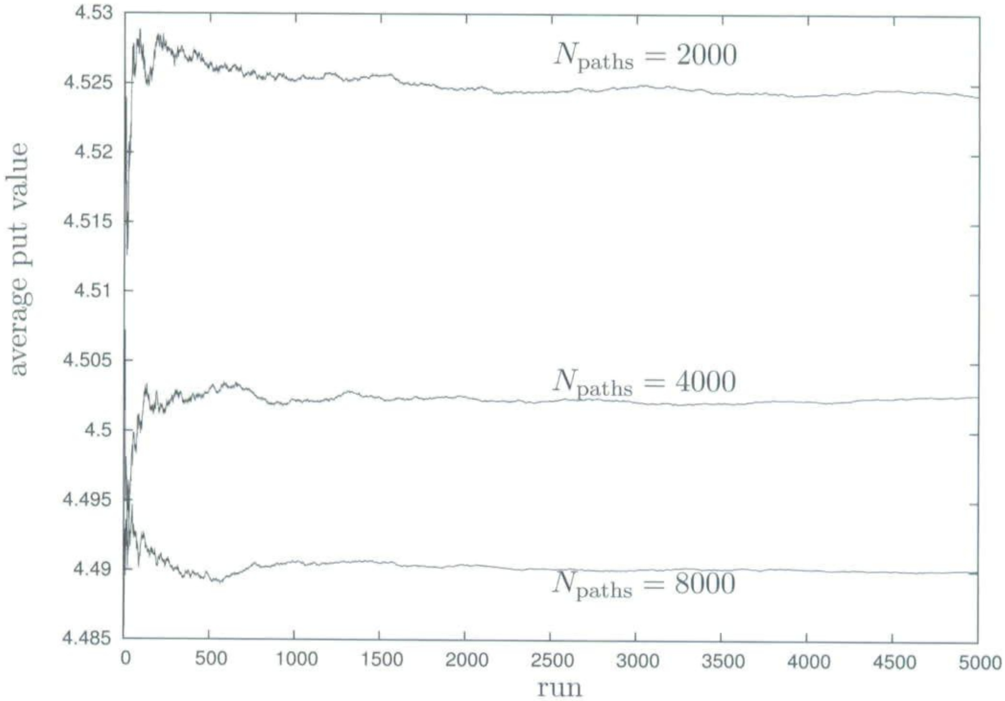
While these asymptotic values are not expected to produce the exact value for the option price, there is a strong suggestion that, as N_{paths} is increased, the values tend monotonically toward the exact value. It turns out this is a universal feature of our procedure. Exhibit 2 shows a clear bias in the results as N_{paths} increases (as runs $\to \infty$).

Garcia [2003] and Haugh and Kogan [2004] address the question of biasing in Monte Carlo methods as applied to pricing early exercise options. This bias is a central issue for us, as it implies a monotonicity in results as N_{paths} increases, thus admitting the possibility of extrapolation procedures.

The familiar inverse-square root error incurred in standard Monte Carlo simulations arises through an application of the central limit theorem (see, for example, Wilmott [1998] and Higham [2004]). This suggests that a reasonable formula to describe the variation of the calculated option value, $P(\text{runs}; N_{\text{paths}})$, with increasing numbers of paths (in the limit of runs $\to \infty$) is:

$$P(\text{runs} \to \infty; N_{\text{paths}}) \approx P(\text{exact}) + \frac{\alpha_1}{\sqrt{N_{\text{paths}}}} + \frac{\alpha_2}{N_{\text{paths}}} \quad (16)$$

EXHIBIT 2
Averaged Values of Put Option



N_{paths} = 2,000, 4,000, 8,000, J = 6, X = 40, S = 36, t_M = 1, M = 50 time observations..

This equation may be viewed as a polynomial (quadratic) fit (in terms of $N_{\text{paths}}^{-\frac{1}{2}}$) to the data [$P(\text{runs} \to \infty; N_{\text{paths}})$]. Even if the functional form of $P(\text{runs} \to \infty; N_{\text{paths}})$ is not of precisely this form, the formula will capture the value of $P(\text{exact})N_{\text{paths}}^{\frac{1}{2}}$ as N_{paths} increases. To enable effective fitting in powers of

$$N_{\text{paths}}^{-\frac{1}{2}}$$

the sole requirement is monotonicity of $P(\text{runs} \to \infty; N_{\text{paths}})$ with variations in N_{paths}. This was seen in every calculation we undertook. Further, this monotonic convergence was seen always to occur from *above*.

Our extrapolation procedure entails three separate sets of simulations, each with different values of N_{paths}. After each run, the averaged put value was evaluated, and (16) was applied, using the resulting three equations to solve for $P_\infty = P(\text{exact})$ and for α_1 and α_2, although these are of more limited interest. This procedure has some similarities with the extrapolation procedure proposed by Widdicks et al. [2002] as applied to binomial tree calculations.

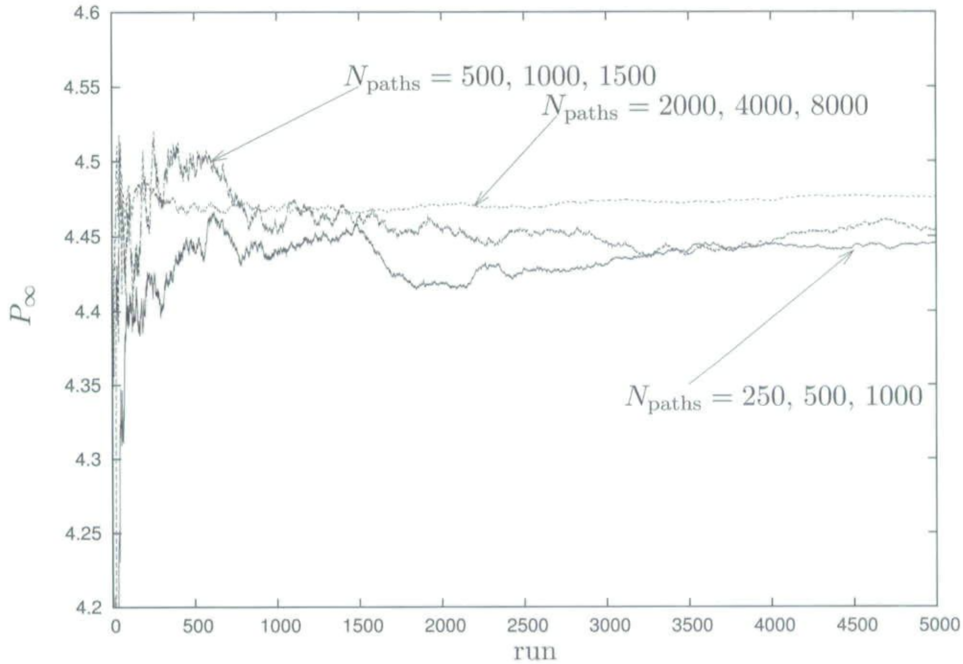
Results for various combinations of N_{paths} are shown in Exhibit 3 (all were obtained with $J = 6$, although the results obtained are rather insensitive to other choices of this numerical parameter). Inspection of Exhibit 3 does suggest that our extrapolation procedure indeed yields accurate and reliable results (compare, for example, the unextrapolated results of Exhibit 2 with the extrapolated results from N_{paths} 2,000, 4,000, and 8,000 in Exhibit 3).

It should also be clear from Exhibit 3 that increasing the number of N_{paths} generally improves the performance of the overall scheme, while at the same time requiring fewer runs or simulations. There is clearly a trade-off between minimizing the number of N_{paths} per simulation and minimizing the numbers of simulations necessary to achieve the desired accuracy.

Exhibit 4 shows root mean square (RMS) percentage errors for eight choices of numerical parameters, based on eight cases—all the permutations of the financial parameters: $S = 36$ and $S = 38$, $X = 38$ and $X = 40$, and $\sigma = 0.2$ and $\sigma = 0.4$. It also shows, for comparison, the corresponding RMS values obtained from the highest N_{paths} (in brackets).

We see that: 1) extrapolated values are generally more accurate than the results from the most N_{paths}; 2) as N_{paths} increases, generally the results settle down after fewer runs; and 3) there is surprisingly little difference between the $J = 4$ and $J = 6$ results as N_{paths} increases.

EXHIBIT 3
Extrapolated Values of Put Option



N_{paths} as shown, $J = 6$, $X = 40$, $S = 36$, $t_M = 1$, $M = 50$ time observations (exact value 4.4778).

EXHIBIT 4
Numerical Parameters and RMS % Errors

set	J	N_{paths}	rms % error (50)	rms % error (500)	rms % error (5000)	time
1	4	250, 500, 1000	3.65 [2.88]	0.86 [2.89]	0.59 [2.89]	1195
2	4	500, 1000, 1500	4.12 [2.24]	1.63 [2.24]	0.49 [2.22]	1409
3	4	2000, 4000, 8000	1.36 [0.93]	0.53 [0.96]	0.50 [0.96]	3433
4	6	250, 500, 1000	3.85 [3.69]	0.63 [3.66]	0.74 [3.63]	1225
5	6	500, 1000, 1500	3.26 [2.88]	1.49 [2.74]	0.55 [2.75]	1458
6	6	2000, 4000, 8000	1.75 [1.11]	0.77 [1.09]	0.53 [1.09]	3648
7	6	25000, 50000, 100000	0.66 [0.65]	0.59 [0.65]	-	5115

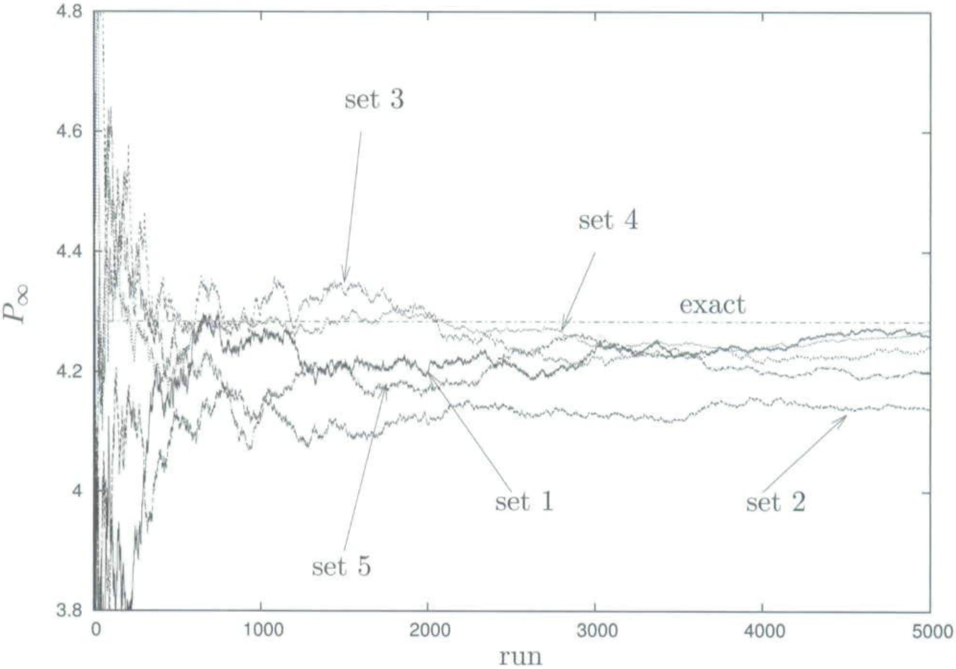
Eight 1D test cases, after 50, 500, 5,000 runs (50 time observations). Corresponding errors with highest N_{paths} shown in brackets. Time indicates total run time in seconds on a 2.5GHz Pentium IV for one test case, 5,000 runs (500 runs for set 7).

The computation times (as indicated in Exhibit 4) are reduced approximately linearly with N_{paths} (although this is not quite the case for lower N_{paths}), and of course linearly with the number of runs. The power and usefulness of the technique is shown vividly by set 4, which involves less than 10% of the number of N_{paths} as set 7 (and fewer polynomials), but achieves comparable accuracy, even after only 500 simulations.

A final point is that the averaging process *even without* extrapolation does appear to enhance the performance of the basic method very significantly, when used with a reasonable number of N_{paths}.

The overall scheme is more accurate and more reliable, and thus more efficient than the original scheme of Longstaff and Schwartz [2001], while at the same time not much more difficult to use.

EXHIBIT 5
Computed Values of 2D Put Option



$X_1 = X_2 = 100$, $S_1 = S_2 = 96$, $t_M = 0.5$, $\sigma_1 = \sigma_2 = \sigma_{12} = 0.2$, $r = 0.06$, $M = 6$ time observations (exact value = 4.2840).

II. MULTIPLE ASSETS

The principle with regard to multiple assets is much the same as for the single-asset case, although the calculations become more challenging as the number of underlyings (the dimension of the problem) increases. Consider first a two-underlying ($N = 2$) put valuation problem: $S_1 = S_2 = 100$, $X_1 = X_2 = 96$, $\sigma_1 = \sigma_2 = 0.2$, $\rho_{12} = \rho_{21} = 0.2$, $r = 0.06$, and $t_M = 0.5$ with $M = 6$ (equally spaced observations, i.e., monthly). The highly accurate multi-asset quadrature procedure of Andricopoulos et al. [2005] values this option as 4.2840.

Exhibit 5 shows results obtained using the numerical parameters in Exhibit 6. These five choices of numerical parameters show the effects of choosing both differing numbers of polynomials (compare set 1 and set 3) and differing numbers of paths (compare set 1 with set 2, and set 3 with set 4). Even with the coarsest choice of grid (set 2), the percentage error (computed by comparing the averaged over 5,000 runs with the so-called exact value) is encouraging.

All the indications are that this two-dimensional example is more challenging than the one-dimensional example (compare the solution variation as the number of runs is increased in Exhibit 3 with that in Exhibit 5). Further evidence may be seen in comparison of Exhibit 4 with Exhibit 6 (especially when comparing like choices of N_{paths}).

A final example is a five-dimensional underlying case, with 12 equally spaced time observations over a period of one year (i.e., monthly observations), with all $S_i = 36$, all $X_i = 40$, all $\sigma_i = 0.2$, $r = 0.06$, and the correlation matrix:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0.1 & 0.2 & 0 & 0 \\ 0.1 & 1 & -0.3 & 0 & 0 \\ 0.2 & -0.3 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{17}$$

Given the challenging nature of the problem, which is quite a severe test of the scheme, a comprehensive and diverse set of numerical parameters is chosen to assess performance. Results for this case are tabulated in Exhibit 7 and graphed in Exhibit 8. In this case, no previous (accurate) results are available, so this exhibit shows absolute

EXHIBIT 6
Numerical Parameters and % Error—2D Test Case

set	J	N_{paths}	% error (50)	% error (500)	% error (5000)
set 1	4	4000, 6000, 8000	13.34	2.69	0.53
set 2	4	250, 500, 1000	8.04	3.66	3.37
set 3	6	4000, 6000, 8000	16.06	0.18	0.96
set 4	6	16000, 24000, 32000	1.60	1.62	0.28
set 5	6	250, 500, 1000	1.76	3.22	1.98

50, 500, 5,000 runs, 6 time observations, $\mathcal{F} = 2J$ (no trimming).

EXHIBIT 7
Numerical Parameters—5D Case

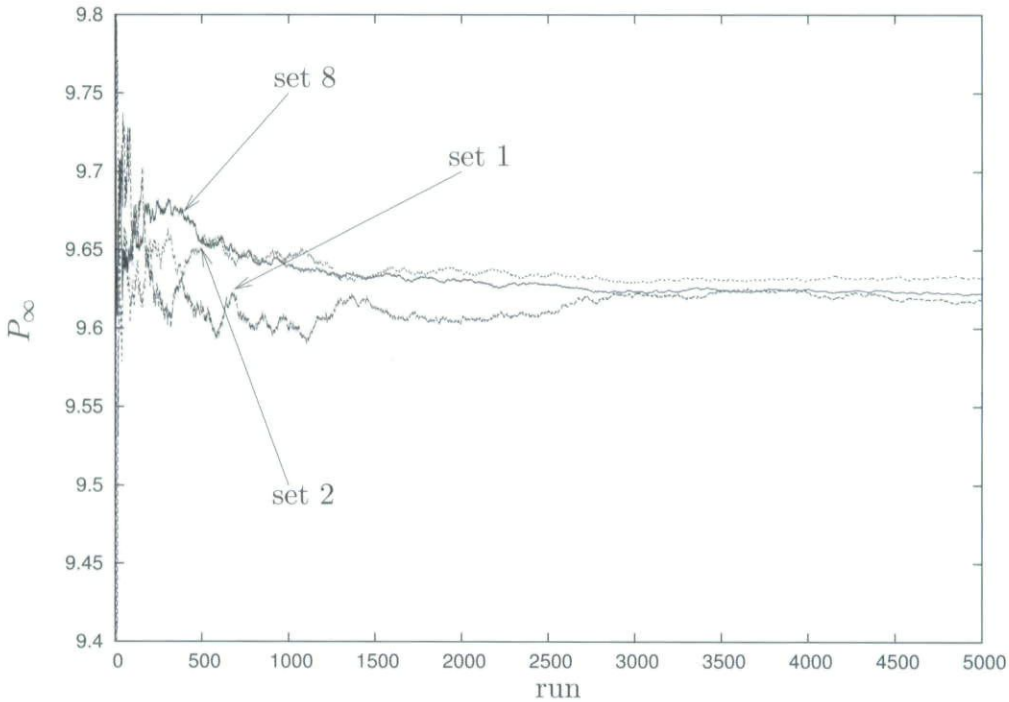
set	J	$\mathcal{F}$	N_{paths}	$P(50)$	$P(500)$	$P(5000)$
set 1	3	3	1000, 2000, 4000	9.723 [9.684]	9.610 [9.680]	9.618 [9.681]
set 2	3	3	2000, 4000, 8000	9.656 [9.664]	9.654 [9.658]	9.633 [9.656]
set 3	6	6	1000, 2000, 4000	9.384 [10.098]	9.389 [10.101]	9.376 [10.100]
set 4	6	6	2000, 4000, 8000	9.558 [9.910]	9.527 [9.908]	9.528 [9.907]
set 5	4	5	1000, 2000, 4000	9.645 [9.910]	9.493 [9.904]	9.518 [9.905]
set 6	4	4	1000, 2000, 4000	9.591 [9.908]	9.505 [9.908]	9.530 [9.906]
set 7	4	4	2000, 4000, 8000	9.535 [9.788]	9.583 [9.790]	9.608 [9.792]
set 8	4	4	4000, 8000, 16000	9.649 [9.724]	9.657 [9.729]	9.623 [9.727]

50, 500, 5,000 runs (highest N_{paths} averaged values shown in brackets).

computed values. Computed values are shown after 50, 500, and 5,000 runs. Close inspection of Exhibit 7 supports some general conclusions:

- Results become more accurate and more reliable as the number of runs increases.
- Results become more accurate and reliable as the number of N_{paths} increases.
- As the order of the polynomial (J) increases, higher values of N_{paths} and more runs are necessary to achieve accuracy (this could be because higher-order polynomials more faithfully replicate the stochastic nature of the results, which is not necessarily a desirable feature; see Moreno and Navas [2003]).
- Detailed numerical experimentation revealed that the trimming parameter ($\mathcal{F}$) has little effect on the numerical results (and much less effect than the choice of J). It is worth pointing out that Equation (12) notionally involves $(J + 1)^N$ terms, which in the case of the five-dimensional computations (sets 3 and 4) involves 16,807 terms in the series. In the 5D case, the trimming parameter ($\mathcal{F}$) was vital to reduce the amount of (unnecessary and lengthy) computation; generally setting $\mathcal{F} = J$ proved sufficient. In the 5D computations (sets 3 and 4), this reduced the number of terms in Equation (12) to a much more manageable 462. This trimming implies the neglect of higher-order Chebyshev terms for higher-order terms. So, for example, if $m_1 = J$, only $m_2 = m_3 = \ldots = 0$ terms were retained, while for $m_1 < J$ more of the other m_i terms were retained, provided the trimming condition

EXHIBIT 8
Computed Values of 5D Put Option



$S_i = 36$, $X_i = 40$, $T = 1.0$, $\sigma_i = 0.2$, $r = 0.06$, 12 time observations.

$\sum_{i=1}^{i=N} m_i \leq \mathcal{F}$

was satisfied.

- The extrapolated results show improved accuracy compared to the averaged value of the highest N_{paths} value (shown again in brackets).
- This higher-dimensional case generally requires more N_{paths} than the 1D and 2D cases.
- For the case computed, the value of this option is within a penny of 9.63.

Note as a further gauge of the effectiveness of the overall procedure, that set 8 took the equivalent of 20 days' computing (on a Pentium IV 2.5 GHz processor), while the corresponding set 2 computation took approximately 27 hours. Of course, running computations for 5,000 runs may be regarded as overkill. Approximately 1,000 runs are generally sufficient, as suggested by Exhibit 5, which shows how the (extrapolated) computed values vary with the number of simulations (for sets 1, 2, and 8).

III. CONCLUSIONS

We have demonstrated a simple yet powerful means to enhance the performance and use of Monte Carlo methods for early exercise options. The technique greatly improves the efficiency, accuracy, and reliability of the basic Longstaff and Schwartz [2001] scheme. It comes into its own as the number of underlyings increases, proving highly capable even in five-dimensional problems.

We have some recommendations on choices of parameters:

1. Approximately 1,000 runs or simulations should be sufficient; very high values of N_{paths} require fewer runs, but 1,000 is considered safe for even modest values of N_{paths}.
2. In the one-dimensional case, the coarsest (set 1) results (Exhibit 4) are perfectly adequate, and yield comparable results as with eight times the number of N_{paths} and three times the amount of computation time (set 3); this performance is mirrored in the set 4 and set 6 results (the corresponding results using $J = 6$).

3. Five Chebyshev polynomials ($J = 4$) should be adequate. Taking higher-order polynomials may necessitate the use of higher values of N_{paths}, because the least squares fit replicates the stochastic variation in data more faithfully.
4. Bear in mind that the choices of the particular values of N_{paths} should be reasonably well separated—choosing values that are too close is likely to diminish accuracy, as small differences in computed values are likely to be dominated by stochastic variations. In our examples, generally the three values of N_{paths} differ by a factor of two, and this appears to work well. Numerous other computations (including sets of N_{paths} that differ by 1,000) also prove a satisfactory choice.
5. For multidimensional calculations, generally higher values of N_{paths} are necessary (compared to the 1D case). For the 5D calculations (the most stringent tackled), the set of N_{paths} = 2,000, 4,000, 8,000 proves effective.
6. For multidimensional calculations, a trimming parameter $\mathcal{F} = J$ should be adequate.

For four or more distinct values of N_{paths} (rather than three) and correspondingly additional terms in the fit (16), the results show no significant improvement over the three-term expansion but longer computation times. Indeed, if more terms are taken in (16), the number of runs must be increased in order to retain consistent accuracy.

Note that *all* the calculations undertaken, both one- and multidimensional, indicate a monotonic variation of the asymptotic results (runs $\to \infty$) with increasing N_{paths}—indeed, this accounts for the success of the method. A better understanding of this behavior would be extremely useful as, if the parametric form of put value (runs $\to \infty$; N_{paths}) were known, a more precise form of this relationship could be invoked to further improve convergence properties

All our numerical results very strongly point to a conclusion that we can achieve convergence of results (to the exact value) under our technique. Indeed, if the Longstaff and Schwartz [2001] procedure converges as the numerical parameters are improved, then so too must our procedure, which is fundamentally based on Longstaff and Schwartz. Every computation undertaken indicates convergence to the apparent exact value from above; that is, these calculations all produce monotonically declining values as N_{paths} is increased (and, of course, runs $\to \infty$).

Analysis of other classes of early exercise options, including calls on dividend-paying underlyings and options on futures, also exhibited this monotonic decline. Such options are therefore amenable to the techniques we have described, indicating a promising degree of universality.

ENDNOTE

Professor Duck gratefully acknowledges funding from the Leverhulme Trust.

REFERENCES

Abramowitz, W., and I. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* New York: Dover, 1972.

Andersen, L., and M. Broadie. "A Primal-Dual Simulation Algorithm for Pricing Multi-Dimensional American Options." *Management Science*, 50 (2004), pp. 1222-1234.

Andricopoulos, A.D., M. Widdicks, P.W. Duck, and D.P. Newton. "Extending Quadrature Methods to Value Multi-Asset and Complex Path Dependent Options." Forthcoming, *Journal of Financial Economics*, 2005.

——. "Universal Option Valuation Using Quadrature Methods." *Journal of Financial Economics,* 67 (2003), pp. 447-471.

Atkinson, K.E. *An Introduction to Numerical Analysis.* New York: John Wiley, 1989.

Broadie, M., and J.B. Detemple. "Option Pricing: Valuation Models and Applications." *Management Science*, 50 (2004), pp. 1145-1177.

Carrière, J. "Valuation of Early-Exercise Price of Options Using Simulations and Nonparametric Regression." *Insurance: Mathematics and Economics*, 19 (1996), pp. 19-30.

Clément, E., D. Lamberton, and P. Protter. "An Analysis of a Least Squares Regression Algorithm of American Option Pricing." *Finance and Stochastics*, 6 (2002), pp. 449-471.

Garcia, D. "Convergence and Biases of Monte Carlo Estimates of American Option Prices Using a Parametric Exercise Rule." *Journal of Economic Dynamics and Control*, 27 (2003), pp. 1855-1879.

Glasserman, P. *Monte Carlo Methods in Financial Engineering.* New York: Springer, 2004.

Haugh, M., and L. Kogan. "Pricing American Options: A Duality Approach." *Operations Research*, 52 (2004), pp. 258-270.

Higham, D.J. *An Introduction to Financial Option Valuation*. Cambridge: Cambridge University Press, 2004.

Ibanez, A., and F. Zapatero. "Monte Carlo Valuation of American Options Through Computation of the Optimal Exercise Frontier." *Journal of Financial and Quantitative Analysis*, 39 (2004), pp. 253-275.

Longstaff, F.A. and E.S. Schwartz. "Valuing American Options by Simulation: A Simple Least-Squares Approach." *The Review of Financial Studies*, 14 (2001), pp. 113-147.

Moreno, M., and J.F. Navas. "On the Robustness of Least-Squares Monte Carlo (LSM) for Pricing American Derivatives." *Review of Derivatives Research*, 6 (2003), pp. 107-128.

Rogers, C. "Monte Carlo Valuation of American Options." *Mathematical Finance*, 12 (2002), pp. 271-286.

Tsitsiklis, J., and B. Van Roy. "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 (1999), pp. 1840-1851.

——. "Regression Methods for Pricing Complex American-Style Options." *IEEE Transactions on Neural Networks*, 12 (2001), pp. 694-703.

Widdicks, M. "Examination, Extension and Creation of Methods for Pricing Options with Early Exercise Features." Ph.D. thesis, University of Manchester, 2002.

Widdicks, M., A.D. Andricopoulos, D.P. Newton, and P.W. Duck. "On the Enhanced Convergence of Standard Lattice Methods for Option Pricing." *Journal of Futures Markets,* 22 (2002), pp. 315-338.

Wilmott, P. *Derivatives*. New York: John Wiley & Sons, 1998.