

# The evaluation of American compound option prices under stochastic volatility and stochastic interest rates

**Carl Chiarella**

Finance Discipline Group, University of Technology, Sydney, PO Box 123, Broadway, NSW 2007, Australia; email: carl.chiarella@uts.edu.au

**Boda Kang**

Finance Discipline Group, University of Technology, Sydney, PO Box 123, Broadway, NSW 2007, Australia; email: boda.kang@uts.edu.au

(Received June 10, 2009; revised May 2, 2011; accepted May 18, 2011)

*A compound option (the mother option) gives the holder the right, but not the obligation, to buy (long) or sell (short) the underlying option (the daughter option). In this paper, we consider the problem of pricing American-type compound options when the underlying dynamics follow Heston's stochastic volatility and with stochastic interest rate driven by Cox–Ingersoll–Ross processes. We use a partial differential equation (PDE) approach to obtain a numerical solution. The problem is formulated as the solution to a two-pass free-boundary PDE problem, which is solved via a sparse grid approach and is found to be accurate and efficient compared with the results from a benchmark solution based on a least-squares Monte Carlo simulation combined with the projected successive over-relaxation method.*

## 1 INTRODUCTION

The compound option goes back to the seminal paper of Black and Scholes (1973). As well as their famous pricing formulas for vanilla European call and put options, they also considered how to evaluate the equity of a company that has coupon bonds outstanding. They argued that the equity can be viewed as a “compound option” because the equity “is an option on an option on ... an option on the firm”. Geske (1979) developed the first closed-form solution for the price of a vanilla European call on a European call. It turns out that a wide variety of important problems are closely related to the valuation of compound options. Some examples include pricing American puts in Geske and Johnson (1984) and hedging volatility risk by trading options on straddles in Brenner *et al* (2006).

A compound option (called the mother option) gives the holder the right, but not the obligation, to buy (long) or sell (short) the underlying option (called the daughter option). For simplicity, we describe the European-type compound option as an example. Suppose that a compound option expires at some date  $T_M$  with the strike price  $K_M$  and the daughter option, on which it is contingent, expires at a later time  $T_D (> T_M)$  with the strike price  $K_D$ . Under geometric Brownian motion dynamics, the price of a European put on a European put (or a put on put for short),  $M(S, t)$ , may be written as a conditional expectation under the risk-neutral measure of the discounted payoff at the maturity of the mother option where the payoff is the positive part of the differences between the price of the daughter option at that time and the strike of the mother option. Similarly, we can define put on call, call on put and call on call.

Compound options are common in many multiphase projects, such as product and drug development, where the initiation of one phase of the project depends on the successful completion of the preceding phase. For example, launching a product that involves a new technology requires successful testing of the technology; drug approval is dependent on successful Phase II trials, which can be conducted only after successful Phase I tests. With compound options, at the end of each phase, we have the option to continue to the next phase, abandon the project or defer it to a later time. Each phase becomes an option that is contingent upon the exercise of earlier options. For phased projects, two or more phases may occur at the same time (parallel options) or in sequence (staged or sequential options). These options are mostly American with the right to buy (call) or sell (put) on or before the expiry of each option. We refer the reader to Kodukula and Papudesu (2006) for more examples of compound options in real option applications.

Derivative securities are commonly written on underlying assets with return dynamics that are not sufficiently well described by the geometric Brownian motion (GBM) process proposed by Black and Scholes (1973). There have been numerous efforts to develop alternative asset return models that are capable of capturing the leptokurtic features found in financial market data, and subsequently to use these models to develop option prices that better reflect the volatility smiles and skews found in market traded options. One of the classical ways to develop option pricing models that are capable of generating such behavior is to allow the volatility to evolve stochastically, for instance according to the square-root process introduced by Heston (1993). Since we consider the pricing of American-type options, the early exercise premium of the option depends on the cost of carry determined by interest rates. Consequently, the volatility of interest rates does affect the decision to exercise this option at any given point in time. Hence the compound options of the type that we consider in this paper are very sensitive not only to the volatility of the underlying but also to the risk-free interest rate, and this is the motivation for considering American-type compound options under stochastic volatility and stochastic interest rates.

Han (2003) and Fouque and Han (2005) introduce a fast, efficient and robust approximation to compute the prices of compound options such as call-on-call options within the context of multiscale stochastic volatility models. However, they only consider the case of a European option on a European option. Furthermore, their method relies on certain expansions, so its range of validity is not entirely clear.

To the best of our knowledge, there is no literature studying the compound option pricing problem under both stochastic volatility and stochastic interest rates, but some authors do discuss the American option pricing problem under these dynamics. Boyarchenko and Levendorski (2007) formulate the option pricing problem by a partial differential equation (PDE) approach and they calculate the option prices with the help of an iteration method based on Wiener–Hopf factorization. Medvedev and Scaillet (2010) introduce a new analytical approach. After using an explicit and intuitive proxy for the exercise rule, they derive tractable pricing formulas using a short-maturity asymptotic expansion. Depending on model parameters, this method can accurately price options with time-to-maturity up to several years.

In the case of European options on European options under GBM dynamics, there exist “almost” explicit integral-form solutions. However, in situations involving more general dynamics (such as stochastic volatility), either explicit solutions do not exist or the integrals become difficult to evaluate. In contrast it turns out that the PDE approach provides a very efficient and flexible way to compute prices of compound options. The use of this approach is not restricted to European-type options, and can also include American-type, Asian-type or other exotic types of options.

In this paper, we demonstrate the PDE approach to the problem of pricing American-type compound options. We assume that both the mother and the daughter options may be American-type. The American meaning of the mother option is the same as the conventional American option, namely the holder of the compound option can exercise the mother option at any time before the maturity  $T_M$ . Upon exercising the mother option, the holder will hold a daughter option from the early exercise time and we assume that the holder can exercise the daughter option any time from then until the maturity  $T_D$ . In principle, the method we develop is able to price all four kinds of compound options, but here we only give details of the case of an American put on an American put.

The remainder of the paper is structured as follows. Section 2 outlines the problem of an American-type compound option, where the underlying asset follows stochastic volatility and stochastic interest rate dynamics. In Section 3 we outline the basic idea of the sparse grid approach and implement a combination technique on a sparse grid to find the price profile of a daughter option and apply the same technique to find the price profile of the mother option based on the results from the previous step. A number of numerical examples that demonstrate the computational advantages of

the sparse grid approach are provided in Section 4. We draw some conclusions in Section 5.

## 2 PROBLEM STATEMENT-COMPOUND OPTION WITH STOCHASTIC VOLATILITY AND STOCHASTIC INTEREST RATES

Let  $D(S, v, r, t)$  denote the price of an American put option (the daughter option) written on a stock of price  $S$  at time  $t$  with maturity time  $T_D$  and strike price  $K_D$ , and let  $M(S, v, r, t)$  denote the price of an American put option written on the daughter option of price  $D(S, v, r, t)$  with maturity time  $T_M (< T_D)$  and strike price  $K_M$ . The variables  $v$  and  $r$  denote the variance of the stock price return and the risk-free rate at time  $t$ , respectively.

Analogously to the setting in Heston (1993) with in addition a stochastic interest rate of the Cox–Ingersoll–Ross (CIR) type, the dynamics for the share price  $S$  under the risk-neutral measure are governed by the stochastic differential equation (SDE) system<sup>1</sup>

$$dS = (r - q)S dt + \sqrt{v}S dZ_1, \quad (2.1)$$

$$dv = \kappa_v(\theta_v - v) dt + \sigma_v \sqrt{v} dZ_2, \quad (2.2)$$

$$dr = \kappa_r(\theta_r - r) dt + \sigma_r \sqrt{r} dZ_3, \quad (2.3)$$

where  $Z_1$ ,  $Z_2$  and  $Z_3$  are standard Wiener processes and  $\mathbb{E}(dZ_i dZ_j) = \rho_{ij} dt$ ,  $i = 1, 2, j = i + 1, \dots, 3$ , with  $\mathbb{E}$  being the expectation operator under the risk-neutral measure. In (2.1),  $r$  is the risk-free rate of interest and  $q$  is the continuously compounded dividend yield. In (2.2) the parameter  $\sigma_v$  is the so called vol-of-vol (in fact,  $\sigma_v^2 v$  is the variance of the variance process  $v$ ). The parameters  $\kappa_v$  and  $\theta_v$  are respectively the rate of mean reversion and long run variance of the process for the variance  $v$ . In (2.3) the parameter  $\sigma_r$  is the volatility of the interest rate process (in fact,  $\sigma_r^2 r$  is the variance of the interest rate process  $r$ ). The parameters  $\kappa_r$  and  $\theta_r$  are respectively the rate of mean reversion and long run interest rate of the process for the interest rate  $r$ .

These parameters are under the risk-neutral measure and are related to the corresponding quantities under the physical measure (that we denote as  $\kappa_v^{\mathbb{P}}$ ,  $\theta_v^{\mathbb{P}}$ ,  $\kappa_r^{\mathbb{P}}$  and  $\theta_r^{\mathbb{P}}$ )

---

<sup>1</sup> Of course, since we are using a numerical technique we could in fact use more general processes for  $S$  and  $v$ . The choice of the Heston processes is driven partly by the fact that this has become a very traditional stochastic volatility model and partly because the transform methods (which are used to derive the benchmark prices by the Fourier cosine expansion method) do not easily handle the more general variance processes.

by two parameters that appear in the market prices of both volatility risk and interest rate risk.<sup>2</sup>

We are also able to write down the above system (2.1)–(2.3) using independent Wiener processes  $W_1$ ,  $W_2$  and  $W_3$  so that

$$\begin{pmatrix} dZ_1 \\ dZ_2 \\ dZ_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \rho_{12} & \sqrt{1 - \rho_{12}^2} & 0 \\ \rho_{13} & \frac{\rho_{23} - \rho_{13}\rho_{12}}{\sqrt{1 - \rho_{12}^2}} & \sqrt{1 - \rho_{13}^2 - \left(\frac{\rho_{23} - \rho_{13}\rho_{12}}{\sqrt{1 - \rho_{12}^2}}\right)^2} \end{pmatrix} \begin{pmatrix} dW_1 \\ dW_2 \\ dW_3 \end{pmatrix}.$$

The price of an American compound option under stochastic volatility at time  $t$ ,  $M(S, v, r, t)$ , can be formulated as the solution to a two-pass free-boundary PDE problem. We first solve the PDE for the value of the daughter option  $D(S, v, r, t)$  given by

$$\mathcal{K}D - rD + \frac{\partial D}{\partial t} = 0, \tag{2.4}$$

on the interval  $0 \leq t \leq T_D$  and subject to the terminal condition

$$D(S, v, r, T_D) = (K_D - S)^+, \tag{2.5}$$

free (early exercise) boundary condition

$$D(d(v, r, t), v, r, t) = K_D - d(v, r, t), \tag{2.6}$$

and the smooth-pasting conditions

$$\lim_{S \rightarrow d(v, r, t)} \frac{\partial D}{\partial S} = -1, \quad \lim_{S \rightarrow d(v, r, t)} \frac{\partial D}{\partial v} = 0, \quad \lim_{S \rightarrow d(v, r, t)} \frac{\partial D}{\partial r} = 0, \tag{2.7}$$

where  $S = d(v, r, t)$  is the early exercise boundary for the daughter option at time  $t$ , variance  $v$  and interest rate  $r$ .

---

<sup>2</sup> In fact, if it is assumed that the market prices of risk associated with the uncertainty driving the variance process and the interest rate process have the form  $\lambda_v \sqrt{v}$  and  $\lambda_r \sqrt{r}$ , respectively, where  $\lambda_v$  is a constant (this was the assumption in Heston (1993)) and  $\lambda_r$  is a constant, then

$$\begin{aligned} \kappa_v &= \kappa_v^{\mathbb{P}} + \lambda_v \sigma_v, & \theta_v &= \frac{\kappa_v^{\mathbb{P}} \theta_v^{\mathbb{P}}}{\kappa_v^{\mathbb{P}} + \lambda_v \sigma_v}, \\ \kappa_r &= \kappa_r^{\mathbb{P}} + \lambda_r \sigma_r, & \theta_r &= \frac{\kappa_r^{\mathbb{P}} \theta_r^{\mathbb{P}}}{\kappa_r^{\mathbb{P}} + \lambda_r \sigma_r}, \end{aligned}$$

where  $\kappa_v^{\mathbb{P}}$ ,  $\theta_v^{\mathbb{P}}$  and  $\kappa_r^{\mathbb{P}}$ ,  $\theta_r^{\mathbb{P}}$  are the corresponding parameters under the physical measure.

In (2.4) the Kolmogorov operator  $\mathcal{K}$  is given by

$$\begin{aligned} \mathcal{K} = & \frac{vS^2}{2} \frac{\partial^2}{\partial S^2} + \frac{\sigma_v^2 v}{2} \frac{\partial^2}{\partial v^2} + \frac{\sigma_r^2 r}{2} \frac{\partial^2}{\partial r^2} \\ & + \rho_{12} \sigma_v v S \frac{\partial^2}{\partial S \partial v} + \rho_{13} \sigma_r \sqrt{r} S \frac{\partial^2}{\partial S \partial r} + \rho_{23} \sigma_v \sigma_r \sqrt{vr} \frac{\partial^2}{\partial v \partial r} \\ & + (\kappa_r(\theta_r - r) - \lambda_r r) \frac{\partial}{\partial r} + (r - q) S \frac{\partial}{\partial S} + (\kappa_v(\theta_v - v) - \lambda_v v) \frac{\partial}{\partial v}, \end{aligned} \quad (2.8)$$

where  $\lambda_v$  and  $\lambda_r$  are the constants appearing in the equation for the market prices of volatility risk and interest rate risk, which as stated in footnote 2 are assumed to be of the form  $\lambda_v \sqrt{v}$  and  $\lambda_r \sqrt{r}$ , respectively.

Given the values for the daughter option, we can then solve the PDE for the mother option  $M(S, v, r, t)$  that satisfies

$$\mathcal{K} M - rM + \frac{\partial M}{\partial t} = 0, \quad (2.9)$$

on the interval  $0 \leq t \leq T_M$  and subject to the terminal condition

$$M(S, v, r, T_M) = (K_M - D(S, v, r, T_M))^+, \quad (2.10)$$

the free (early exercise) boundary condition

$$M(m(v, r, t), v, r, t) = K_M - D(m(v, r, t), v, r, t), \quad (2.11)$$

and the smooth-pasting conditions

$$\lim_{S \rightarrow m(v, r, t)} \frac{\partial M}{\partial S} = -\frac{\partial D}{\partial S}, \quad \lim_{S \rightarrow m(v, r, t)} \frac{\partial M}{\partial v} = -\frac{\partial D}{\partial v}, \quad \lim_{S \rightarrow m(v, r, t)} \frac{\partial M}{\partial r} = -\frac{\partial D}{\partial r}, \quad (2.12)$$

where  $m(v, r, t)$  is the early exercise boundary for the mother option at variance  $v$ , interest rate  $r$  and time  $t$ .

### 3 SPARSE GRID IMPLEMENTATION

In order to tackle the computationally demanding task of solving the two nested PDEs (2.4)–(2.6) and (2.9)–(2.11) we apply the sparse grid approach, which turns out to be quite fast and accurate. The sparse grid combination technique for solving PDEs was first introduced by Griebel *et al* (1992), after which Reisinger (2004), Reisinger and Wittum (2007) and Leentvaar and Oosterlee (2008a,b) discussed the application of this approach to various option pricing problems. The combination technique requires the solution of the original equation only on a set of conventional subspaces defined on Cartesian grids specified in a certain way and a subsequent extrapolation step, but still retains a certain order convergence.

In fact we can identify three desirable properties of the combined solution. First, in comparison with the standard full grid approach the number of grid points can be reduced significantly from  $O(2^{n \cdot d})$  to  $O(2^n \cdot n^{d-1})$  at refinement level  $n$  in the  $d$ -dimensional case, whereas the point-wise accuracy of the approximation to the solution of the PDE is  $O(n^{d-1} \cdot 2^{-n \cdot p})$ , which is only slightly worse than  $O(2^{-n \cdot p})$ . Here,  $p$  includes the order of the underlying discretization scheme, as well as the influence of singularities. Furthermore, each of the Cartesian grids setting up the sparse grid only consists of  $O(2^n)$  nodes. Thus, the efficient usage of sparse grids for the computational solution of the PDE greatly reduces storage requirements and computing time at a moderate cost of accuracy.

Second, we have to point out the simplicity of the combination concept: we have seen that the sparse grid combined solution represents a linear combination of numerical solutions on Cartesian grids corresponding to the components of a sparse grid at the same refinement level. Thus, the combination technique allows for the integration of existing solvers for partial differential equations on traditional full grids. In contrast to the discretization on a real sparse grid, which requires hierarchical data structures and thus specially designed solvers, the combined solution is built on simple data structures and can be based on any “black box solver”. Only the final linear combination of these simple solutions has to be newly implemented.

From the combined solution as a linear combination of traditional full grid discretizations we can also deduce a further advantage of the combination technique. Since the  $O(n^{d-1})$  problems solved on the Cartesian grids  $\Omega_l$  that set up the sparse grids are independent from one another, these problems can be solved in parallel on different workstations. Communication has to take place only at the end, when the summation and the extrapolation by linear combination of the different solutions is performed.

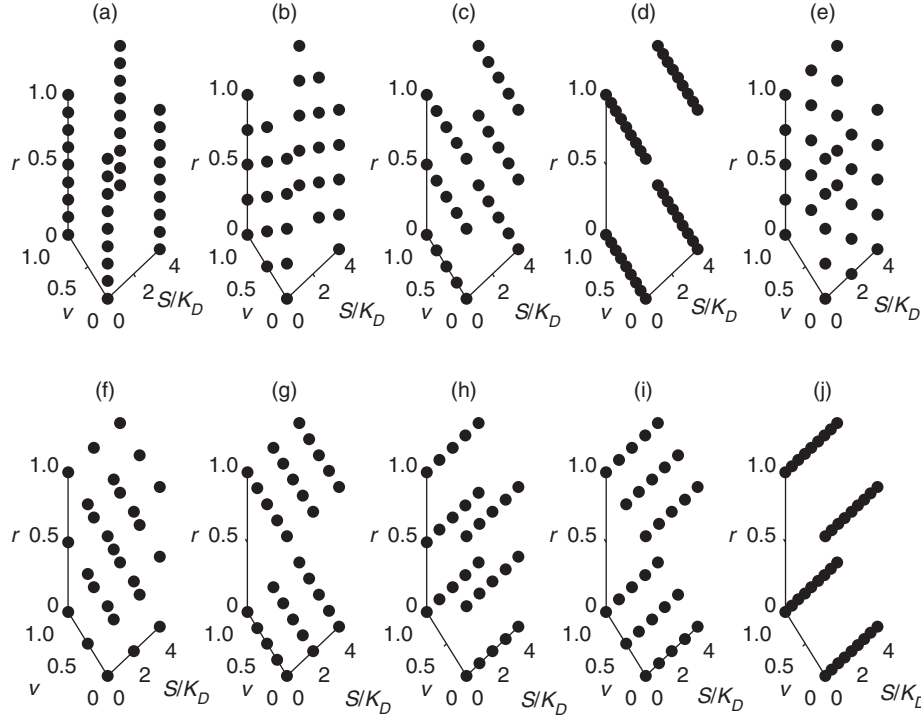
### 3.1 The sparse grid combination technique

We incorporate the techniques and algorithms used in Chiarella *et al* (2010) and the sparse grid approach to solve the linked PDEs (2.4)–(2.6) and (2.9)–(2.11) with suitable initial and boundary conditions.

In a general  $d$ -dimensional unit cube and the family of grids with grid sizes  $h_j = 2^{-l_j}$  in direction  $j$ ,  $l_j \in \mathbb{N}_0$ . We write the vector of grid sizes as  $\mathbf{h} = 2^{-\mathbf{l}}$  with  $\mathbf{l} = (l_1, \dots, l_d) \in \mathbb{N}_0^d$  and denote the solution of the PDE on those grids by  $p_{\mathbf{h}}$ . The sparse grid solution at level  $l$  is then defined as

$$p_l = \sum_{k=l}^{l+d-1} a_{l-k} \sum_{l_1+\dots+l_d=k} p_{\mathbf{h}}, \tag{3.1}$$

**FIGURE 1** A sparse grid hierarchy of level 3 with respect to each combination.



These are (a) (0,0,3), (b) (0,1,2), (c) (0,2,1), (d) (0,3,0), (e) (0,1,2), (f) (1,1,1), (g) (1,2,0), (h) (2,0,1), (i) (2,1,0), (j) (3,0,0).

with

$$a_k = (-1)^{d-1-k} \binom{d-1}{k}, \quad 0 \leq k \leq d-1. \quad (3.2)$$

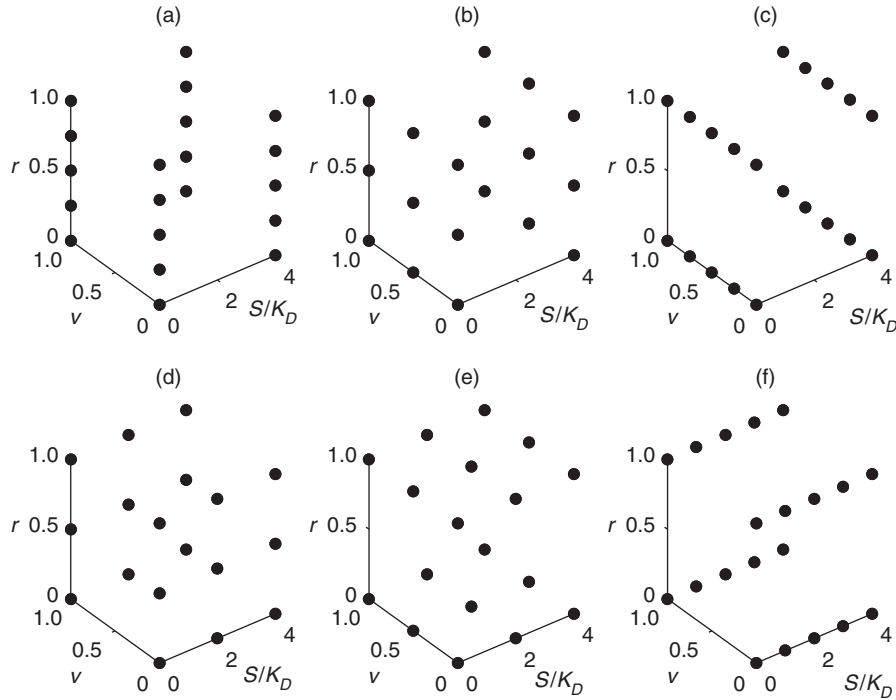
In our case  $d = 3$ . Hence, we consider a truncated three-dimensional cube  $\Omega := [0, S_{\max}] \times [0, v_{\max}] \times [0, r_{\max}]$  and a Cartesian grid with mesh size  $h_j = 2^{-l_j}$  (corresponding to a level  $l_j \in \mathbb{N}_0$ ) in the directions  $j = 1, 2, 3$ . The indices  $j = 1, j = 2$  and  $j = 3$  represent the directions of the stock price  $S$ , the variance  $v$  and the interest rate  $r$ , respectively.

For a vector  $\mathbf{h} = (h_1, h_2, h_3)$  we denote by  $p_{\mathbf{h}}$  the representation of a function on such a grid with points

$$\mathbf{x}_h = (i_1 \cdot h_1, i_2 \cdot h_2, i_3 \cdot h_3), \quad 1 \leq i_j \leq N_j, \quad N_j = 1/h_j = 2^{l_j}, \quad \text{for } j = 1, 2, 3.$$



**FIGURE 2** A sparse grid hierarchy of level 2 with respect to each combination.



These are (a) (0,0,2), (b) (0,1,1), (c) (0,2,0), (d) (1,0,1), (e) (1,1,0), (f) (2,0,0).

For a given level  $l$ , the above grid consists of all possible combinations of  $(l_1, l_2, l_3)$  with  $0 \leq l_1, l_2, l_3 \leq l$ . Hence, in total, there are  $(2^l + 1)^3$  points in the grid. The number of total points in the full grid increases significantly with the increase of the level  $l$ . It will be quite expensive to solve the two-pass PDE system on the above full grid.

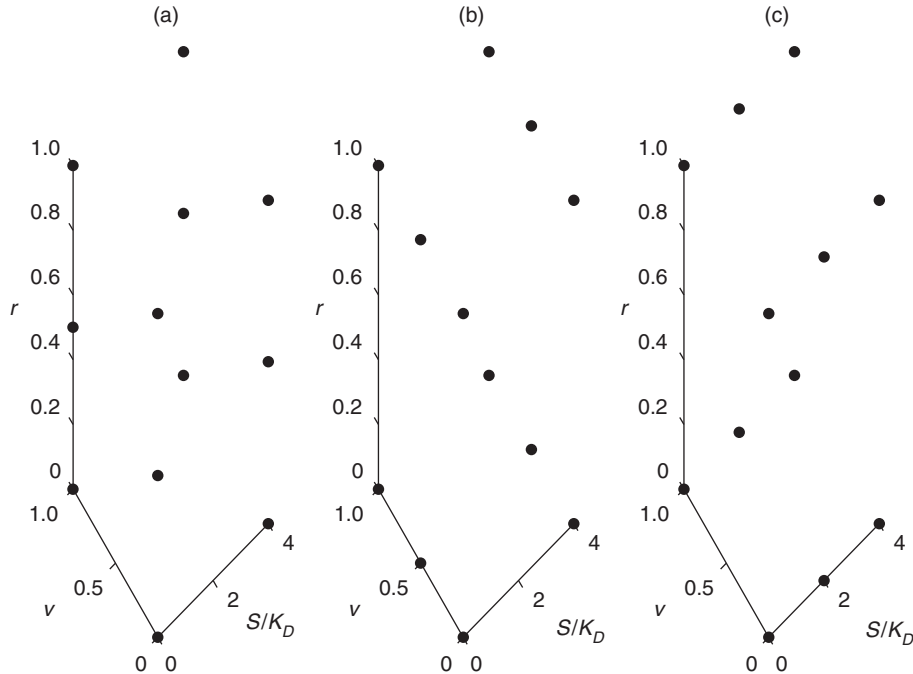
However, with the same level  $l$ , the sparse grid will consist of the following points:

$$\mathbf{x}_h = (i_1 \cdot h_1, i_2 \cdot h_2, i_3 \cdot h_3), \quad 1 \leq i_j \leq N_j, \quad N_j = 1/h_j = c_j 2^{l_j}, \quad \text{for } j = 1, 2, 3.$$

satisfying  $l_1 + l_2 + l_3 = l$  and where  $c_j$  are some positive constants with the help of which it is possible to construct a nonequidistant grid.

It is not hard to see that there are  $\binom{l+d-1}{d-1}$  choices of such combinations of  $(l_1, l_2, l_3)$  such that  $l_1 + l_2 + l_3 = l$ . Figure 1 on the facing page, Figure 2 and Figure 3 on the next page provide an example of a standard sparse grid hierarchy with level  $l = 3$ ,

**FIGURE 3** A sparse grid hierarchy of level 1 with respect to each combination.



These are (a) (0,0,1), (b) (0,1,0), (c) (1,0,0).

$l = 2$  and  $l = 1$  with respect to 10, 6 and 3 different combinations corresponding to each level, respectively.

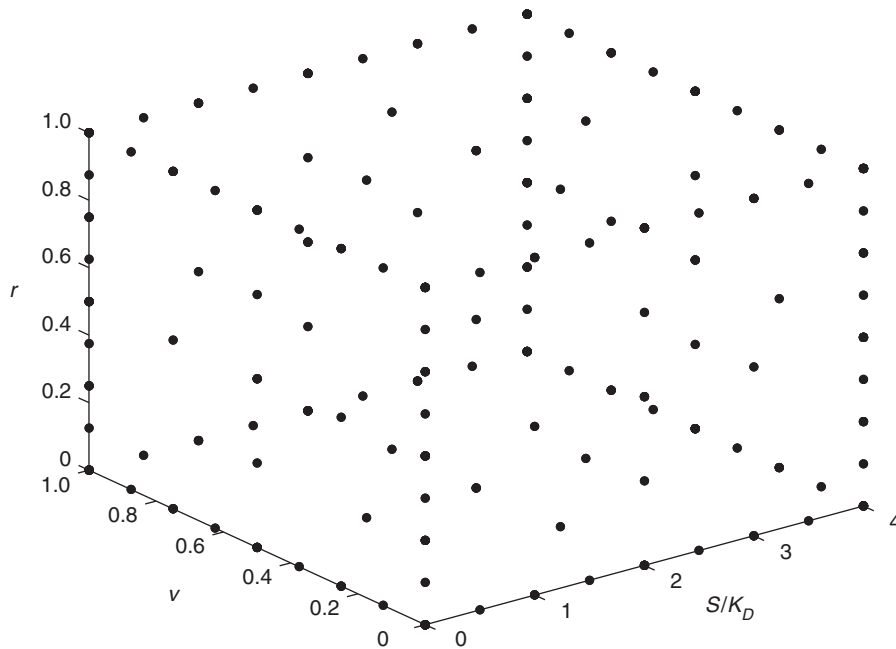
Obviously, the above grids share the common property that they are dense in one direction but sparse in the other directions. If we put all of the above grids together, we will obtain the standard sparse grid shown in Figure 4 on the facing page.

Let  $\mathbf{p}_h$  be the discrete vector of function values at the grid points of the standard sparse grid. In general,  $\mathbf{p}_h$  is the finite difference solution to the PDE of interest on the corresponding grid  $\mathbf{h}$ . The solution can be extended to  $\Omega$  by a suitable multilinear interpolation operator  $\mathcal{I}^3$  in the point wise sense according to

$$p_h(S, v, r, \tau) = \mathcal{I}^3 \mathbf{p}_h, \quad \forall (S, v, r) \in \Omega.$$

<sup>3</sup>A thorough error analysis of the multilinear interpolation operator can be found in Reisinger (2008), who gives a generic derivation for linear difference schemes through an error correction technique employing semidiscretizations and obtains error formulas as well.

**FIGURE 4** A combined sparse grid solution.



Next, we define the family  $P$  of solutions corresponding to the different sparse grids (as in Figure 1 on page 78 for instance) by  $P = (P(\mathbf{i}))_{\mathbf{i} \in \mathbb{N}^3}$  with

$$P(\mathbf{i}) := p_{2^{-\mathbf{i}}},$$

that is the family of numerical approximations (after proper interpolation)  $p_{\mathbf{h}}$  on tensor product grids with  $h_k = 2^{-i_k}$ . For example, the solution on the first grid in Figure 1 on page 78 would be  $P(0, 0, 3)$ , etc. The combination technique in Reisinger and Wittum (2007) tells us that the solution  $p_l$  ( $l$  is the level of the sparse grid) of the corresponding PDE is

$$p_l = \sum_{l_1+l_2+l_3=l} P(l_1, l_2, l_3) - 2 \sum_{l_1+l_2+l_3=l+1} P(l_1, l_2, l_3) + \sum_{l_1+l_2+l_3=l+2} P(l_1, l_2, l_3). \quad (3.3)$$

The above is a special case of (3.1) and (3.2) when  $d = 3$ .

The procedure involves solving the PDE in parallel on each of the sparse grids of level  $l$ ,  $l + 1$  and level  $l + 2$ , respectively. See Figure 1 on page 78, Figure 2 on page 79

and Figure 3 on page 80 for  $l = 1$  and  $d = 3$  as an example. Thus, we have

$$\sum_{k=l}^{l+2} \binom{k+d-1}{d-1} = \sum_{k=1}^3 \binom{k+2}{2} = 3 + 6 + 10 = 19,$$

which means that there will be 19 PDE solvers running simultaneously when  $l = 1$ . The theory developed by Reisinger and Wittum (2007) shows that (3.3) combines all solutions together to yield a more accurate solution to the PDE.

The essential principle of the extrapolation is that all lower order error terms cancel out in the combination formula (3.3) and only the highest order terms

$$h_1^2 \cdot h_2^2 \cdot h_3^2 = (2^{-l_1} \cdot 2^{-l_2} \cdot 2^{-l_3})^2 = 4^{-l}$$

remain. Taking advantage of this cancellation mechanism, (3.3) is able to produce quite accurate results fairly quickly. The details of the error analysis can be found in Reisinger (2004) and Reisinger and Wittum (2007).<sup>4</sup>

We implement the above sparse grid combination technique to solve the PDE (2.4) in order to obtain the desired daughter option prices. We thus have the terminal and boundary conditions for the PDE governing the price of the mother option. Next, we apply the technique again to solve the PDE (2.9) to obtain the prices of the mother option. We solve the PDEs (2.4)–(2.6) and (2.9)–(2.11) in each of the subspaces on a parallel cluster, which makes the process very efficient.

### 3.2 Finite difference method with PSOR

In the implementation, a standard Crank–Nicolson finite difference method with the projected successive over-relaxation (PSOR) method has been applied to each of the sparse grids in Figure 1 on page 78, Figure 2 on page 79 and Figure 3 on page 80 to work out the solution of both PDEs (2.4)–(2.6) and (2.9)–(2.11) on the grid points; solutions at other nongrid points are obtained by the same multilinear interpolation as in Reisinger (2008). The implementation of PSOR is detailed in this section. As an example, we show the detail of solving the PDE followed by the daughter option prices. We apply the same approach to the PDE followed by prices of the mother option with different terminal and boundary conditions. The boundary conditions for the mother option are quite similar to those for daughter options, hence for brevity, we only mention boundary conditions for the daughter option.

---

<sup>4</sup>The combination method requires, theoretically, smoothness of mixed derivatives of the solution. This is obviously not the case here due to the nonsmooth payoff. However, if the payoff is aligned with the grid, which is the case in our problem, then good results have been observed for the combination method (see Leentvaar and Oosterlee 2008b). This is probably due to the rapid smoothing of the payoff in the first few time steps.

It is convenient to consider the time-to-maturity  $\tau = T - t$  instead of time  $t$ . The three space variables  $S, v, r$  and time-to-maturity  $\tau$  are discretized according to

$$\begin{aligned} S_i &= (i - 1) \cdot \Delta S, & i &= 1, \dots, N_1 + 1, \\ v_j &= (j - 1) \cdot \Delta v, & j &= 1, \dots, N_2 + 1, \\ r_k &= (k - 1) \cdot \Delta r, & k &= 1, \dots, N_3 + 1, \\ \tau_l &= T - l \Delta t, & l &= 1, \dots, N_\tau, \end{aligned}$$

where  $N_1, N_2, N_3$  and  $N_\tau$  are the number of grid points in the directions  $S, v, r$  and  $\tau$ , respectively.

The option prices at the discrete points are thus

$$D_{i,j,k}^l = D(S_i, v_j, r_k, \tau_l).$$

Similar to the discussions in Ekstrom *et al* (2009), we use central differences to approximate most of the first and second derivatives in the PDE but use the forward and backward finite difference approximations on the boundaries other than the time derivative in (2.4). Thus, we set

$$\begin{aligned} \frac{\partial D}{\partial S} &= \frac{D_{i+1,j,k}^l - D_{i-1,j,k}^l}{2\Delta S}, & \frac{\partial^2 D}{\partial S^2} &= \frac{D_{i+1,j,k}^l - 2D_{i,j,k}^l + D_{i-1,j,k}^l}{\Delta S^2}, \\ \frac{\partial D}{\partial v} &= \frac{D_{i,j+1,k}^l - D_{i,j-1,k}^l}{2\Delta v}, & \frac{\partial^2 D}{\partial v^2} &= \frac{D_{i,j+1,k}^l - 2D_{i,j,k}^l + D_{i,j-1,k}^l}{\Delta v^2}, \\ \frac{\partial D}{\partial r} &= \frac{D_{i,j,k+1}^l - D_{i,j,k-1}^l}{2\Delta r}, & \frac{\partial^2 D}{\partial r^2} &= \frac{D_{i,j,k+1}^l - 2D_{i,j,k}^l + D_{i,j,k-1}^l}{\Delta r^2}, \end{aligned}$$

and

$$\begin{aligned} \frac{\partial D}{\partial S} \Big|_{S=S_1} &= \frac{D_{1,j,k}^l - D_{0,j,k}^l}{\Delta S}, & \frac{\partial D}{\partial S} \Big|_{S=S_{N_1}} &= \frac{D_{N_1,j,k}^l - D_{N_1-1,j,k}^l}{\Delta S}, \\ \frac{\partial D}{\partial v} \Big|_{v=v_1} &= \frac{D_{i,1,k}^l - D_{i,0,k}^l}{\Delta v}, & \frac{\partial D}{\partial v} \Big|_{v=v_{N_2}} &= \frac{D_{i,N_2,k}^l - D_{i,N_2-1,k}^l}{\Delta v}, \\ \frac{\partial D}{\partial r} \Big|_{r=r_1} &= \frac{D_{i,j,1}^l - D_{i,j,0}^l}{\Delta r}, & \frac{\partial D}{\partial r} \Big|_{r=r_{N_3}} &= \frac{D_{i,j,N_3}^l - D_{i,j,N_3-1}^l}{\Delta r}. \end{aligned}$$

Similar to Boyarchenko and Levendorski (2007), we use the boundary conditions  $\partial^2 D / \partial v^2 = 0$  at the boundaries  $v = 0$  and  $v = v_{\max}$  and boundary conditions  $\partial^2 D / \partial r^2 = 0$  at the boundaries  $r = 0$  and  $r = r_{\max}$ .<sup>5</sup> The discretized analogs for all

<sup>5</sup>The Fichera theory for degenerate parabolic equation suggests that the complete equation (2.4) should hold on  $v = 0$  and  $r = 0$  instead of  $\partial^2 D / \partial^2 v = 0$  and  $\partial^2 D / \partial^2 r = 0$  (see the explanation in Meyer (2013)). However the benchmark prices indicate that the influence of these boundary conditions on the solution is small far enough away from the boundary.

$i = 1, \dots, N_1 + 1$  are

$$\begin{aligned} D_{i,1,k}^l &= 2 \cdot D_{i,2,k}^l - D_{i,3,k}^l, D_{i,N_2+1,k}^l \\ &= 2 \cdot D_{i,N_2,k}^l - D_{i,N_2-1,k}^l, \quad \forall k = 1, \dots, N_3 + 1, \\ D_{i,j,1}^l &= 2 \cdot D_{i,j,2}^l - D_{i,j,3}^l, D_{i,j,N_3+1}^l \\ &= 2 \cdot D_{i,j,N_3}^l - D_{i,j,N_3-1}^l, \quad \forall j = 1, \dots, N_2 + 1, \end{aligned}$$

which is essentially an extrapolation scheme.

All other derivative terms in the  $(v, r)$ -direction vanish at  $v = 0$  or  $r = 0$  due to the factor  $(v, r)$  occurring in (2.4); hence, these terms do not require further treatment.

We follow Ikonen and Toivanen (2007) to indicate which grid point values we use to approximate the second-order mixed derivative in order to obtain nonpositive off-diagonal weights in the finite difference stencil, which makes the matrix an  $M$  matrix as much as possible.

In fact, to simplify the algorithm and take consideration of the correlations  $\rho_{ij}$ , in each two dimensional space, we use a seven-point stencil and the mixed derivatives when  $\rho_{12} > 0$  are approximated as

$$\begin{aligned} \frac{\partial D^2}{\partial S \partial v} &\approx \frac{1}{2} \left( \frac{D_{i+1,j+1,k}^l - D_{i,j+1,k}^l - (D_{i+1,j,k}^l - D_{i,j,k}^l)}{\Delta S \Delta v} \right. \\ &\quad \left. + \frac{D_{i,j,k}^l - D_{i-1,j,k}^l - (D_{i,j-1,k}^l - D_{i-1,j-1,k}^l)}{\Delta S \Delta v} \right), \end{aligned}$$

and when  $\rho_{12} < 0$  we have

$$\begin{aligned} \frac{\partial D^2}{\partial S \partial v} &\approx \frac{1}{2} \left( \frac{D_{i,j+1,k}^l - D_{i-1,j+1,k}^l - (D_{i,j,k}^l - D_{i-1,j,k}^l)}{\Delta S \Delta v} \right. \\ &\quad \left. + \frac{D_{i+1,j,k}^l - D_{i,j,k}^l - (D_{i+1,j-1,k}^l - D_{i,j-1,k}^l)}{\Delta S \Delta v} \right). \end{aligned}$$

The other mixed derivatives  $\partial^2/\partial S \partial r$  and  $\partial^2/\partial v \partial r$  are handled in a similar way, depending on the sign of the corresponding correlations.

Boundary conditions at the boundaries  $S = 0$  and  $S = S_{\max}$  for a put option are

$$\begin{aligned} D_{1,j,k}^l &= K_D \cdot e^{-(k-1) \cdot \Delta r \cdot \tau_l}, \quad D_{N_1+1,j,k}^l = 0, \\ &\quad \forall j = 1, \dots, N_2 + 1, k = 1, \dots, N_3 + 1; \end{aligned}$$

while those for a call option would be set as

$$\begin{aligned} D_{1,j,k}^l &= 0, \quad D_{N_1+1,j,k}^l = S_{\max} - K_D \cdot e^{-(k-1) \cdot \Delta r \cdot \tau_l}, \\ &\quad \forall j = 1, \dots, N_2 + 1, k = 1, \dots, N_3 + 1. \end{aligned}$$

The spatial discretization above leads to a semidiscrete equation which has the matrix representation

$$\frac{\partial \mathbf{D}}{\partial \tau} + \mathbf{A}\mathbf{D} = 0 \quad (3.4)$$

where  $\mathbf{A}$  is a block tridiagonal  $(N_1 + 1)(N_2 + 1)(N_3 + 1) \times (N_1 + 1)(N_2 + 1)(N_3 + 1)$  matrix and  $\mathbf{D}$  is a vector of length  $(N_1 + 1)(N_2 + 1)(N_3 + 1)$ .

Next, we implement a more general  $\theta$  scheme, which includes the implicit ( $\theta = 1$ ), the Crank–Nicolson ( $\theta = \frac{1}{2}$ ) and the explicit ( $\theta = 0$ ) approaches to discretize the semidiscrete problem (3.4) as

$$(\mathbf{I} + \theta \Delta \tau \mathbf{A})\mathbf{D}^{(l+1)} = (\mathbf{I} - (1 - \theta)\Delta \tau \mathbf{A})\mathbf{D}^{(l)}, \quad l = 0, \dots, N_\tau - 1, \quad (3.5)$$

where  $N_\tau$  is the number of time steps and  $\mathbf{I}$  is the identity matrix.

After the discretization of the underlying PDE with three spatial variables an approximate price of an American option can be obtained by solving a sequence of linear complementarity problems (LCPs)

$$\left. \begin{aligned} \mathbf{B}\mathbf{D}^{(l+1)} &\geq \mathbf{E}\mathbf{D}^{(l)}, & \mathbf{D}^{(l+1)} &\geq \mathbf{g}, \\ (\mathbf{B}\mathbf{D}^{(l+1)} - \mathbf{E}\mathbf{D}^{(l)})^\top (\mathbf{D}^{(l+1)} - \mathbf{g}) &= 0, \end{aligned} \right\} \quad (3.6)$$

for  $l = 0, \dots, N_\tau - 1$ . The matrixes  $\mathbf{B}$  and  $\mathbf{E}$  in (3.6) are defined by the left-hand side and right-hand side of (3.5), respectively. The initial value  $\mathbf{D}^{(0)}$  is given by the discrete form  $\mathbf{g}$  of the payoff function  $g$  of the option, so that the  $i$ th element of  $\mathbf{D}^{(0)}$  is given by

$$\mathbf{D}_i^{(0)} = \max(K - S_i, 0). \quad (3.7)$$

In order to avoid the oscillations that often occur with the CN scheme, we use the implicit Euler scheme ( $\theta = 1$ ) for the first three time steps and then switch to the CN scheme ( $\theta = \frac{1}{2}$ ) in the rest of the time steps. We implemented a PSOR finite difference scheme to solve the sequence (3.6) of LCPs efficiently.

In order to accelerate the convergence of the PSOR, we need to select the over-relaxation parameter  $\omega$  in the algorithm (see Kwok 2008, Section 6.2.3). We note that we may not choose the same  $\omega$  when we apply PSOR to different discretized grids in Figure 1 on page 78. Table 2 on the next page demonstrates how the optimal  $\omega$  should be chosen for different grids on a specific numerical example.

## 4 NUMERICAL EXAMPLES

To demonstrate the performance of the sparse grid algorithm outlined in Section 3 we implement the method for a given set of parameter values. For comparison purposes,

**TABLE 1** Parameter values used for the American put daughter option.

Option parameter	Value	SV parameter	Value	SI parameters	Value
$q$	0.0	$\theta_v$	0.02	$\theta_r$	0.04
$T_D$	1.0	$\kappa_v$	1.5	$\kappa_r$	0.3
$K_D$	100	$\sigma_v$	0.15	$\sigma_r$	0.1
$T_M$	0.50	$\lambda_v$	0.00	$\lambda_r$	0.00
$K_M$	4	$\rho_{12}$	-0.50	$\rho_{13}, \rho_{23}$	0.0

The stochastic volatility (SV) and stochastic interest rate (SI) parameters are those used in Medvedev and Scaillet (2010) to facilitate comparisons.

**TABLE 2** Values of  $\omega$  for different sparse grids are used for the American put daughter option when  $l = 6$ .

$(l_1, l_2, l_3)$	$\omega$	$(l_1, l_2, l_3)$	$\omega$	$(l_1, l_2, l_3)$	$\omega$	$(l_1, l_2, l_3)$	$\omega$	$(l_1, l_2, l_3)$	$\omega$
(0, 0, 6)	1.6	(0, 1, 5)	1.3	(0, 2, 4)	1.1	(0, 3, 3)	1.1	(0, 4, 2)	1.3
(0, 5, 1)	1.6	(0, 6, 0)	1.8	(1, 0, 5)	1.3	(1, 1, 4)	1.1	(1, 2, 3)	1.1
(1, 3, 2)	1.1	(1, 4, 1)	1.3	(1, 5, 0)	1.6	(2, 0, 4)	1.1	(2, 1, 3)	1.1
(2, 2, 2)	1.1	(2, 3, 1)	1.1	(2, 4, 0)	1.3	(3, 0, 3)	1.1	(3, 1, 2)	1.1
(3, 2, 1)	1.1	(3, 3, 0)	1.1	(4, 0, 2)	1.2	(4, 1, 1)	1.2	(4, 2, 0)	1.2
(5, 0, 1)	1.4	(5, 1, 0)	1.4	(6, 0, 0)	1.7	—	—	—	—

those parameters are the same as one set of parameters used in Medvedev and Scaillet (2010).<sup>6</sup> The parameter values used are listed in Table 1.

In Section 3.2, we mentioned that the over-relaxation parameter  $\omega$  is important to the convergence of the sparse grid approach and it usually depends on the shape of the grids as well. Table 2 illustrates this dependence in the case when  $l = 6$ . It can be seen from the table that the parameter  $\omega$  is usually higher for relatively less balanced grids, such as (0, 0, 6), (0, 6, 0), etc, in which the calculation will take more time as well.

In order to compare the option prices with Medvedev and Scaillet (2010) as well as demonstrate the efficiency and accuracy of our sparse grid approach, we first work out European put option prices by implementing the sparse grid (SG) approach and compare them with the solution from Grzelak and Oosterlee (2011) using the Fourier cosine expansion (COS) approach. Table 3 on the facing page reports the numerical

<sup>6</sup>The source code for all methods was implemented using NAG Fortran with the IMSL library running on the UTS, Faculty of Business F&E HPC Linux Cluster, which consists of eight nodes running Red Hat Enterprise Linux 4.0 (64bit) with  $2 \times 3.33$  GHz,  $2 \times 6$  MB cache Quad Core Xeon X5470 processors with 1333MHz FSB 8GB DDR2-667 RAM.

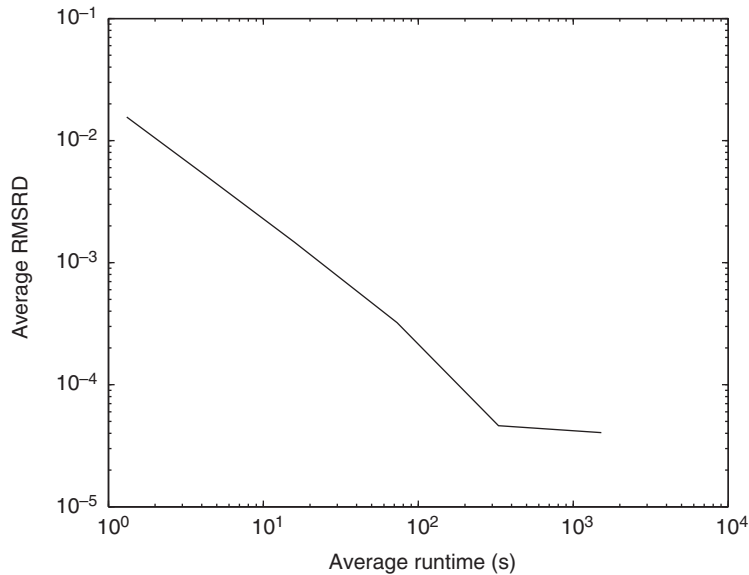


**TABLE 3** Daughter prices (European put) computed using the sparse grid with ( $c_1 = 16, c_2 = 8, c_3 = 4$ ), Fourier cosine expansion approach.

Level $l$	$S$					RMSRD	Runtime (s)
	80	90	100	110	120		
1	17.0446	9.7664	4.9921	2.3992	1.1530	$1.56 \times 10^{-2}$	1.31
2	16.9834	9.6920	4.9852	2.4424	1.1738	$5.11 \times 10^{-3}$	4.30
3	16.9640	9.6821	5.0100	2.4612	1.1796	$1.47 \times 10^{-3}$	15.84
4	16.9550	9.6828	5.0210	2.4625	1.1779	$3.22 \times 10^{-4}$	72.97
5	16.9520	9.6849	5.0226	2.4616	1.1774	$4.61 \times 10^{-5}$	329.12
6	16.9514	9.6855	5.0225	2.4616	1.1773	$4.05 \times 10^{-5}$	1517.34
COS	16.9512	9.6856	5.0227	2.4618	1.1774	—	—

Parameter values are given in Table 1 on the facing page, with  $v_0 = 0.04$  and  $r_0 = 0.04$ .

**FIGURE 5** Efficiency plot of the prices of the daughter European put option prices under the SV and SI models.



Daughter put option prices under Heston and stochastic interest rates when  $\rho = -0.5$

results and Figure 5 demonstrates the rate of the convergence of our SG approach with the increase of the level  $l$ . Table 3 has the following columns:  $l$  defines the level of the sparse grid; the put prices when the spot price goes from 80 to 120; the error

**TABLE 4** Mother prices (European put on European put) computed using the sparse grid with ( $c_1 = 16$ ,  $c_2 = 8$ ,  $c_3 = 4$ ), Monte Carlo with the Fourier cosine expansion approach.

Level $l$	$S$					RMSRD	Runtime (s)
	80	90	100	110	120		
1	-0.0326	0.2837	1.2164	2.3713	3.0846	$2.10 \times 10^{-1}$	1.83
2	0.2190	0.8083	1.7392	2.5594	3.1422	$1.55 \times 10^{-1}$	3.48
3	0.1012	0.5743	1.4898	2.4287	3.0852	$2.07 \times 10^{-2}$	12.17
4	0.0724	0.5619	1.4910	2.4144	3.0777	$2.96 \times 10^{-2}$	51.06
5	0.1064	0.6105	1.5229	2.4341	3.0867	$2.50 \times 10^{-2}$	220.37
6	0.1030	0.6052	1.5196	2.4322	3.0861	$1.88 \times 10^{-3}$	820.16
MC+COS	0.1032	0.6079	1.5228	2.4339	3.0889	—	2388.97
MC upper bound	0.1053	0.6129	1.5295	2.4405	3.0942	—	—
MC lower bound	0.1011	0.6030	1.5161	2.4273	3.0835	—	—

There are 1 000 000 sample paths and 200 time steps in the MC simulation. Parameter values are given in Table 1 on page 86, with  $v_0 = 0.04$  and  $r_0 = 0.04$ .

column gives the root mean square relative differences (RMSRD<sup>7</sup>) compared with the COS method; and the last column gives the runtime of our approach.

Using the level 6 prices in Table 3 on the preceding page as suitable terminal and boundary conditions for the mother options, we are able to work out the prices of the mother put option on the daughter option. Since we do not have a closed-form solution for the mother option, we compare our mother prices against the prices from Monte Carlo simulation using the daughter prices calculated from the COS method. The comparisons are displayed in Table 4.<sup>8</sup>

We see from the values given in Table 4 that the compound option prices from the sparse grid all lie between the lower and upper bounds of the MC simulation

<sup>7</sup> RMSRD is calculated as

$$\sqrt{\frac{1}{5} \sum_{i=1}^5 \left( \frac{\hat{P}(S_i) - P(S_i)}{P(S_i)} \right)^2},$$

where  $S_i = 80 + 10 \cdot (i - 1)$ ,  $\hat{P}(S)$  is the estimate of the price, and  $P(S)$  is the true price in the last row of Table 3 on the preceding page.

<sup>8</sup> Note that there is a negative price when  $S = 80$ ,  $l = 1$ . This is partly because, according to (3.2), the sequence of the weights are positive and negative interchangeably, for instance, if  $d = 3$ , we have  $a_0 = 1$ ,  $a_1 = -2$ ,  $a_2 = 1$ . Hence, it is possible to produce negative prices especially when the grid is very coarse. However, the “combination” technique ensures convergence when the grid becomes finer.

**TABLE 5** Daughter prices (American put) computed using the sparse grid with ( $c_1 = 16, c_2 = 8, c_3 = 4$ ), PSOR in a fine grid.

Level $l$	$S$					RMSRD	Runtime (s)
	80	90	100	110	120		
1	19.3227	10.0150	4.9658	2.4678	1.2260	$6.62 \times 10^{-2}$	1.34
2	20.0596	10.6755	5.2288	2.4966	1.1889	$3.80 \times 10^{-2}$	4.15
3	20.0275	11.0899	5.5819	2.7110	1.2712	$2.01 \times 10^{-2}$	18.33
4	20.0025	11.1125	5.5018	2.6028	1.2264	$8.36 \times 10^{-3}$	79.08
5	20.0232	11.0393	5.4922	2.6377	1.2432	$3.21 \times 10^{-3}$	383.10
6	20.0067	10.9663	5.4969	2.6299	1.2385	$8.85 \times 10^{-4}$	1708.86
PSOR	19.9987	10.9820	5.4899	2.6295	1.2388		124875.00

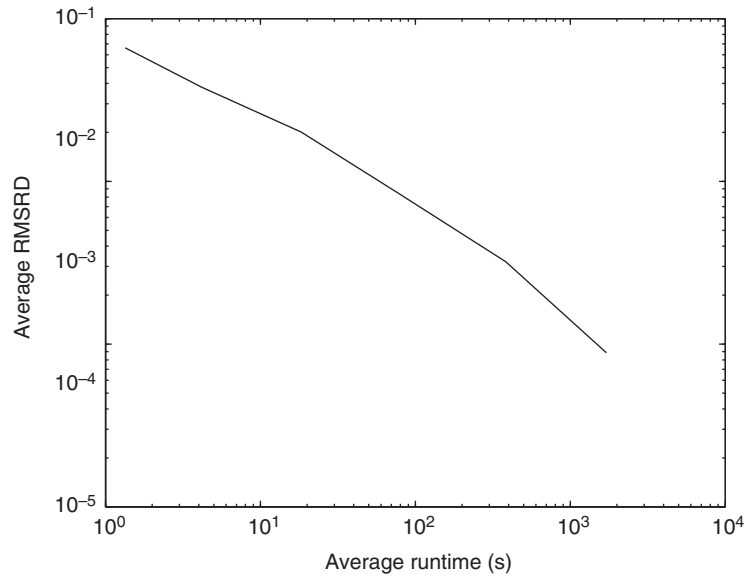
Parameter values are given in Table 1 on page 86, with  $v_0 = 0.04$  and  $r_0 = 0.04$ .

approach.<sup>9</sup> However, it should be emphasized that the runtime of the sparse grid approach involves the calculations of all compound option prices within the grid whereas the runtime of the simulation approach is just that required to work out the five prices at the five stock price values given in Table 4 on the facing page. It is clear that the sparse grid approach attains the same accuracy in far less time compared with the simulation approach.

Table 5 and Figure 6 on the next page demonstrate the convergence property of the sparse grid approach calculating the American put option prices under Heston with stochastic interest rate. A finite difference approach with PSOR with a fine grids discretization ( $N_1 = 1024, N_2 = 512, N_3 = 512, N_t = 128$ ) are treated as the benchmark prices. It is clear that the convergence of the sparse grid approach is quite fast.

Since the mother option can be exercised any time prior to the maturity of the compound option, at each time before maturity, in order to make the decisions on the early exercise, we need to compare the continuation value and the immediate exercise payoff, which is derived from the price of the daughter option at that particular time. However, with the increase in the number of the time step of the mother option, the computational burden to work out the American daughter option prices at different time to maturity will increase dramatically as well. Hence, a repeated Richardson extrapolation technique as described in Chang *et al* (2007) is used to find the prices of the compound option when the mother option is allowed to be exercised early. In this case, we are able to treat the mother option as a Bermudan option that can be exercised at a number of fixed time points before the maturity and here we use the

<sup>9</sup>The RMSRD in Table 4 on the facing page and Table 6 on the next page are only used to indicate that the results from the sparse grid “converge” in a correct direction.

**FIGURE 6** Efficiency plot of the prices of the daughter American put option prices under the SV and SI models.**TABLE 6** Mother prices (American put on American put) computed using the sparse grid with ( $c_1 = 16$ ,  $c_2 = 8$ ,  $c_3 = 4$ ).

Level $l$	$S$					RMSRD	Runtime (s)
	80	90	100	110	120		
1	0.0487	0.4439	1.3844	2.4962	3.2094	$2.78 \times 10^{-1}$	2.32
2	0.1241	0.6260	1.6071	2.5854	3.2090	$7.33 \times 10^{-2}$	4.38
3	0.0773	0.5538	1.5204	2.4944	3.1782	$1.33 \times 10^{-1}$	14.07
4	0.0924	0.5902	1.5482	2.5080	3.1835	$6.41 \times 10^{-2}$	61.42
5	0.1021	0.6168	1.5721	2.5179	3.1865	$2.16 \times 10^{-2}$	311.16
6	0.1080	0.6108	1.5675	2.5179	3.1862	$4.02 \times 10^{-3}$	1064.55
MC+PSOR	0.1072	0.6119	1.5618	2.5233	3.1928	—	8388.97
MC upper bound	0.1097	0.6175	1.5688	2.5298	3.1988	—	—
MC lower bound	0.1047	0.6063	1.55548	2.5168	3.1868	—	—

MC refers to the Longstaff and Schwartz (2001) with 1 000 000 sample paths, 500 time steps and 50 exercise dates. Parameter values are given in Table 1 on page 86, with  $v_0 = 0.04$  and  $r_0 = 0.04$ .

geometric-spaced exercise points employed in the modified Geske–Johnson formula described in Chang *et al* (2007). Table 6 on the facing page shows the results of the Richardson extrapolation technique compared with the least squares Monte Carlo approach by Longstaff and Schwartz (2001) with the daughter option prices calculated from the PSOR with a very fine grid ( $N_1 = 1024$ ,  $N_2 = 512$ ,  $N_3 = 512$ ,  $N_t = 128$ ). Boyarchenko and Levendorski (2007) and Medvedev and Scaillet (2010) have implemented the least squares Monte Carlo approach as a benchmark solution.

## 5 CONCLUSION

We have studied the pricing of American compound options by solving the corresponding PDE using both a sparse grid approach and a benchmark based on Monte Carlo simulation together with the PSOR approach.

It turns out that application of the standard SG approach is able to provide fairly accurate and efficient results for the compound option prices under Heston with stochastic interest rate. We apply this method twice to solve both PDEs followed by the price of the daughter option and the price of the mother option. We develop a benchmark solution by applying the finite difference with PSOR approach to solve the PDE followed by the daughter option in a fine grid. Then using these results to specify the appropriate boundary conditions, we employ least squares Monte Carlo simulation to find the prices of the mother option. The numerical results clearly show the computational advantage of the SG approach compared with the MC/PSOR approach.

In future research, the sparse grid approach can be speeded up by using better PDE solvers, such as the operator splitting in Ikonen and Toivanen (2004) and Ikonen and Toivanen (2007) with a modified adaptive sparse grid. The method may be applied to tackle specific examples in real options applications such as multistage investment projects, where it is important to take account of both stochastic volatility and stochastic interest rate due to the sensitivity of the compound options to volatility.

## REFERENCES

- Black, F., and Scholes, M. (1973). The pricing of corporate liabilities. *Journal of Political Economy* **81**, 637–659.
- Boyarchenko, S., and Levendorski, S. (2007). American options in the Heston model with stochastic interest rate and its generalization. URL: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1031282](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1031282).
- Brenner, M., Ou, E., and Zhang, J. (2006). Hedging volatility risk. *Journal of Banking and Finance* **30**(3), 811–821.
- Chang, C.-C., Chung, S.-L., and Stapleton, R. (2007). Richardson extrapolation techniques for the pricing of American-style options. *Journal of Futures Markets* **27**(8), 791–817.

- Chiarella, C., Ziogas, A., and Ziveyi, J. (2010). Representation of American option prices under Heston stochastic volatility dynamics using integral transforms. In *Contemporary Quantitative Finance: Essays in Honour of Eckhard Platen*, pp. 281–315. Springer.
- Ekstrom, E., Lotstedt, P., and Tysk, J. (2009). Boundary values and finite difference methods for the single factor term structure equation. *Applied Mathematical Finance* **16**(3), 253–259.
- Fouque, J. P., and Han, C. H. (2005). Evaluation of compound options using perturbation approximation. *The Journal of Computational Finance* **9**(1), 41–61.
- Geske, R. (1979). The valuation of compound options. *Journal of Financial Economics* **7**, 63–81.
- Geske, R., and Johnson, H. E. (1984). The American put option valued analytically. *Journal of Finance* **39**(5), 1511–1524.
- Griebel, M., Schneider, M., and Zenger, C. (1992). Iterative methods in linear algebra. In *A Combination Technique for the Solution of Sparse Grid Problems*, pp. 263–281. Elsevier.
- Grzelak, L., and Oosterlee, C. (2011). On the Heston model with stochastic interest rates. *SIAM Journal of Financial Mathematics* **2**, 255–286.
- Han, C. H. (2003). Singular perturbation on non-smooth boundary problems in finance. Thesis, North Carolina State University.
- Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* **6**(2), 327–343.
- Ikonen, S., and Toivanen, J. (2004). Operator splitting methods for American options with stochastic volatility. *Applied Mathematics Letters* **17**, 809–814.
- Ikonen, S., and Toivanen, J. (2007). Componentwise splitting methods for pricing American options under stochastic volatility. *International Journal of Theoretical and Applied Finance* **10**(2), 331–361.
- Kodukula, P., and Papudesu, C. (2006). *Project Valuation Using Real Options: A Practitioner's Guide*. J. Ross Publishing, Plantation, FL.
- Kwok, Y.-K. (2008). *Mathematical Models of Financial Derivatives*, 2nd edn. Springer.
- Leentvaar, C., and Oosterlee, C. (2008a). Multi-asset option pricing using a parallel fourier-based technique. *The Journal of Computational Finance* **12**(1), 1–26.
- Leentvaar, C., and Oosterlee, C. (2008b). On coordinate transformation and grid stretching for sparse grid pricing of basket options. *Journal of Computational and Applied Mathematics* **222**, 193–209.
- Longstaff, F., and Schwartz, E. (2001). Valuing American options by simulation: a simple least-squares approach. *Review of Financial Studies* **14**, 113–147.
- Medvedev, A., and Scaillet, O. (2010). Pricing American options under stochastic volatility and stochastic interest rates. *Journal of Financial Economics* **98**, 145–159.
- Meyer, G. (2013). Pricing options and bonds with the method of lines. Web notes <http://people.math.gatech.edu/~meyer/MOL-notes/>
- Reisinger, C. (2004). Numerische Methoden für hochdimensionale parabolische Gleichungen am Beispiel von Optionspreisaufgaben. Doctoral Thesis, Universität Heidelberg.
- Reisinger, C. (2008). Analysis of linear difference schemes in the sparse grid combination technique. Technical Report, Mathematical Institute, Oxford, UK.
- Reisinger, C., and Wittum, G. (2007). Efficient hierarchical approximation of high-dimensional option pricing problems. *SIAM Journal on Scientific Computing* **29**(1), 440–458.