Hedging with a Correlated Asset: Solution of a Nonlinear Pricing PDE *

H. Windcliff[†] J. Wang[‡] P.A. Forsyth[§] K.R. Vetzal[¶]

June 14, 2005

Abstract

Hedging a contingent claim with an asset which is not perfectly correlated with the underlying asset results in unhedgeable residual risk. Even if the residual risk is considered diversifiable, the option writer is faced with the problem of uncertainty in the estimation of the drift rates of the underlying and the hedging instrument. If the residual risk is not considered diversifible, then this risk can be priced using an actuarial standard deviation principle in infinitesmal time. In both cases, these models result in the same nonlinear partial differential equation (PDE). A fully implicit, monotone discretization method is developed for solution of this pricing PDE. This method is shown to converge to the viscosity solution. Certain grid conditions are required to guarantee monotonicity. An algorithm is derived which, given an initial grid, inserts a finite number of nodes in the grid to ensure that the monotonicity condition is satisfied. At each timestep, the nonlinear discretized algebraic equations are solved using an iterative algorithm, which is shown to be globally convergent. Monte Carlo hedging examples are given to illustrate the standard deviation of the profit and loss distribution at the expiry of the option.

Keywords: Hedging, contingent claim, basis risk, viscosity solution, monotone discretization
AMS Classification: 65M12, 65M60, 91B28

1 Introduction

In this paper we consider the problem of hedging a contingent claim in a case where the underlying asset cannot be traded. As a specific example, consider the following situation. Segregated funds are contracts offered by Canadian insurers which provide guarantees on mutual funds held in pension plan investment accounts [38]. In many cases, due to both legal and practical considerations, the insurance company hedges these guarantees using index futures in place of the actual mutual fund. The index, of course, will not be perfectly correlated with the underlying mutual fund, giving rise to basis risk.

In this situation, it is well known that it is possible to construct a best local hedge, in the sense that the residual risk is orthogonal to the risk which is hedged [20]. If an index is used to construct the hedge, and the residual risk is not correlated with the market index, it can be argued that this residual risk is firm specific and so can be diversified away. However, the pricing equation contains an effective drift rate which is a

*This work was supported by the Natural Sciences and Engineering Research Council of Canada and RBC Financial Group.

[†]Equity Trading Lab, Morgan Stanley, 1585 Broadway, New York, NY, U.S.A. 10036, email: Heath.Windcliff@morganstanley.com

[‡]School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1, email: j27wang@uwaterloo.ca

[§]School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1, email: paforsyt@uwaterloo.ca

[¶]Centre for Advanced Studies in Finance, University of Waterloo, Waterloo ON, Canada N2L 3G1, email: kvetzal@watarts.uwaterloo.ca

function of the actual drift rates of both the underlying mutual fund and the index. Drift rates are difficult to estimate. It is therefore natural to consider a worst case pricing approach, assuming only that the drift parameter lies between known bounds, but is otherwise uncertain. This approach gives rise to a nonlinear PDE [37].

However, the assumption of diversifible residual risk is questionable, especially since insurers are mandated to have sufficient reserves to guarantee solvency. As a result, the usual approach in the industry is to build up a reserve to provide for unhedgeable risk. In this paper, we will follow along the lines suggested in [25, 36], where the expected return of the hedging portfolio is adjusted to reflect a risk premium due to the unhedgeable risk. This approach is based on a common actuarial valuation principle [25, 36]. Essentially, insurers charge premia larger than the expected payoff of the hedging portfolio (in incomplete markets) so that sufficient reserves are built up to ensure solvency. This is known as *safety loading*. In [36], this valuation principle is translated into a measure of preferences. This measure can then be used in an indifference argument to generate a financial premium principle. A similiar pricing method was also suggested in [10].

More precisely, we use local risk minimization [34, 35, 9] to determine the best local hedge. We then use the *modified standard deviation principle* [26] in infinitesimal time to account for the residual risk. The standard deviation principle is used (as opposed to the variance principle) since it gives a value which is linear in terms of the number of units traded [26]. Applying this principle in infinitesmal time results in a method which is easily extended to American style contracts with complex path-dependent features, such as typically found in pension portfolio guarantees offered by insurers. This method gives rise to a nonlinear pricing PDE.

It is interesting to observe that the nonlinear PDE which results from worst case pricing with an uncertain drift term and the PDE which prices a contingent claim using the actuarial safety loading principle are identical in form. Hence, both the risk premium for bearing unhedgeable risk and the risk associated with uncertain parameter estimation may be taken into account using the same pricing PDE.

The nonlinear PDE gives a different value depending on whether the hedger is long or short the contingent claim. This is similar to the situation which arises in other nonlinear PDEs in finance, such as uncertain volatility and transaction cost models [4, 37, 30].

Since the pricing PDE is nonlinear, questions of convergence to the financially relevant solution arise. We develop a monotone, implicit scheme for discretization of the nonlinear pricing PDE. The results in [5, 15] can then be used to guarantee that the discrete solution converges to the viscosity solution. In order to ensure that the scheme is monotone, the grid must satisfy certain conditions. Given an initial grid, a node insertion algorithm is developed which ensures that the monotonicity conditions hold. We show that the insertion algorithm inserts a finite number of nodes, and that the grid aspect ratio of the grid after node insertion is only slightly increased compared to the grid aspect ratio of the original grid.

At each timestep, the implicit discretization leads to a nonlinear set of algebraic equations. An iterative algorithm is described for solution of the algebraic equations. The iterative method is designed so that existing PDE pricing software can be easily modified to solve the nonlinear algebraic equations. We prove that this algorithm is globally convergent. Moreover, convergence is quadratic in a sufficiently small neighborhood of the solution. We also prove that the discrete scheme satisfies certain arbitrage inequalities.

Finally, we include some numerical examples demonstrating that convergence of the nonlinear iteration at each timestep is rapid. We also include some Monte Carlo hedging simulations, where the optimal hedge parameters are given from the solution of the pricing PDE. The hedging simulation computations can then be used to determine the standard deviation, mean and value-at-risk (VaR) of the profit and loss distribution of the hedging portfolio at the expiry time of the contingent claim.

Although we focus specifically on the nonlinear PDE which arises in the context of uncertain drift rates and/or pricing of unhedgeable risk using an actuarial principle, this PDE has many of the charactersitics which arise in other nonlinear models in finance, including uncertain volatility [4], passport options [3],

utility-based pricing models [27], transaction cost models [23], and large investor effects [2]. As a result, we expect that many of the numerical methods developed here can be extended to these other nonlinear PDEs in finance.

2 Formulation

Let $V(S,t)$ be the value of a contingent claim written on asset S which follows the stochastic process

$$dS = \mu S\,dt + \sigma S\,dZ, \tag{2.1}$$

where μ is the drift rate, σ is volatility, and dZ is the increment of a Wiener process.

Suppose that we cannot trade in the underlying S, but only in a correlated asset H with price process

$$dH = \mu' H\,dt + \sigma' H\,dW, \tag{2.2}$$

where dW is the increment of a Wiener process. In the following we will use the usual Wiener process properties $dW^2 = dt, dZ^2 = dt, dZ\,dW = \rho\,dt$, where ρ is the correlation between dW and dZ. Consider a case where we wish to hedge a short position in the claim with value $V = V(S,t)$. Construct the portfolio

$$\Pi = -V + xH + B, \tag{2.3}$$

where x is the number of units of H held in the portfolio, and B is a risk free bond. We assume that $B = V - xH$ at time t, so that $\Pi(t) = 0$. The change in the portfolio value is given by (note that x is held constant in $[t, t+dt]$)

$$\begin{aligned} d\Pi &= -\left[V_t + \mu S V_S + \frac{\sigma^2 S^2}{2} V_{SS}\right] dt - \sigma S V_S\,dZ + r(V - xH)\,dt + x(\mu' H\,dt + \sigma' H\,dW) \\ &= -\left[V_t + \mu S V_S + \frac{\sigma^2 S^2}{2} V_{SS} + (xH - V)r - x\mu' H\right] dt - \sigma S V_S\,dZ + x\sigma' H\,dW. \end{aligned} \tag{2.4}$$

The variance of $d\Pi$ is given by

$$E^P\left[(x\sigma' H\,dW - \sigma S V_S\,dZ)^2\right] = \left[x^2(\sigma')^2 H^2 + \sigma^2 S^2 V_S^2 - 2\sigma S V_S x\sigma' H\rho\right] dt \tag{2.5}$$

where E^P is the expectation operator under the objective or P measure. Choosing x to minimize equation (2.5) gives

$$x = \left(\frac{\sigma S\rho}{\sigma' H}\right) V_S. \tag{2.6}$$

Substituting equation (2.6) into equation (2.4) gives

$$d\Pi = -\left[V_t + \mu S V_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV + \left(\frac{r\sigma S\rho}{\sigma'}\right) V_S - \left(\frac{\sigma S\rho\mu'}{\sigma'}\right) V_S\right] dt - \sigma S V_S\,dZ + \sigma S V_S \rho\,dW. \tag{2.7}$$

Defining

$$r' = \mu - (\mu' - r)\frac{\sigma\rho}{\sigma'} \tag{2.8}$$

in equation (2.7) gives

$$d\Pi = -\left[V_t + r' S V_S - rV + \frac{\sigma^2 S^2}{2} V_{SS}\right] dt + \sigma S V_S(\rho\,dW - dZ). \tag{2.9}$$

Note that to avoid arbitrage, we must have $r' \to r$ as $|\rho| \to 1$ [16]. Substituting equation (2.6) into equation (2.5) results in

$$\text{var}[d\Pi] = (1-\rho^2)\sigma^2 V_S^2 S^2\, dt. \tag{2.10}$$

Noting that $\text{cov}[\rho\, dW - dZ, dW] = 0$, we obtain $\text{cov}[d\Pi, dW] = 0$, so that the residual risk is orthogonal (in this sense) to the hedging instrument.

Define a new Brownian increment

$$dX = \frac{1}{\sqrt{1-\rho^2}}\left[\rho dW - dZ\right] \tag{2.11}$$

with the property $dX^2 = dt$. This allows us to write equation (2.9) as

$$d\Pi = -\left[V_t + r'SV_S - rV + \frac{\sigma^2 S^2}{2} V_{SS}\right] dt + SV_S\sqrt{1-\rho^2}\,\sigma\, dX. \tag{2.12}$$

Based on equation (2.12), one possible pricing approach is to require that the portfolio be mean self-financing

$$E^P[d\Pi] = 0. \tag{2.13}$$

This results in the linear PDE

$$V_t + r'SV_S - rV + \frac{\sigma^2 S^2}{2} V_{SS} = 0. \tag{2.14}$$

2.1 Uncertain Drift Rate

Unfortunately, equation (2.14) contains the term r' which is a function of the drift rates μ, μ'. In the usual complete market setting, the drift rate of the underlying asset disappears from the final PDE. However, in the cross hedging case, we are required to estimate μ, μ', which are notoriously difficult to determine. It might therefore be prudent to assume only that we can estimate a range of possible values for r',

$$r' \in [r'_{min}, r'_{max}]. \tag{2.15}$$

This is similar to the uncertain drift rate/dividend model described in [37]. The worst case price for an short position in the claim is given by [37]

$$V_t + \max_{r' \in [r'_{max}, r'_{min}]} \left(r'SV_S - rV + \frac{\sigma^2 S^2}{2} V_{SS}\right) = 0, \tag{2.16}$$

with the optimal choice for r' being

$$r' = \begin{cases} r'_{max} & \text{if } V_S > 0, \\ r'_{min} & \text{if } V_S \le 0. \end{cases} \tag{2.17}$$

Letting

$$\begin{aligned} r^* &= \frac{r'_{max} + r'_{min}}{2} \\ \lambda^* &= \frac{r'_{max} - r'_{min}}{2}, \end{aligned} \tag{2.18}$$

we can write equations (2.16) and (2.17) as

$$V_t + \left[r^* + \text{sgn}(V_S)\lambda^*\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0. \tag{2.19}$$

A similar argument for a worst case long position gives

$$V_t + \left[r^* - \text{sgn}(V_S)\lambda^*\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0. \tag{2.20}$$

For future reference, note that the two cases are

$$\begin{aligned} &\text{Short Position:} \qquad V_t + \left[r^* + \lambda^* \,\text{sgn}(V_S)\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0 \\ &\text{Long Position:} \qquad V_t + \left[r^* - \lambda^* \,\text{sgn}(V_S)\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0. \end{aligned} \tag{2.21}$$

2.2 Risk Loading

An insurance company which charged premia based only on equation (2.13) could soon have solvency problems [19]. As discussed in [25], insurance companies typically charge a premium for unhedgeable risk. If the residual risk is not diversifiable, then the option writer should be compensated for this risk. In this incomplete market situation, there are many possible approaches to the pricing problem. We will use the actuarial standard deviation principle in infinitesmal time. In our notation, this becomes

$$E^P[d\Pi] = \lambda \sqrt{\frac{\text{var}[d\Pi]}{dt}}\, dt \tag{2.22}$$

where λ is the *risk loading* parameter, which has units of $(time)^{-1/2}$ (the same units as a market price of risk). Note that we have specified that the expectation is under the P measure in equation (2.22). In order words, during each interval $[t, t+dt]$, the portfolio should earn a premium at a rate proportional to its instantaneous standard deviation. Note that the premium is based on the instantaneous properties of the portfolio, which means that this approach is trivially generalized to the path-dependent case. A similar idea was used in [1], in the context of a hedging strategy in the presence of transaction costs. In [1], the hedging strategy was constrained so that in each small time interval the expected gains from the hedging portfolio were proportional to the standard deviation of the gain.

From equation (2.10) we have that

$$\sqrt{\frac{\text{var}[d\Pi]}{dt}} = \sigma S |V_S| \sqrt{1-\rho^2}. \tag{2.23}$$

Combining equations (2.12,2.22,2.23) gives

$$V_t + r' SV_S + S|V_S|\lambda\sigma\sqrt{1-\rho^2} + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0, \tag{2.24}$$

or equivalently

$$V_t + \left[r' + \lambda\sigma\sqrt{1-\rho^2}\,\text{sgn}(V_S)\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0. \tag{2.25}$$

Note that the definition of Π in equation (2.3) assumes that the hedger is short the contingent claim V. Consequently, equation (2.25) is valid for a short position in V. Repeating the above arguments for a long position gives

$$V_t + \left[r' - \lambda\sigma\sqrt{1-\rho^2}\,\text{sgn}(V_S)\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0. \tag{2.26}$$

For future reference, note that the two cases are

$$\begin{aligned}
&\text{Short Position:} \qquad V_t + \left[r' + \lambda\sigma\sqrt{1-\rho^2}\,\operatorname{sgn}(V_S)\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0 \\
&\text{Long Position:} \qquad V_t + \left[r' - \lambda\sigma\sqrt{1-\rho^2}\,\operatorname{sgn}(V_S)\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0.
\end{aligned} \tag{2.27}$$

From equations (2.12) and (2.27) we have

$$\begin{aligned}
&\text{Short Position:} \qquad d\Pi = \lambda\sigma\sqrt{1-\rho^2}\, S|V_S|\, dt + SV_S\sqrt{1-\rho^2}\,\sigma\, dX \\
&\text{Long Position:} \qquad d\Pi = \lambda\sigma\sqrt{1-\rho^2}\, S|V_S|\, dt - SV_S\sqrt{1-\rho^2}\,\sigma\, dX.
\end{aligned} \tag{2.28}$$

Note that equations (2.27) have the same form as equations (2.21), if we make the identification

$$\begin{aligned}
\lambda\sigma\sqrt{1-\rho^2} &\to \lambda^* \\
r' &\to r^*.
\end{aligned} \tag{2.29}$$

In fact, we can combine both models (uncertain drift rate and actuarial risk-loading for unhedgeable risk) by defining

$$\begin{aligned}
r_c^* &= \frac{r'_{max} + r'_{min}}{2} \\
\lambda_c^* &= \lambda\sigma\sqrt{1-\rho^2} + \frac{r'_{max} - r'_{min}}{2}.
\end{aligned} \tag{2.30}$$

As a result, the combined model which takes both effects into account becomes (for worst case prices)

$$\begin{aligned}
&\text{Short Position:} \qquad V_t + \left[r_c^* + \lambda_c^* \operatorname{sgn}(V_S)\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0 \\
&\text{Long Position:} \qquad V_t + \left[r_c^* - \lambda_c^* \operatorname{sgn}(V_S)\right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0.
\end{aligned} \tag{2.31}$$

2.3 The Nonlinear Pricing PDE

For expositional simplicity in the following, we will consider the nonlinear PDE (2.28) which results only from the risk loading model. Of course, as outlined above, these nonlinear PDEs can as well be viewed as models of uncertain drift rates with suitable redefinition of the parameters.

Assuming $\lambda \geq 0$, then equations (2.27) are equivalent to

$$\begin{aligned}
&\text{Short Position:} \qquad V_\tau = \max_{q\in\{-1,+1\}} \left[\left(r' + q\lambda\sigma\sqrt{1-\rho^2}\right) SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV\right] \\
&\text{Long Position:} \qquad V_\tau = \min_{q\in\{-1,+1\}} \left[\left(r' + q\lambda\sigma\sqrt{1-\rho^2})\right) SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV\right],
\end{aligned} \tag{2.32}$$

where $\tau = T - t$, with T being the expiry time of the contingent claim. Note that the optimal choice for q in equation (2.32) is

$$q = \begin{cases} +\operatorname{sgn}(V_S) & \text{if short,} \\ -\operatorname{sgn}(V_S) & \text{if long.} \end{cases} \tag{2.33}$$

If we write (for a short position)

$$\mathscr{L}V \equiv V_\tau - \max_{q\in\{-1,+1\}} \left[\left(r' + q\lambda\sigma\sqrt{1-\rho^2}\right) SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV\right] \tag{2.34}$$

with the payoff denoted by $V = V^*$, then the price of a short contingent claim with an American early exercise feature would be given by $\min(\mathscr{L}V, V - V^*) = 0$. We will focus on European options in this paper, but much of the analysis can be extended to the American case as well.

2.4 Boundary Conditions

At $\tau = 0$, we set $V(S,0)$ to the payoff. As $S \to 0$, equation (2.27) reduces to

$$V_\tau = -rV. \tag{2.35}$$

In fact, in order to ensure certain properties of the discrete equations, we will impose condition (2.35) at some finite value $S_{min} > 0$, and let S_{min} tend to zero as the mesh is refined. We will demonstrate the effectiveness of this approximation through numerical tests.

As $S \to \infty$, we make the common assumption that $V_{SS} \simeq 0$, meaning that

$$V \simeq A(\tau)S + B(\tau); \qquad S \to \infty. \tag{2.36}$$

Assuming equation (2.36) holds, then substituting equation (2.36) into equation (2.27) gives ordinary differential equations for $A(\tau), B(\tau)$ with solution

$$V = A(0)S \exp\left[\left(r' - r + q\lambda\sigma\sqrt{1-\rho^2}\right)\tau\right] + B(0)\exp\left[-r\tau\right], \tag{2.37}$$

where q is given from equation (2.33) at $\tau = 0$. The initial conditions for $A(0), B(0)$ are given from the option payoffs.

2.5 Overview of Previous Work

We can relate equation (2.4) to the work in [34] by noting that for $\lambda = 0$, $d\Pi$ is the incremental profit of hedging. (In [34], the incremental cost is defined as $-d\Pi$.) In a complete market $d\Pi = 0$. In general, in incomplete markets, it is not possible to construct self-financing portfolios which perfectly replicate a contingent claim.

Consider the case where $\lambda = 0$. Let $\Pi(t + dt^-) = \Pi(t) + d\Pi(t)$. In general, $\Pi(t + dt^-)$ will not be zero, given that $\Pi(t) = 0$. In order to reset the portfolio value back to zero, cash is added to or subtracted from the portfolio so that

$$\Pi(t + dt^+) = \Pi(t + dt^-) - d\Pi(t) = 0, \tag{2.38}$$

hence this portfolio is not self-financing.

If $\lambda = 0$, then the approach used above is based on local risk minimization [34], i.e. we choose the trading strategy to minimize the variance of the incremental hedging profit/loss at each hedging time. Note that if $\lambda = 0$, then from equation (2.28) we have $E^P[d\Pi] = 0$, so this strategy is mean self-financing.

Given that the payoff of the option is used as an initial condition for equations (2.32) at $t = T$, cash must be infused into the portfolio during the hedging strategy in order to ensure that the payoff is met (the trading gains do not exactly balance the change in the option value during each infinitesimal step). As noted in [12], using the hedging parameters (2.6) given from the solution to equation (2.27), we can define a self-financing portfolio related to the locally risk minimizing portfolio, which in general will suffer from a shortfall at expiry. We will use this approach in our hedging simulations reported in Section 9.

The local risk minimization approach can be contrasted with the mean variance hedging or total risk minimization approach [35, 22]. In this strategy, a self-financing portfolio is constructed which minimizes the expected value of the square of the difference between the hedging portfolio and the payoff at the expiry

date. As discussed in [12], total risk minimization is a dynamic stochastic programming problem which is difficult, in general, to solve. In this paper, we will consider local risk minimization only, since this strategy attempts to control the riskiness of the hedging strategy at all times during the life of the contingent claim. This local risk minimization also appears natural in a context where the nature of the short contingent claim may change frequently, due to American style features [38]. Note that a similar combination of local risk minimization and a risk premium proportional to the standard deviation of the hedging portfolio was applied to real estate derivatives in [28].

In addition to the actuarial approaches mentioned above for optimal hedging with basis risk, another possible pricing method is based on maximizing exponential utility [16, 27]. It is interesting to note that if we had specified an actuarial variance principle

$$E^P[d\Pi] = \lambda^V \left[\frac{\text{var}[d\Pi]}{dt} \right] dt, \tag{2.39}$$

then we would obtain a nonlinear PDE identical to the PDE derived in [27]. (Note that the PDE in [27] is written for the case $r = 0$.)

3 Discretization

For discretization purposes, PDEs (2.32) can be written as

$$V_\tau = \left[r' + q\lambda\sigma\sqrt{1-\rho^2} \right] SV_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV, \tag{3.1}$$

where the nonlinear term q is given from equation (2.33). Define a grid $\{S_0, S_1, \ldots, S_p\}$, and let $V_i^n = V(S_i, \tau^n)$. Equation (3.1) can be discretized using forward, backward or central differencing in the S direction, coupled with a fully implicit timestepping to give

$$V_i^{n+1} - V_i^n = \alpha_i^{n+1} V_{i-1}^{n+1} + \beta_i^{n+1} V_{i+1}^{n+1} - (\alpha_i^{n+1} + \beta_i^{n+1} + r\Delta\tau) V_i^{n+1}, \tag{3.2}$$

where α_i, β_i are defined in Appendix A. We can also write the discrete equations in a manner consistent with the local max/min control problem (2.32). Let

$$\begin{aligned} \alpha_i^n &= \alpha_i' - q_{i,cent}^n \gamma_{i,cent}' - q_{i,back}^n \gamma_{i,back}' \\ \beta_i^n &= \beta_i' + q_{i,cent}^n \gamma_{i,cent}' + q_{i,for}^n \gamma_{i,for}', \end{aligned} \tag{3.3}$$

where $\alpha', \beta', \gamma', q_i^n$ are defined in Appendix A. Note that $q_i^n = \pm 1$ (see Appendix A).

In the following analysis, it will also be convenient to express discretization (3.2) in the form

$$\begin{aligned} V_i^{n+1} - V_i^n = {} & \alpha_i' V_{i-1}^{n+1} + \beta_i' V_{i+1}^{n+1} - (\alpha_i' + \beta_i' + r\Delta\tau) V_i^{n+1} \\ & + \kappa\gamma_{i,back}' |V_i^{n+1} - V_{i-1}^{n+1}| + \kappa\gamma_{i,for}' |V_{i+1}^{n+1} - V_i^{n+1}| + \kappa\gamma_{i,cent}' |V_{i+1}^{n+1} - V_{i-1}^{n+1}|, \end{aligned} \tag{3.4}$$

where

$$\kappa = \begin{cases} +1 & \text{if short,} \\ -1 & \text{if long.} \end{cases} \tag{3.5}$$

We approximate the infinite computational domain $S \in [0, \infty)$ by the finite domain $S \in [S_{min}, S_{max}]$. Denote the node corresponding to $S_i = S_{max}$ as $S_i = S_{imax}$. Let the discrete Dirichlet condition (2.37) at $S = S_{imax}$ be given by

$$D_{imax}^{n+1} = A(0) S_{imax} \exp\left[\left(r' - r + q\lambda\sigma\sqrt{1-\rho^2} \right) \tau^{n+1} \right] + B(0) \exp\left[-r\tau^{n+1} \right]. \tag{3.6}$$

For further notational convenience, we can write equation (3.2) in matrix form. Let

$$V^{n+1} = [V_0^{n+1}, V_1^{n+1}, \ldots, V_{imax}^{n+1}]'$$
$$V^n = [V_0^n, V_1^n, \ldots, V_{imax}^n]' \tag{3.7}$$

and

$$\left[\hat{M}^n V^n\right]_i = \left[(\alpha_i^n + \beta_i^n + r\Delta\tau)V_i^n - \alpha_i^n V_{i-1}^n - \beta_i^n V_{i+1}^n\right]; \qquad i < imax. \tag{3.8}$$

The first and last rows of $\hat{M}$ are modified as needed to handle the boundary conditions. The boundary condition at $S = S_{min}$ (equation (2.35)) is enforced by setting $\alpha_i = \beta_i = 0$ at $i = 0$. Let $D^{n+1} = [0, \ldots, D_{imax}^{n+1}]'$, and let I^* be the matrix which is identically zero, except for a one in the diagonal of the last row. The boundary condition at $i = imax$ is enforced by setting the last row of $\hat{M}$ to be identically zero. With a slight abuse of notation, we denote this last row as $(\hat{M})_{imax} \equiv 0$. In the following, it will be understood that equations of type (3.8) hold only for $i < imax$, with $(\hat{M})_{imax} \equiv 0$.

The discrete equations (3.2) can then be written as

$$\left[I + (1-\theta)\hat{M}^{n+1}\right] V^{n+1} = \left[I - \theta\hat{M}^n\right] V^n + I^*(D^{n+1} - V^n), \tag{3.9}$$

where the term $I^*(D^{n+1} - V^n)$ enforces the boundary condition at $S = S_{imax}$, and we have generalized the discretization (3.2) to the Crank Nicolson ($\theta = 1/2$) or fully implicit ($\theta = 0$) cases. Note that the discrete equations (3.9) are nonlinear since $\hat{M}^{n+1} = \hat{M}(V^{n+1})$.

4 Convergence to the Viscosity Solution

In [30], examples were given in which seemingly reasonable discretizations of nonlinear option pricing PDEs were unstable or converged to the incorrect solution. It is important to ensure that we can generate discretizations which are guaranteed to converge to the viscosity solution [5, 15]. Equation (2.32) satisfies the strong comparison property [6, 7, 11]. Hence from [8, 5], a numerical scheme converges to the viscosity solution if the method is consistent, stable (in the l_∞ norm) and monotone. For the convenience of the reader, we include a brief intuitive explanation of viscosity solutions in Appendix B.

4.1 Stability

We can ensure stability by requiring that discretization (3.2) be a positive coefficient method, $\alpha_i^n, \beta_i^n \geq 0$. This can be enforced by selecting a grid, and choosing forward, backward or central differencing so that the following condition is satisfied:

Condition 4.1 (Positive Coefficient Condition).

$$\beta_i' - \gamma_{i,cent}' - \gamma_{i,for}' \geq 0; \qquad i = 0, \ldots, imax-1$$
$$\alpha_i' - \gamma_{i,cent}' - \gamma_{i,back}' \geq 0; \qquad i = 0, \ldots, imax-1 \tag{4.1}$$

Note from the definitions of γ_i' in equations (A.17-A.19) that at each node only one of $\gamma_{i,cent}', \gamma_{i,for}', \gamma_{i,back}'$ is nonzero, and that $\gamma_i' \geq 0$. Condition (4.1) is based on the worst case choice of q_i^n in equation (3.3), hence this condition is independent of the solution. In other words, a grid is constructed, and central, forward or backward differencing is selected so that condition (4.1) is always satisfied. We emphasize that the choice of difference scheme is fixed, and does not depend on the solution. This is an important property [29] which will be used below. We will also give an algorithm in Section 7 which, given an arbitrary grid, can satisfy condition (4.1) by inserting a finite number of nodes.

Given condition (4.1), we have the following stability result

Lemma 4.1 (Stability of Discretization (3.2)). *Provided that*

- $r \geq 0$,
- *condition (4.1) is satisfied, and*
- *Dirichlet boundary conditions (2.35) and (2.36) are imposed,*

then the fully implicit discretization (3.2) is unconditionally stable in the sense that

$$\|V^{n+1}\|_\infty \leq \max\left(\|V^n\|_\infty, D^{n+1}_{imax}\right) \tag{4.2}$$

independent of the timestep size.

Proof. If conditions (4.1) are satisfied and $r \geq 0$, then it follows from equation (3.3) that α_i^n, β_i^n in discretization (3.2) are nonnegative, independent of the solution. The result then follows from a straightforward maximum analysis. ∎

4.2 Monotonicity

As discussed above, another important property of a discretization is monotonicity [5]. We write equations (3.2-3.4) as

$$\begin{aligned} g_i\left(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n\right) &= -\left(V_i^{n+1} - V_i^n\right) + \alpha_i^{n+1} V_{i-1}^{n+1} + \beta_i^{n+1} V_{i+1}^{n+1} - \left(\alpha_i^{n+1} + \beta_i^{n+1} + r\Delta\tau\right) V_i^{n+1} \\ &= -\left(V_i^{n+1} - V_i^n\right) + \alpha_i' V_{i-1}^{n+1} + \beta_i' V_{i+1}^{n+1} - \left(\alpha_i' + \beta_i' + r\Delta\tau\right) V_i^{n+1} \\ &\quad + \kappa\gamma'_{i,back}\left|V_i^{n+1} - V_{i-1}^{n+1}\right| + \kappa\gamma'_{i,for}\left|V_{i+1}^{n+1} - V_i^{n+1}\right| + \kappa\gamma'_{i,cent}\left|V_{i+1}^{n+1} - V_{i-1}^{n+1}\right| \\ &= 0, \qquad i = 0, \ldots, imax - 1, \end{aligned} \tag{4.3}$$

where κ is defined in equation (3.5).

Definition 4.1 (Monotonicity). *A discretization of the form (4.3) is monotone if the following conditions hold*

$$g_i\left(V_i^{n+1}, V_{i-1}^{n+1} + \varepsilon_1, V_{i+1}^{n+1} + \varepsilon_2, V_i^n + \varepsilon_3\right) \geq g_i\left(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n\right); \qquad \forall \varepsilon_i \geq 0, \tag{4.4}$$

$$g_i\left(V_i^{n+1} + \varepsilon_4, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n\right) < g_i\left(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n\right); \qquad \forall \varepsilon_4 \geq 0. \tag{4.5}$$

Observe that definition (4.1) includes condition (4.5), whereas only condition (4.4) is usually specified in the viscosity solution literature [5]. Condition (4.5) leads to a more intuitively appealing interpretation, and is a consequence of condition (4.4) and consistency [18].

Lemma 4.2 (Monotonicity). *If condition (4.1) is satisfied, then discretization (4.3) is monotone.*

Proof. The result holds trivially at $i = imax$ since $g_{imax} = -\left(V_{imax}^{n+1} - D_{imax}^{n+1}\right)$. For $i < imax$, from equation (4.3) we have (for $\varepsilon \geq 0$, and noting that $\gamma_i' \geq 0$; see Appendix A)

$$\begin{aligned} g_i\left(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1} + \varepsilon, V_i^n\right) - g_i\left(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n\right) &= \beta_i'\varepsilon \\ &\quad + \kappa\gamma'_{i,cent}\left[\left|V_{i+1}^{n+1} + \varepsilon - V_{i-1}^{n+1}\right| - \left|V_{i+1}^{n+1} - V_{i-1}^{n+1}\right|\right] \\ &\quad + \kappa\gamma'_{i,for}\left[\left|V_{i+1}^{n+1} + \varepsilon - V_i^{n+1}\right| - \left|V_{i+1}^{n+1} - V_i^{n+1}\right|\right] \\ &\geq \beta_i'\varepsilon - \gamma'_{i,cent}\varepsilon - \gamma'_{i,for}\varepsilon \\ &= \varepsilon\left(\beta_i' - \gamma'_{i,cent} - \gamma'_{i,for}\right) \geq 0 \end{aligned} \tag{4.6}$$

which follows from condition (4.1). Similarly,

$$g_i\left(V_i^{n+1}, V_{i-1}^{n+1}+\varepsilon, V_{i+1}^{n+1}, V_i^n\right) - g_i\left(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n\right) \geq \varepsilon(\alpha_i' - \gamma_{i,cent}' - \gamma_{i,back}') \geq 0 \tag{4.7}$$

and

$$\begin{aligned} g_i\left(V_i^{n+1}+\varepsilon, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n\right) - g_i\left(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n\right) &\leq -\varepsilon - \varepsilon(\alpha_i' + \beta_i' + r\Delta\tau) + \varepsilon\gamma_{i,back}' + \varepsilon\gamma_{i,for}' \\ &= -\varepsilon(1+r\Delta\tau) - \varepsilon(\alpha_i' - \gamma_{i,back}') - \varepsilon(\beta_i' - \gamma_{i,for}') \\ &\leq 0. \end{aligned} \tag{4.8}$$

Finally, it is obvious from equation (4.3) that

$$g_i\left(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n+\varepsilon\right) - g_i\left(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n\right) > 0. \tag{4.9}$$

∎

4.3 Consistency

The discrete scheme (3.9) is locally consistent with PDE (3.1) if the discrete operator applied to any C^∞ function converges to the equation (3.1) as the mesh size and timestep vanishes.

Lemma 4.3 (Consistency). *The discrete scheme (3.9) is locally consistent.*

Proof. From the definitions of the discrete coefficients α_i, β_i in equation (3.2) and Appendix A, a simple Taylor series verifies consistency. ∎

4.4 Convergence

Letting $\Delta\tau = \max_n\left(\tau^{n+1} - \tau^n\right)$, $\Delta S = \max_i\left(S_{i+1} - S_i\right)$, we can now state our convergence result.

Theorem 4.1 (Convergence of the Fully Implicit Discretization). *Provided that*

- $r \geq 0$,
- *the Dirichlet boundary conditions (2.35-2.36) are imposed, and condition (2.35) is imposed at S_{min}, $S_{min} \to 0$ as $\Delta S \to 0$, and*
- *the positive coefficient condition (4.1) holds,*

the fully implicit discretization (3.2) converges unconditionally to the viscosity solution of the nonlinear PDE (3.1) as $\Delta S, \Delta\tau \to 0$.

Proof. Since PDE (3.1) satisfies the strong comparison principle, a consistent, stable, and monotone discretization converges to the viscosity solution of PDE (3.1) [5]. Hence Theorem 4.1 follows directly from the results in [5] and Lemmas 4.1, 4.2 and 4.3. ∎

5 Solution of the Nonlinear Algebraic Equations

Although we have shown that the discretization converges to the viscosity solution, it is not clear that scheme (3.2) is practical since we must solve a set of nonlinear, nonsmooth algebraic equations at each timestep. The following iterative method is used to solve the nonlinear discretized algebraic equations (3.9):

Iterative Solution of the Discrete Equations

Let $(V^{n+1})^0 = V^n$

Let $\hat{V}^k = (V^{n+1})^k$

For $k = 0, 1, 2, \ldots$ until convergence

Solve

$$\left[I + (1-\theta)\hat{M}(\hat{V}^k)\right]\hat{V}^{k+1} = \left[I - \theta\hat{M}(V^n)\right]V^n + I^*(D^{n+1} - V^n) \tag{5.1}$$

If $(k > 0)$ and $\left(\max_i \dfrac{|\hat{V}_i^{k+1} - \hat{V}_i^k|}{\max\left(scale, |\hat{V}_i^{k+1}|\right)} < tolerance\right)$ then quit

EndFor

The scale factor in algorithm (5.1) is selected so that small option values are not determined with impractical precision. For example, if the option is valued in dollars, then $scale = 1$ would be a reasonable value for this parameter.

Some manipulation of algorithm (5.1) results in

$$\left[I + (1-\theta)\hat{M}^k\right]\left(\hat{V}^{k+1} - \hat{V}^k\right) = (1-\theta)\left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k, \tag{5.2}$$

where $\hat{M}^k = \hat{M}(\hat{V}^k)$. A key property which can be used to establish convergence of algorithm (5.1) concerns the sign of the right hand side of equation (5.2). We utilize a result obtained in [29]:

Lemma 5.1 (Single Signed Update). *If $\hat{M}^n V^n$ is given by equation (3.8), with nonlinear coefficients determined by a local control problem of form (3.3), and the choice of forward, backward, or central differencing is independent of the solution (i.e. preselected at each node independent of solution values), then*

$$\textit{Short Position:} \qquad \left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k \geq 0 \tag{5.3}$$

$$\textit{Long Position:} \qquad \left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k \leq 0. \tag{5.4}$$

Proof. For convenience, we summarize the proof in [29]. The result holds trivially at $i = imax$, since $(\hat{M}^{k-1} - \hat{M}^k)_{imax} \equiv 0$. Writing out $\left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k$ in component form gives ($i < imax$)

$$\begin{aligned}\left[\left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k\right]_i &= \left(\alpha_i^k\hat{V}_{i-1}^k + \beta_i^k\hat{V}_{i+1}^k - (\alpha_i^k + \beta_i^k + r\Delta\tau)\hat{V}_i^k\right) \\ &\quad - \left(\alpha_i^{k-1}\hat{V}_{i-1}^k + \beta_i^{k-1}\hat{V}_{i+1}^k - (\alpha_i^{k-1} + \beta_i^{k-1} + r\Delta\tau)\hat{V}_i^k\right).\end{aligned} \tag{5.5}$$

Consider a short position so that in terms of the local control problem (2.32), α_i^k, β_i^k are selected so that

$$\alpha_i^k\hat{V}_{i-1}^k + \beta_i^k\hat{V}_{i+1}^k - (\alpha_i^k + \beta_i^k + r\Delta\tau)\hat{V}_i^k \tag{5.6}$$

is maximized. Any other choice of coefficients, for example $\alpha_i^{k-1}, \beta_i^{k-1}$, cannot exceed the maximum produced by expression (5.6). Thus

$$\left(\alpha_i^k \hat{V}_{i-1}^k + \beta_i^k \hat{V}_{i+1}^k - (\alpha_i^k + \beta_i^k + r\Delta\tau)\hat{V}_i^k\right) - \left(\alpha_i^{k-1} \hat{V}_{i-1}^k + \beta_i^{k-1} \hat{V}_{i+1}^k - (\alpha_i^{k-1} + \beta_i^{k-1} + r\Delta\tau)\hat{V}_i^k\right) \geq 0, \quad (5.7)$$

so that for a short position $\left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k \geq 0$. A similar argument for a long position verifies (5.4). ∎

It is also useful to note the following property of the matrix $\left[I + (1-\theta)\hat{M}^{n+1}\right]$.

Lemma 5.2 (M-matrix). *If the positive coefficient condition (4.1) is satisfied, $r \geq 0$, and boundary conditions (2.35,2.36) are imposed at $S = S_{min}, S_{max}$, then $\left[I + (1-\theta)\hat{M}^{n+1}\right]$ is an M-matrix.*

Proof. As in the proof of Lemma 4.1, condition (4.1) implies that α_i^n, β_i^n in equation (3.8) are non-negative. Hence $\left[I + (1-\theta)\hat{M}^{n+1}\right]$ has positive diagonals, non-positive offdiagonals, and is diagonally dominant, so it is an M-matrix. ∎

Remark 5.1 (Properties of M-Matrices). *An M-matrix Q has the important properties that $Q^{-1} \geq 0$ and $diag(Q^{-1}) > 0$.*

We can now state our main result concerning the convergence of iteration (5.1).

Theorem 5.1 (Convergence of Iteration (5.1)). *Provided that the conditions required for Lemmas 5.1 and 5.2 are satisfied, then the nonlinear iteration (5.1) converges to the unique solution of equation (3.9) for any initial iterate $\hat{V}^0$. Moreover, the iterates converge monotonically, and for $\hat{V}^k$ sufficiently close to the solution, convergence is quadratic.*

Proof. Given Lemmas 5.1 and 5.2, the proof of this result is similar to the proof of convergence given in [30]. We give a brief outline of the steps in this proof, and refer readers to [30] for details. A straightforward maximum analysis of scheme (5.1) can be used to bound $\|\hat{V}^k\|_\infty$ independent of iteration k. From Lemma 5.1, we have that the right hand side of equation (5.2) is non-decreasing (non-increasing) for short (long) positions. Noting that $\left[I + (1-\theta)\hat{M}^k\right]$ is an M-matrix (from Lemma 5.2) and hence $\left[I + (1-\theta)\hat{M}^k\right]^{-1} \geq 0$, it is easily seen that the iterates form a bounded non-decreasing (short) or non-increasing (long) sequence. In addition, if $\hat{V}^{k+1} = \hat{V}^k$ the residual is zero. Hence the iteration converges to a solution. It follows from the M-matrix property of $\left[I + (1-\theta)\hat{M}^k\right]$ that the solution is unique. The iteration (5.1) can be regarded as a non-smooth Newton iteration. Since the non-smooth algebraic nonlinear equations (3.9) are strongly semi-smooth [32], convergence is quadratic in a sufficiently small neighborhood of the solution [31]. ∎

6 Arbitrage Inequalities

It is interesting to verify that the discrete equations satisfy discrete arbitrage inequalities [13, 14], independent of the choice of grid or timestep size. In other words, inequalities in option payoffs translate to inequalities in option values. More precisely, if V^n, W^n are two solutions of the fully implicit equations (4.3) and if $V^0 > W^0$ and $V_{imax}^k > W_{imax}^k, (k = 0, \ldots, n)$, then $V^n > W^n$.

Leting $D_V^{n+1} = [0, \ldots, V_{imax}^{n+1}]'$, $D_W^{n+1} = [0, \ldots, W_{imax}^{n+1}]'$, we have the following result:

Theorem 6.1 (Discrete Comparison Principle). *The fully implicit discretization (3.9) satisfies a discrete comparison principle, i.e. if $V^n > W^n, D_V^{n+1} > D_W^{n+1}$ and V^{n+1}, W^{n+1} satisfy equation (3.9) and the conditions for Lemmas 5.1 and 5.2 are satisfied, then $V^{n+1} > W^{n+1}$.*

Proof. V, W satisfy

$$\begin{aligned}\left[I+\hat{M}(V^{n+1})\right]V^{n+1} &= V^n + I^*(D_V^{n+1}-V^n)\\ \left[I+\hat{M}(W^{n+1})\right]W^{n+1} &= W^n + I^*(D_W^{n+1}-W^n).\end{aligned} \tag{6.1}$$

Some manipulation of equation (6.1) gives

$$\begin{aligned}\left[I+\hat{M}(W^{n+1})\right](V^{n+1}-W^{n+1}) &= (I-I^*)(V^n-W^n)+\left[\hat{M}(W^{n+1})-\hat{M}(V^{n+1})\right]V^{n+1}\\ &\quad + I^*(D_V^{n+1}-D_W^{n+1}) \qquad (6.2)\\ \left[I+\hat{M}(V^{n+1})\right](V^{n+1}-W^{n+1}) &= (I-I^*)(V^n-W^n)-\left[\hat{M}(V^{n+1})-\hat{M}(W^{n+1})\right]W^{n+1}\\ &\quad + I^*(D_V^{n+1}-D_W^{n+1}). \qquad (6.3)\end{aligned}$$

Consider a short position. From Lemma 5.1 (after relabeling $\hat{V}^{k-1}=W^{n+1}, \hat{V}^k = V^{n+1}$) we have that $\left[\hat{M}(W^{n+1})-\hat{M}(V^{n+1})\right]V^{n+1}\geq 0$. If $V^n > W^n$ and $D_V^{n+1} > D_W^{n+1}$, then from Lemma 5.2 and equation (6.2)

$$\left[I+\hat{M}(W^{n+1})\right]^{-1}\left[(I-I^*)(V^n-W^n)+I^*(D_V^{n+1}-D_W^{n+1})+\left(\hat{M}(W^{n+1})-\hat{M}(V^{n+1})\right)V^{n+1}\right]>0, \tag{6.4}$$

so $V^{n+1} > W^{n+1}$. For a long position, a similar argument using Lemmas 5.1 and 5.2 and equation (6.3) gives the same result. ∎

Remark 6.1 (Use of Lemma 5.1). *Note that a key property in the above proof is Lemma 5.1. This Lemma holds if we ensure that we solve a discrete version of the control problem (2.32), i.e. we maximize or minimize the discrete equations for a finite mesh and timesteps, not just in the limit of vanishing grid and timestep size. This illustrates the importance of maximizing or minimizing the discrete equations directly.*

7 Positive Coefficient Grid Condition

In this section, we develop an algorithm to ensure that grid condition (4.1) can be satisfied by insertion of a finite number of nodes in any initial grid. Some algebra shows that condition (4.1) is satisfied by at least one of forward or backward differencing at node i if

$$\sigma^2 S_i + (S_{i+1}-S_{i-1})\left(|r'|-\lambda\sigma\sqrt{1-\rho^2}\right)\geq 0. \tag{7.1}$$

Equation (7.1) is always satisfied if $\left(|r'|-\lambda\sigma\sqrt{1-\rho^2}\right)\geq 0$. Consequently, we will examine the case when $\left(|r'|-\lambda\sigma\sqrt{1-\rho^2}\right)<0$. Suppose $S_{i+1}-S_i=\Delta S, \forall i$, and so $S_i = i\Delta S$. Then condition (7.1) reduces to

$$\sigma^2 i + 2\left(|r'|-\lambda\sigma\sqrt{1-\rho^2}\right)\geq 0. \tag{7.2}$$

Clearly, for sufficiently large S_i condition (7.2) can be satisfied if $\sigma^2>0$. Equation (7.2) simplifies at $i=1$ to

$$\sigma^2 + 2\left(|r'|-\lambda\sigma\sqrt{1-\rho^2}\right)\geq 0. \tag{7.3}$$

Consequently, as $S_i\to 0$, condition (7.2) may not be satisfied, no matter how small ΔS is chosen. From equation (7.3), we can see that the problem arises since $S_0=0$. Instead, suppose we choose $S_i = S_0 + i\Delta S, S_0>0$. In this case condition (7.2) becomes

$$\sigma^2 S_0 + \Delta S\left(\sigma^2 i + 2|r'|\right) - 2\Delta S\lambda\sigma\sqrt{1-\rho^2}\geq 0, \tag{7.4}$$

which can always be satisfied if ΔS is sufficiently small and $S_0 > 0$. More generally, suppose

$$h = \max_i \left(S_{i+1} - S_i \right). \tag{7.5}$$

Condition (7.1) is always satisfied if

$$h \leq \frac{\sigma^2 S_0^2}{2\left| |r'| - \lambda\sigma\sqrt{1-\rho^2} \right|}. \tag{7.6}$$

Note that a grid constructed by enforcing condition (7.6) is not required in practice (as we shall see below). Condition (7.6) simply ensures that given $S_0 > 0$, a grid with a finite number of nodes can always be constructed which ensures that the positive coefficient condition (4.1) is satisfied.

In the following, we will develop an algorithm which, given an initial grid with $S_0 > 0$, will insert a finite number of nodes to ensure that condition (4.1) is satisfied. For a given grid with $S_0 > 0$, we will apply the boundary condition (2.35) at $S = S_0$. In order to carry out a convergence study, finer grids can be constructed by inserting nodes between each two coarse grid nodes, and reducing S_0 by half. In this way, the effect of applying boundary condition (2.35) at S_0 is reduced at each grid refinement. In fact, for practical values of σ, r' we expect that the effect of this approximation at $S = S_0$ is very small. This will be verified in some numerical examples.

The node insertion algorithm is given below:

Node Insertion Algorithm

If $\left((|r'| - \lambda\sigma\sqrt{1-\rho^2}) \geq 0 \right)$ Then
    Return // *Original grid satisfies condition*
Endif
If $\left([S_0 = 0] \text{ and } [\sigma^2 S_1 + \min(S_2, 2S_1)\left(|r'| - \lambda\sigma\sqrt{1-\rho^2} \right) < 0] \right)$ Then
    Exit // *Algorithm fails, need* $S_0 > 0$
Endif
$i = 1$
While (S_i is not the largest node)
    If $\left(\sigma^2 S_i + (S_{i+1} - S_{i-1})\left(|r'| - \lambda\sigma\sqrt{1-\rho^2} \right) < 0 \right)$ Then
        If $\left(\sigma^2 S_i + 2(S_i - S_{i-1})\left(|r'| - \lambda\sigma\sqrt{1-\rho^2} \right) < 0 \right)$ Then (7.7)
            Insert node at $(S_{i-1} + S_i)/2$
            //*New node labeled i*
        Else
            Insert node at $(S_i + S_{i+1})/2$
            //*New node labeled* $i+1$
        Endif
    Else
        Increment i
    Endif
Endwhile

If $S_0 \neq 0$, then algorithm (7.7) is guaranteed to produce a fine grid such that condition (7.1) holds for all nodes. From equation (7.6), the total number of nodes inserted must be finite.

If $S_0 = 0$ and $\sigma^2 S_1 + \min(S_2, 2S_1)\left(|r'| - \lambda\sigma\sqrt{1-\rho^2}\right) < 0$, then a new grid satisfying condition (7.1) does not exist. Consequently, in the case that $\sigma^2 + \left(|r'| - \lambda\sigma\sqrt{1-\rho^2}\right) < 0$, we must have $S_0 > 0$ in order for algorithm (7.7) to succeed. In this case, we can set S_0 to be a small number, and apply boundary condition (2.35) at S_0. We will verify that this does not cause any significant error at asset values of interest through some numerical experiments to be reported in subsequent sections. Algorithm (7.7) has the desirable property that the grid aspect ratio does not become too large after the node insertion is completed. More precisely, if the original grid has the property that

$$\begin{aligned} p_0 \leq \frac{S_{i+1} - S_i}{S_i - S_{i-1}} \leq q_0; \qquad & i = 1, \ldots, n-1 \\ q_0 \geq p_0 > 0, & \end{aligned} \tag{7.8}$$

we prove the following result in Appendix C.

Theorem 7.1 (Grid Aspect Ratio after Application of Algorithm (7.7)). *Given an initial grid with n nodes and p_0, q_0 given by equation (7.8), after application of algorithm (7.7) with $S_0 > 0$, the new grid (with m nodes, $m \geq n+1$) satisfies*

$$p \leq \frac{S_{i+1} - S_i}{S_i - S_{i-1}} \leq q \qquad 1 \leq i \leq m-1, \tag{7.9}$$

where $p = \min(1/3, p_0)$ and $q = \max(5, 2q_0)$.

Proof. See Appendix C. ∎

Note that algorithm (7.7) is based on testing only forward and backward differencing. However, in practice, we carry out the following steps

- Given an initial grid, construct a new grid from algorithm (7.7).
- Each node i of the new grid is processed, and the discretization coefficients α_i, β_i are constructed (equation (3.2)). First, central differencing is tested. If the positive coefficient condition (4.1) is satisfied, then we use central differencing at this node. If central differencing does not result in a positive coefficient discretization, then one of forward or backward differencing must satisfy this condition (from algorithm 7.7). Forward or backward differencing is then used at this node.

Different nodes may use different discretization methods. In this way, central differencing is used as much as possible. In practice, for normal market parameters, only a few nodes with forward or backward differencing are required. Usually these nodes are near $S = 0$, so that accuracy in regions of interest is unaffected by low order discretization methods.

8 Convergence Tests

This section presents a number of numerical examples which illustrate the performance and convergence of our iteration scheme. We also examine both fully implicit and Crank-Nicolson methods, and experiment with the minimum value in the asset grid (S_0), when algorithm (7.7) is applied. We show that the solution is insensitive to small positive S_0.

Constant timesteps are usually quite inefficient, so variable timesteps are desired. A simple and very effective timestep selector is discussed in [21]. Given an initial timestep $\Delta\tau^{n+1}$, a new timestep $\Delta\tau^{n+2}$ is selected so that

$$\Delta\tau^{n+2} = \left[\min_i \left(\frac{dnorm}{\frac{|V(S_i,\tau^n+\Delta\tau^{n+1})-V(S_i,\tau^n)|}{\max(D,|V(V(S_i,\tau^n+\Delta\tau^{n+1})|,|V(S_i,\tau^n)|)}} \right) \right] \Delta\tau^{n+1}, \tag{8.1}$$

where *dnorm* is a target relative change (during the timestep) specified by the user. The scale D is selected so that the timestep selector does not take an excessive number of timesteps in regions where the value is small (for options valued in dollars, $D = 1$ is often used).

Recall from equation (2.8) that the drift term in our PDE is

$$r' = \mu - (\mu' - r)\frac{\sigma\rho}{\sigma'}, \tag{8.2}$$

which implies

$$\frac{\mu - r'}{\sigma} = \frac{\rho(\mu' - r)}{\sigma'}. \tag{8.3}$$

When $|\rho| = 1$, the drift term r' must equal the risk free interest rate r [16]. Therefore, r, ρ, μ, σ, μ' and σ' cannot be determined independently. We arbitrarily choose μ' as the dependent variable. From equation (8.3) we see that if $\rho = 1$ and $r' = r$, we obtain

$$\mu' = r + (\mu - r)\frac{\sigma'}{\sigma}, \tag{8.4}$$

while if $\rho = -1$ ($r' = r$), we have

$$\mu' = r - (\mu - r)\frac{\sigma'}{\sigma}. \tag{8.5}$$

This suggests that we could interpolate μ' as

$$\mu' = r + f(\rho)(\mu - r)\frac{\sigma'}{\sigma}, \tag{8.6}$$

where, to avoid arbitrage, $f(-1) = -1, f(1) = 1$. In our numerical examples we simply choose

$$f(\rho) = \rho, \tag{8.7}$$

although any other interpolant could be used which satisfies $f(-1) = -1, f(1) = 1$. In the following numerical examples, we assume

$$\mu' = r + (\mu - r)\frac{\sigma'\rho}{\sigma}. \tag{8.8}$$

Substituting equation (8.6) into equation (8.2) gives

$$r' = (1 - \rho f(\rho))\mu + r\rho f(\rho). \tag{8.9}$$

Assuming equation (8.7) holds, we obtain

$$r' = (1 - \rho^2)\mu + \rho^2 r. \tag{8.10}$$

r	0.05
ρ	0.9
σ	0.2
μ	0.07
σ'	0.3
$\mu' = r + (\mu - r)\frac{\sigma'\rho}{\sigma}$	0.077
λ	0.2
$r' = \mu - (\mu' - r)\frac{\sigma\rho}{\sigma'}$	0.0538
Strike price K	100
Payoff	straddle
Time to expiry T	1 year

TABLE 1: *Parameters used in the straddle option examples. These parameters give* $|r'| - \lambda\sigma\sqrt{1-\rho^2} = 0.03636 > 0$*, so no new node is inserted into the asset grid when algorithm (7.7) is applied.*

8.1 Fully Implicit and Crank-Nicolson Comparison

In this section, we will examine the convergence as the grid and timesteps are refined for fully implicit and Crank-Nicolson timestepping. The parameters are given in Table 1. In this example we will assume a European straddle with payoff

$$V(S, \tau = 0) = \max(K - S, 0) + \max(S - K, 0). \tag{8.11}$$

where K is the strike price. The derivative (V_S) of the payoff changes sign, so the PDE is truly nonlinear. Note that $|r'| - \lambda\sigma\sqrt{1-\rho^2} = 0.03636 > 0$, so no new node is inserted into the asset grid when algorithm (7.7) is applied. The tolerance in algorithm (5.1) is set to 10^{-6}.

Table 2 shows the convergence results for fully implicit and Crank-Nicolson timestepping (using the modification suggested in [33]). We use variable timestepping as given in equation (8.1). As expected, fully implicit timestepping displays first order convergence and Crank-Nicolson method exhibits quadratic convergence. Recall from Theorem 4.1 that convergence to the viscosity solution is guaranteed only for fully implicit timestepping. In this case, Crank-Nicolson timestepping also converges to the viscosity solution. From algorithm (5.1), we can see the minimum number of iterations per timestep is two. In Table 2, we see that the average number of nonlinear iterations per timestep is only slightly larger than two, indicating that the nonlinear algebraic equations are very easily solved.

8.2 Positive S_0 Tests

In Section 7, it was shown that when $\sigma^2 + (|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0$, we need the minimum value for the asset grid $S_0 > 0$ in order for algorithm (7.7) to succeed. Table 3 shows parameters which require $S_0 > 0$ to ensure that algorithm (7.7) completes successfully.

Table 4 shows the option prices, deltas (V_S) and gammas (V_{SS}) under different S_0 values for various asset price values. We see that as asset price gets smaller, the effect of positive S_0 becomes more pronounced (recall that the strike of this option is \$100). However, for $S = 30$, the effect of changing S_0 from 2 to 0.1 is very small. The data in Table 3 was used for this test. Observe that this data requires $S_0 > 0$ for algorithm (7.7) to succeed.

Table 5 presents a convergence study using parameters from Table 3. As the asset grid size doubles and S_0 goes to zero, we obtain quadratic convergence as before. In order to have $\sigma^2 + (|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0$, we have assigned large values to σ, λ. This makes hedging with an imperfectly correlated asset very risky,

Nodes	*dnorm*	Timesteps	Nonlinear iterations	Option value	Change	Ratio
Fully Implicit						
51	0.1	37	81	17.02070		
101	0.05	72	151	17.05760	0.03689	
201	0.025	147	294	17.08743	0.02985	1.2365
401	0.0125	301	602	17.10857	0.02113	1.4120
801	0.00625	602	1204	17.11964	0.01108	1.9078
1601	0.003125	1169	2338	17.12508	0.00544	2.0349
Crank-Nicolson						
51	0.1	37	80	17.10144		
101	0.05	72	147	17.12367	0.02224	
201	0.025	147	294	17.12899	0.00532	4.1826
401	0.0125	301	602	17.13021	0.00122	4.3654
801	0.00625	602	1204	17.13050	0.00030	4.1054
1601	0.003125	1169	2338	17.13058	0.00007	3.9906

TABLE 2: *Convergence for fully implicit and Crank-Nicolson timestepping using variable timesteps (equation (8.1)). Crank-Nicolson incorporates the modification suggested in [33]. Input parameters are given in Table 1. No new nodes are inserted into the asset grid. Straddle payoff (8.11), short position, option values reported at* $S = 100$.

and the hedger is very risk averse (i.e. seeks high compensation for bearing the unhedgeable risk). These parameter values make the option prices extremely high as shown in Tables 4 and 5. From the data in Tables 4 and 5, we can conclude that small positive S_0 has little effect on the solution.

9 Hedging Simulations

In this section we use a Monte Carlo method to simulate the hedging process. We illustrate the results by showing histograms of the hedging portfolio at the expiry time, i.e. the profit and loss (P&L) distribution.

9.1 Algorithm Description

We make a slight change from the description of the hedging portfolio in equation (2.3). In the numerical examples, we will assume that the portfolio has initial value of zero, but no cash is injected into the portfolio as time progresses.

As in Section 2, consider the case where we wish to hedge a short position in a claim with value $V = V(S,t)$. Then, a portfolio $\mathscr{P}^i$ at time $t_i = i\Delta t$ has three components:

- a short claim position worth V^i;
- a long position of x^i shares of asset H (where x^i is the number of units of H held in the portfolio); and
- an amount $\mathscr{B}^i$ in a risk free account.

Hence,

$$\mathscr{P}^i = -V^i + x^i H^i + \mathscr{B}^i. \tag{9.1}$$

In contrast to the hedging portfolio Π in Section 2, we do not inject any cash into this portfolio $\mathscr{P}$ to ensure that $\mathscr{P} = 0$ after the initial time. In the case $\lambda = 0$, this portfolio is then self-financing (on $[0,T)$, where

r	0.03
ρ	0.5
σ	0.7
μ	0.04
σ'	0.25
$\mu' = r + (\mu - r)\frac{\sigma'\rho}{\sigma}$	0.0317857
λ	0.9
$r' = \mu - (\mu' - r)\frac{\sigma\rho}{\sigma'}$	0.0375
Strike price K	100
Payoff	straddle
Time to expiry T	1 year

TABLE 3: *Parameters used for positive S_0 tests. These parameters give $\sigma^2 + (|r'| - \lambda\sigma\sqrt{1-\rho^2}) = -0.0181 < 0$. In this case S_0 must be positive in order for algorithm (7.7) to succeed. When algorithm (7.7) is applied, new nodes may be inserted into the asset grid.*

Asset Price	S_0	Option Price	Delta (V_S)	Gamma (V_{SS})
10	0.1	91.2063	-0.583641	0.000119
	2	91.1604	-0.570818	-0.004019
	5	90.4519	-0.449766	-0.028363
20	0.1	85.3930	-0.576180	0.001791
	2	85.3879	-0.575238	0.001591
	5	85.2286	-0.553956	-0.001682
30	0.1	79.7849	-0.537174	0.006621
	2	79.7839	-0.537029	0.006598
	5	79.7360	-0.531765	0.005933

TABLE 4: *The effect of positive S_0 on the solution at low asset values. Crank-Nicolson method used with variable timesteps and the modification suggested in [33]. Input parameters are given in Table 3. Straddle payoff (8.11), short position. There are 401 nodes in the original grid. Seven new nodes are inserted for $S_0 = 0.1$, and no new node is inserted for $S_0 = 2$ or $S_0 = 5$.*

S_0	Nodes	*dnorm*	Option Value	Change	Ratio
5	58	0.1	102.69536		
2.5	106	0.05	102.83341	0.1381	
1.25	206	0.025	102.86821	0.0348	3.9678
0.625	409	0.0125	102.87715	0.0089	3.8902
0.3125	817	0.00625	102.87939	0.0022	3.9922
0.15625	1633	0.003125	102.87996	0.0006	3.9541
0.078125	3265	0.0015625	102.88010	0.0001	4.0071

TABLE 5: *Convergence in a case where S_0 has to be positive in order for the node insertion algorithm (7.7) to succeed. Crank-Nicolson timestepping is used with variable timesteps and the modification suggested in [33]. Input parameters are given in Table 3. These parameters imply that $\sigma^2 + (|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0$, so that S_0 has to be positive. Extreme values of λ, σ are required to force the necessity of making $S_0 > 0$. The option price is very large with these extreme parameters. When algorithm (7.7) is applied, new nodes may be inserted into the asset grid. The sizes of the original asset grids are 51, 102, 204, 408, 816, 1632, and 3264 respectively. Straddle payoff (8.11), short position, option values reported at $S = 100$.*

T is the expiry time), but in general it will not meet the exact obligations of the contingent claim at expiry. Note that PDE (2.27) does not contain B (the risk free account), so that use of x^i given by equation (2.6) minimizes the local risk, regardless of B. We have denoted the risk free account in the portfolio $\mathscr{P}$ by $\mathscr{B}$ to distinguish it from the account in equation (2.3). In this case, $\mathscr{P}$ will not necessarily be zero after the initial time, since we will not inject cash into this portfolio.

As discussed in [24] for the case $\lambda = 0$, this strategy is self-financing on $[0,T)$, with a single payment at time T. It is also observed in [24] that a disadvantage of this approach is that at any time $t < T$, the value of the portfolio will not equal the conditional expected value of the payoff. However, in our case with $\lambda > 0$, the value of the portfolio is increased by systematic gains to compensate for the risk of the hedge, and therefore this simple strategy may in fact be quite practical. In any case, we present the results of this strategy since it is easy to interpret the resulting P&L diagrams. These diagrams show the distribution of the future value of the incremental profit/loss of the hedge portfolio.

Given the option value at $t = 0$, which comes from the solution of the PDEs (2.27), the initial portfolio is given by

$$\mathscr{P}^0 = -V^0 + x^0 H^0 + \mathscr{B}^0. \tag{9.2}$$

We choose $\mathscr{B}^0 = V^0 - x^0 H^0$. Let

$$V_S^i = \frac{\partial V}{\partial S}(S^i, t_i). \tag{9.3}$$

According to equation (2.6), to minimize the local variance we choose x^i at time t_i to be

$$x^i = \left(\frac{\sigma S^i \rho}{\sigma' H^i}\right) V_S^i. \tag{9.4}$$

Let ϕ_H^i, ϕ_S^i be random draws from a normal distribution with zero mean and unit variance. The values of the underlying asset and the correlated hedging asset at time t_{i+1} are given by

$$\begin{aligned} S^{i+1} &= S^i \exp\left[(\mu - \sigma^2/2)\,\Delta t + \sigma \phi_S^i \sqrt{\Delta t})\right] \\ H^{i+1} &= H^i \exp\left[(\mu' - (\sigma')^2/2)\,\Delta t + \sigma' \phi_H^i \sqrt{\Delta t})\right], \end{aligned}$$

where

$$E^P(\phi_S\ \phi_H) = \rho. \tag{9.5}$$

Initially, we solve equation (2.32) numerically backward in time from $t = T$ to $t = 0$. At each timestep, the option values and deltas are stored in data tables. Then asset paths are generated by Monte Carlo simulation. The hedging information is recovered from the stored tables. The hedging algorithm for one Monte Carlo simulation is given in algorithm (9.6).

Hedging Algorithm

$\mathscr{P}^0 = 0$

Interpolate V^0 and V_S^0 from the stored tables

$x^0 = \left(\dfrac{\sigma S^0 \rho}{\sigma' H^0}\right) V_S^0$

$\mathscr{B}^0 = V^0 - x^0 H^0$

For each hedging time $0 < t_i \leq T, t_i = i\Delta t$

Calculate current asset price S^i and H^i from equation (9.5)

Interpolate V_S^i from the stored tables

$x^i = \left(\dfrac{\sigma S^i \rho}{\sigma' H^i}\right) V_S^i$

Update the portfolio by buying $x^i - x^{i-1}$ shares

$\mathscr{B}^i = e^{r\Delta t}\mathscr{B}^{i-1} - H^i(x^i - x^{i-1})$

Endfor

$\mathscr{P}(T) = -V(T) + x(T)H(T) + \mathscr{B}(T)$ (9.6)

Recall that

$$\begin{aligned} &\text{Short Position:} && d\Pi = \lambda\sigma\sqrt{1-\rho^2}S|V_S|\,dt + SV_S\sqrt{1-\rho^2}\sigma\,dX \\ &\text{Long Position:} && d\Pi = \lambda\sigma\sqrt{1-\rho^2}S|V_S|\,dt - SV_S\sqrt{1-\rho^2}\sigma\,dX. \end{aligned} \tag{9.7}$$

Considering only the short position, we have

$$d\Pi = \lambda\sigma\sqrt{1-\rho^2}S|V_S|\,dt + SV_S\sqrt{1-\rho^2}\sigma\,dX. \tag{9.8}$$

We will show histograms of $\mathscr{P}(T)$, i.e. the future P&L distribution. Since the cash shortfall is only realized at the expiry time in the portfolio $\mathscr{P}$, the final value of $\mathscr{P}$ can be determined in terms of the solution V by considering the future value of $d\Pi$ at each instant, i.e.

$$\mathscr{P}(T) = \lambda\int_0^T e^{r(T-t)}\sigma\sqrt{1-\rho^2}S|V_S|\,dt + \int_0^T e^{r(T-t)}SV_S\sqrt{1-\rho^2}\sigma\,dX. \tag{9.9}$$

This means that

$$E^P[\mathscr{P}(T)] = E^P\left[\lambda\int_0^T e^{r(T-t)}\sigma\sqrt{1-\rho^2}|V_S|\,dt\right] \tag{9.10}$$

in the limit as the rebalancing interval tends to zero.

9.2 Hedging Simulations

Hedging experiments are carried out using 1,000,000 Monte Carlo parameters with a fixed hedging interval of 2 days. There are many parameters which affect the hedging results, but we are primarily interested in the risk loading factor λ and the correlation ρ. We will show the results of the hedging simulations in terms

λ	ρ	Mean	VaR (95%)	CVaR (95%)	Std. Dev.	$V(S=100,\, t=0)$
0.0	0.5	-0.0034	-23.9135	-34.7796	12.524	16.4795
0.0	0.7	0.0085	-19.1239	-27.6244	10.284	16.3238
0.0	0.9	-0.001	-11.0917	-15.7675	6.293	16.1306

TABLE 6: *Hedging simulation results with $\lambda = 0.0$ and ρ varying. Other input parameters are given in Table 1. Straddle payoff (8.11), short position, hedging interval of 2 days, 1,000,000 simulation runs.*

λ	ρ	Mean	VaR (95%)	CVaR (95%)	Std. Dev.	$V(S=100,\, t=0)$
0.1	0.9	0.5081	-10.5158	-15.1710	6.3014	16.6233
0.3	0.9	1.6175	-9.1778	-13.6682	6.3155	17.6516
0.5	0.9	2.7685	-7.9180	-12.3103	6.3809	18.7388

TABLE 7: *Hedging simulation results with λ varying and $\rho = 0.9$. Other input parameters are given in Table 1. Straddle payoff (8.11), short position, hedging interval of 2 days, 1,000,000 simulation runs.*

of mean, standard deviation, VaR and CVaR of the P&L. If the probability density of the P&L x is denoted by $p(x)$, then the ξ% VaR and CVaR are defined as

$$\int_{\text{VaR}}^{\infty} p(x)\,dx = \frac{\xi}{100}$$
$$\text{CVaR} = \frac{\int_{-\infty}^{\text{VaR}} x p(x)\,dx}{\int_{-\infty}^{\text{VaR}} p(x)\,dx}. \tag{9.11}$$

When $\lambda = 0$, $E^P[\mathscr{P}(T)] = 0$ from equation (9.10). According to equation (2.22), increasing λ implies a greater reward for bearing the unhedgeable risk, hence the mean P&L (i.e. $E^P[\mathscr{P}(T)]$) should also increase (when $|\rho| \neq 1$). Table 6 shows the case in which λ is fixed at zero and ρ increases. Since $\lambda = 0$, the mean of the P&L stays at zero (it is not exactly zero because of finite rebalancing and Monte Carlo sampling error). As ρ increases, standard deviation decreases, which causes VaR and CVaR to increase. Table 7 shows the case where λ increases and the other parameters are held constant. As λ increases (i.e. we require greater reward for bearing unhedgeable risk), the mean, VaR, and CVaR of P&L increase, while standard deviation is nearly constant. These results are also depicted in panels (a), (c), and (d) of Figure 1.

When $|\rho| = 1$, asset H provides a perfect hedge and equation (2.27) reverts back to the usual Black-Scholes equation. In this case, the hedging simulation should be the same as standard discrete delta hedging, and thus the mean and standard deviation of the P&L should be zero. Some results for the case $|\rho| = 1$ are given in Table 8. Note that the standard deviation is not identically zero in this case due to the finite (two day) rebalancing interval.

Table 9 shows the results obtained when ρ increases from 0.7 to 0.9 and $\lambda \neq 0$. When $|\rho|$ increases, the hedging results become closer to that given by standard delta hedging. The mean shifts closer to zero (the

λ	ρ	Mean	VaR (95%)	CVaR (95%)	Std. Dev.	$V(S=100,\, t=0)$
0.5	1.0	-0.001	-1.9365	-2.7062	1.1752	16.0237
0.5	-1.0	-0.001	-2.2683	-3.0510	1.3583	16.0237

TABLE 8: *Hedging simulation results with $|\rho| = 1$ and $\lambda = 0.5$. Other input parameters are given in Table 1. Straddle payoff (8.11), short position, hedging interval of 2 days, 1,000,000 simulation runs. Note that the standard deviation of the P&L is nonzero due to the finite rebalancing interval.*

λ	ρ	Mean	VaR (95%)	CVaR (95%)	Std. Dev.	$V(S=100,\ t=0)$
0.2	0.7	1.8032	-17.0839	-25.4860	10.2861	18.0288
0.2	0.8	1.4828	-14.0163	-20.7803	8.6506	17.6383
0.2	0.9	1.0646	-9.8292	-14.3816	5.9792	17.1302

TABLE 9: *Hedging simulations with $\lambda = 0.2$ and ρ varying. Other input parameters are given in Table 1. Straddle payoff (8.11), short position, hedging interval of 2 days, 1,000,000 simulation runs.*

Hedging interval	Mean	VaR (95%)	CVaR (95%)	Std. Dev.
8	1.0626	-10.2150	-14.7470	6.5583
4	1.0523	-9.9895	-14.5614	6.3981
2	1.0646	-9.8292	-14.3816	6.3049
1	1.0662	-9.8029	-14.3994	6.2815

TABLE 10: *Convergence of the standard deviation as the hedging interval (measured in days) is decreased. Input parameters are given in Table 1. Straddle payoff (8.11), short position, 1,000,000 simulation runs.*

mean decreases, since we take less risk), and the standard deviation of the P&L decreases. These results are also illustrated in panels (b), (e), and (f) of Figure 1.

9.3 The Convergence of the Standard Deviation

If $|\rho| < 1$, there is unavoidable residual risk. As the hedging interval goes to zero and the number of simulations goes to infinity, the standard deviation of the portfolio at time T converges to a finite value. Table 10 provides a numerical example of this convergence.

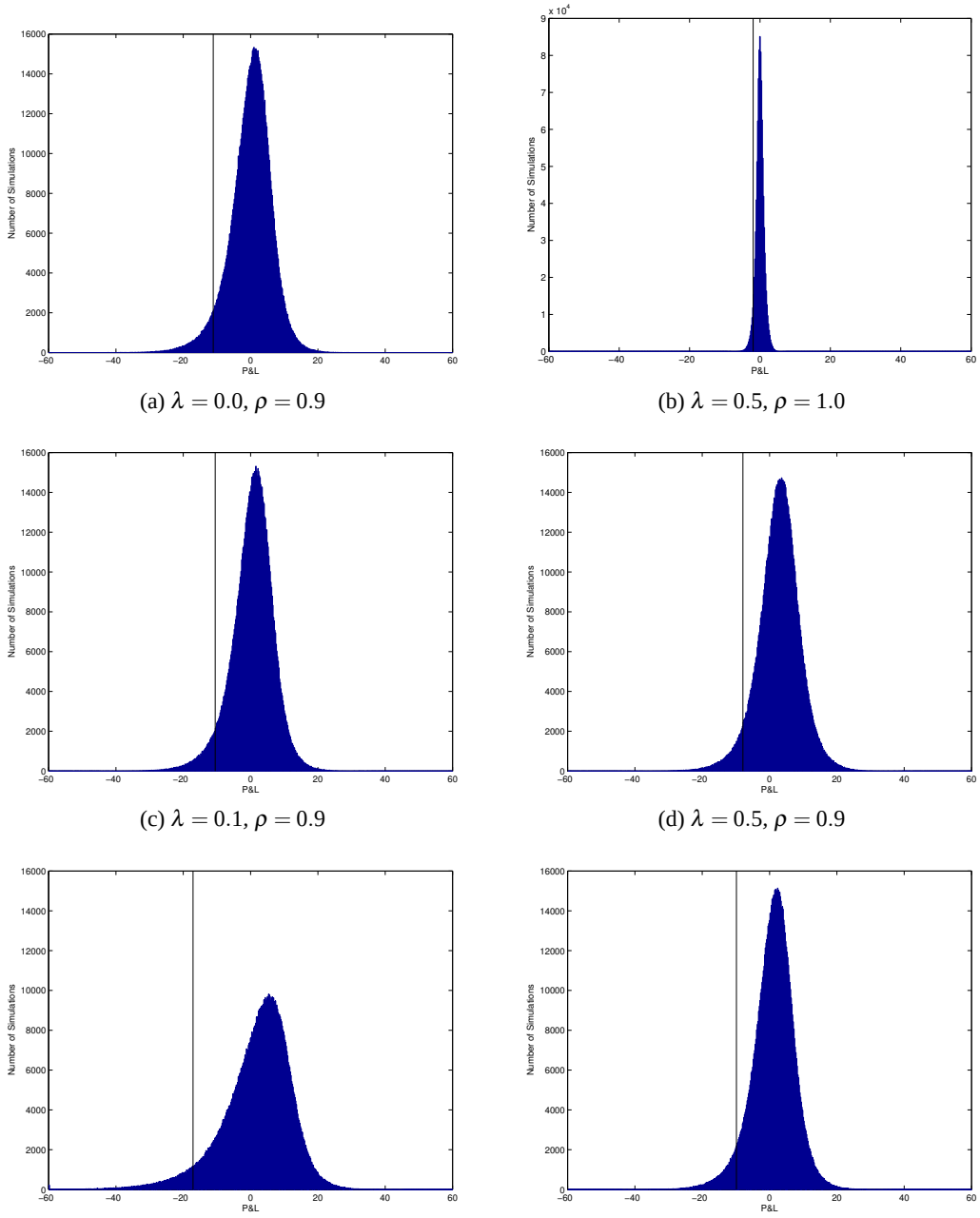
9.4 An American Example

The price of an American claim is given by equation (2.34). We can generalize the numerical methods described in this work to the American case using the penalty method described in [21, 17]. The proofs of convergence to the viscosity solution are easily extended to handle this case. As a numerical example, consider an American contingent claim, using the parameters in Table 1. Table 11 shows the values for long and short American/European straddle positions.

From equation (2.27), it is clear that the value of a short position should always be higher than that of a corresponding long position. Table 11 clearly shows this fact at a particular value of S.

Option Type	$V(S=100,\ t=0)$
European Short	17.13
European Long	15.19
American Short	17.39
American Long	15.70

TABLE 11: *Values for long and short positions of a straddle payoff (8.11). Input parameters are given in Table 1. Results are correct to the number of digits shown.*



(a) $\lambda = 0.0$, $\rho = 0.9$

(b) $\lambda = 0.5$, $\rho = 1.0$

(c) $\lambda = 0.1$, $\rho = 0.9$

(d) $\lambda = 0.5$, $\rho = 0.9$

(e) $\lambda = 0.2$, $\rho = 0.7$

(f) $\lambda = 0.2$, $\rho = 0.9$

FIGURE 1: *P& L distributions from Monte Carlo hedging simulations for various values of λ and ρ. Other input parameters are given in Table 1. Straddle payoff (8.11), short position, hedging interval of 2 days, 1,000,000 simulation runs. Note that the vertical scale for panel (b) differs from that in the other panels. The vertical line in each panel represents the 95% VaR of the P&L distribution.*

Option Type	$V(S=100,\ t=0)$
European Short Call	11.86
European Short Put	6.08
European Short Straddle	17.13

TABLE 12: *Call, put, and straddle values. Input parameters are given in Table 1. Results are correct to the number of digits shown. Note that the payoff of the straddle (8.11) is the sum of the call and put payoffs.*

9.5 Nonlinearity and Reinsurance

Suppose there are two firms, *A* and *B*, and a reinsurer *C*. Further assume that all of these firms value short positions using the parameters in Table 1. In particular, *A*, *B*, and *C* all have the same estimates for drift rates and the risk loading factor.

Suppose *A* needs to hedge a short call, and *B* needs to hedge a short put. *A* and *B* can hedge these positions, or purchase reinsurance from *C*. *C* would then have a short straddle position. The values from individually hedging a call, a put, and a straddle are given in Table 12 (calculated using PDE (2.27)). If *A* and *B* individually hedge their positions, their total charge to an end customer would be $11.86 + 6.08 = 17.94$. On the other hand, the total charge to *A* and *B* if *C* hedges a straddle is 17.13. In this case, *C* can charge a lower fee for this insurance than *A* and *B* can do by themselves. This result is due to the fact that the pricing PDE is nonlinear.

10 Conclusions

In this paper, we have considered the situation where a financial institution selling a contingent claim cannot hedge directly with the asset underlying the claim. At each infinitesmal time interval, the best local hedge is constructed. Even if the residual risk is diversifiable, the option writer may be exposed to uncertainty in parameter estimation. Assuming that the parameters are uncertain but lie within upper and lower bounds, a worst case pricing approach can be used. This results in a nonlinear PDE.

However, since the hedge is not perfect, the writer may not be able to diversify the unhedgeable risk. In this case, this risk can be priced using an actuarial standard deviation principle in infinitesmal time. The risk preferences of the issuing firm enter into the valuation through a risk loading parameter. For non-zero risk-loading, the PDE is nonlinear, producing different values for long or short positions. Note that in contrast to many other approaches, the values are linear in terms of the number of units bought/sold.

In both cases (uncertain parameters and actuarial standard deviation principle), the nonlinear PDE has the same form. We have developed a discretization scheme for this nonlinear PDE which is monotone, consistent and stable; hence convergence to the viscosity solution is guaranteed. In order to ensure the discretization is monotone, a node insertion algorithm is derived which guarantees monotonicity by insertion of a finite number of nodes in a given initial grid. An iterative method for solution of the nonlinear discrete algebraic equations at each timestep is developed. We have proven that this iteration is globally convergent. Existing PDE option pricing software can be modified in a straightforward fashion to value options using this model, simply by adding an updating step to the American pricing iteration.

If we interpret the PDE as accounting for the unhedgeable risk, then the solution of the PDE gives a trading strategy for the best possible local hedge, as well as providing systematic gains to compensate for the residual risk. Monte Carlo hedging experiments are given which demonstrate the use of this hedging strategy. These examples clearly show that the unhedgeable risk is compensated by a reserve which is built up over time.

Finally, we note that many of the numerical methods discussed here can be extended to other nonlinear pricing PDEs in financial applications.

A Discrete Equation Coefficients

The detailed form of the discrete equation coefficients used in equation (3.3) are given here. In the case of a central discretization

$$\begin{aligned} \alpha^n_{i,cent} &= \alpha'_{i,cent} - \gamma_{i,cent}\, q^n_{i,cent} \\ \beta^n_{i,cent} &= \beta'_{i,cent} + \gamma_{i,cent}\, q^n_{i,cent}, \end{aligned} \tag{A.1}$$

where

$$q^n_{i,cent} = \begin{cases} \operatorname{sgn}(V_S)^n_{i,cent} & \text{if short} \\ -\operatorname{sgn}(V_S)^n_{i,cent} & \text{if long} \end{cases}, \tag{A.2}$$

$$\begin{aligned} \gamma_{i,cent} &= \left(\frac{S_i \lambda \sigma \sqrt{1-\rho^2}}{S_{i+1} - S_{i-1}} \right) \Delta\tau \\ (V_S)^n_{i,cent} &= \frac{V^n_{i+1} - V^n_{i-1}}{S_{i+1} - S_{i-1}}, \end{aligned} \tag{A.3}$$

and

$$\begin{aligned} \alpha'_{i,cent} &= \left[\frac{\sigma^2 S_i^2}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} - \frac{r' S_i}{S_{i+1} - S_{i-1}} \right] \Delta\tau \\ \beta'_{i,cent} &= \left[\frac{\sigma^2 S_i^2}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \frac{r' S_i}{S_{i+1} - S_{i-1}} \right] \Delta\tau. \end{aligned} \tag{A.4}$$

Note that the above definitions ensure that we are solving a discrete version of the local control problem (2.33).

In the case of forward differencing, we obtain

$$\begin{aligned} \alpha^n_{i,for} &= \alpha'_{i,for} \\ \beta^n_{i,for} &= \beta'_{i,for} + \gamma_{i,for}\, q^n_{i,for}, \end{aligned} \tag{A.5}$$

where

$$q^n_{i,for} = \begin{cases} \operatorname{sgn}(V_S)^n_{i,for} & \text{if short} \\ -\operatorname{sgn}(V_S)^n_{i,for} & \text{if long} \end{cases}, \tag{A.6}$$

$$\begin{aligned} \gamma_{i,for} &= \left(\frac{S_i \lambda \sigma \sqrt{1-\rho^2}}{S_{i+1} - S_i} \right) \Delta\tau \\ (V_S)^n_{i,for} &= \frac{V^n_{i+1} - V^n_i}{S_{i+1} - S_i}, \end{aligned} \tag{A.7}$$

and

$$\begin{aligned} \alpha'_{i,for} &= \left(\frac{\sigma^2 S_i^2}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} \right) \Delta\tau \\ \beta'_{i,for} &= \left[\frac{\sigma^2 S_i^2}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \frac{r' S_i}{S_{i+1} - S_i} \right] \Delta\tau. \end{aligned} \tag{A.8}$$

Again, note that we have used definition (A.6), so that we solve a discrete version of the local control problem (2.33).

In the case of backward differencing we have

$$\begin{aligned} \alpha^n_{i,back} &= \alpha'_{i,back} - \gamma_{i,back}\, q^n_{i,back} \\ \beta^n_{i,back} &= \beta'_{i,back}, \end{aligned} \tag{A.9}$$

where

$$q^n_{i,back} = \begin{cases} \operatorname{sgn}(V_S)^n_{i,back} & \text{if short} \\ -\operatorname{sgn}(V_S)^n_{i,back} & \text{if long} \end{cases}, \tag{A.10}$$

$$\begin{aligned} \gamma_{i,back} &= \left(\frac{S_i \lambda \sigma \sqrt{1-\rho^2}}{S_i - S_{i-1}} \right) \Delta\tau \\ (V_S)^n_{i,back} &= \frac{V^n_i - V^n_{i-1}}{S_i - S_{i-1}}, \end{aligned} \tag{A.11}$$

and

$$\begin{aligned} \alpha'_{i,back} &= \left[\frac{\sigma^2 S_i^2}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} - \frac{r' S_i}{S_i - S_{i-1}} \right] \Delta\tau \\ \beta'_{i,back} &= \left[\frac{\sigma^2 S_i^2}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} \right] \Delta\tau. \end{aligned} \tag{A.12}$$

For future reference, it is convenient to define *generic* coefficients

$$\alpha^n_i = \begin{cases} \alpha^n_{i,cent} & \text{if central differencing} \\ \alpha^n_{i,for} & \text{if forward differencing} \\ \alpha^n_{i,back} & \text{if backward differencing} \end{cases}, \tag{A.13}$$

$$\beta^n_i = \begin{cases} \beta^n_{i,cent} & \text{if central differencing} \\ \beta^n_{i,for} & \text{if forward differencing} \\ \beta^n_{i,back} & \text{if backward differencing} \end{cases}, \tag{A.14}$$

$$\alpha'_i = \begin{cases} \alpha'_{i,cent} & \text{if central differencing} \\ \alpha'_{i,for} & \text{if forward differencing} \\ \alpha'_{i,back} & \text{if backward differencing} \end{cases}, \tag{A.15}$$

$$\beta'_i = \begin{cases} \beta'_{i,cent} & \text{if central differencing} \\ \beta'_{i,for} & \text{if forward differencing} \\ \beta'_{i,back} & \text{if backward differencing} \end{cases}. \tag{A.16}$$

We also define

$$\gamma'_{i,cent} = \begin{cases} \gamma_{i,cent} & \text{if central differencing} \\ 0 & \text{otherwise} \end{cases}, \tag{A.17}$$

$$\gamma'_{i,for} = \begin{cases} \gamma_{i,for} & \text{if forward differencing} \\ 0 & \text{otherwise} \end{cases}, \tag{A.18}$$

$$\gamma'_{i,back} = \begin{cases} \gamma_{i,back} & \text{if backward differencing} \\ 0 & \text{otherwise} \end{cases}. \tag{A.19}$$

Recalling equation (3.2)

$$V_i^{n+1} - V_i^n = \alpha_i^{n+1} V_{i-1}^{n+1} + \beta_i^{n+1} V_{i+1}^{n+1} - (\alpha_i^{n+1} + \beta_i^{n+1} + r\Delta\tau) V_i^{n+1}, \tag{A.20}$$

we can write the generic coefficients α_i^n, β_i^n as

$$\begin{aligned} \alpha_i^n &= \alpha'_i - \gamma'_{i,cent}\, q_{i,cent}^n - \gamma'_{i,back}\, q_{i,back}^n \\ \beta_i^n &= \beta'_i + \gamma'_{i,cent}\, q_{i,cent}^n + \gamma'_{i,for}\, q_{i,for}^n. \end{aligned} \tag{A.21}$$

B Viscosity Solution

In this appendix, we give a brief intuitive explanation of the ideas behind the definition of a viscosity solution. For more details, we refer the reader to [15].

Consider a short position, so that we can write equation (2.32) as

$$g(V, V_S, V_{SS}, V_\tau) = -V_\tau + \max_{q \in \{-1,+1\}} \left[\left(r' + q\,\lambda\,\sigma\sqrt{1-\rho^2} \right) S V_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV \right] = 0. \tag{B.1}$$

We assume that $g(x,y,z,w)$ ($x = V, y = V_S, z = V_{SS}, w = V_\tau$) satisfies the ellipticity condition

$$g(x,y,z+\varepsilon,w) \geq g(x,y,z,w) \qquad \forall \varepsilon \geq 0, \tag{B.2}$$

which is our case simply means that $\sigma^2 \geq 0$. Suppose for the moment that smooth solutions to equation (B.1) exist, i.e. $V \in C^{2,1}$, where $C^{2,1}$ refers to a continuous function $V = V(S,\tau)$ having continuous first and second derivatives in S, and continuous first derivatives in τ. Let ϕ be a set of $C^{2,1}$ test functions. Suppose $\phi - V \geq 0$ and $\phi(S_0,\tau_0) = V(S_0,\tau_0)$ at the single point (S_0,τ_0). Then the single point (S_0,τ_0) is a global minimum of $(\phi - V)$

$$\begin{aligned} \phi - V &\geq 0 \\ \min(\phi - V) &= \phi(S_0,\tau_0) - V(S_0,\tau_0) = 0. \end{aligned} \tag{B.3}$$

Consequently, at (S_0,τ_0)

$$\begin{aligned} \phi_\tau &= V_\tau \\ \phi_S &= V_S \\ \phi_{SS} &\geq V_{SS}. \end{aligned} \tag{B.4}$$

Hence, from equations (B.2,B.4) we have

$$\begin{aligned} g(V(S_0,\tau_0), \phi_S(S_0,\tau_0), \phi_{SS}(S_0,\tau_0), \phi_\tau(S_0,\tau_0)) &= g(V(S_0,\tau_0), V_S(S_0,\tau_0), \phi_{SS}(S_0,\tau_0), V_\tau(S_0,\tau_0)) \\ &\geq g(V(S_0,\tau_0), V_S(S_0,\tau_0), V_{SS}(S_0,\tau_0), V_\tau(S_0,\tau_0)) \\ &= 0, \end{aligned} \tag{B.5}$$

or, to summarize,

$$\begin{aligned} g(V(S_0,\tau_0), \phi_S(S_0,\tau_0), \phi_{SS}(S_0,\tau_0), \phi_\tau(S_0,\tau_0)) &\geq 0 \\ \phi - V &\geq 0 \\ \min(\phi - V) = \phi(S_0,\tau_0) - V(S_0,\tau_0) &= 0. \end{aligned} \tag{B.6}$$

Now, suppose that χ is a $C^{2,1}$ test function with $V-\chi \geq 0$, and $V(S_0,\tau_0) = \chi(S_0,\tau_0)$ at the single point (S_0,τ_0). Then (S_0,τ_0) is the global minimum of $V-\chi$,

$$\begin{aligned} V-\chi &\geq 0 \\ \min(V-\chi) &= V(S_0,\tau_0) - \chi(S_0,\tau_0) \\ &= 0. \end{aligned} \tag{B.7}$$

Repeating the above arguments we have

$$\begin{aligned} g(V(S_0,\tau_0),\chi_S(S_0,\tau_0),\chi_{SS}(S_0,\tau_0),\chi_\tau(S_0,\tau_0)) &\leq 0 \\ V-\chi &\geq 0 \\ \min(V-\chi) = V(S_0,\tau_0) - \chi(S_0,\tau_0) &= 0. \end{aligned} \tag{B.8}$$

Now suppose that V is continuous but not smooth. This means that we cannot define V as the solution to $g(V,V_S,V_{SS},V_\tau) = 0$. However, we can still use conditions (B.6) and (B.8) to define a viscosity solution to equation (B.1) since all derivatives are applied to smooth test functions. Informally, a viscosity solution V to equation (B.1) is defined such that

- For any $C^{2,1}$ test function ϕ, such that

$$\phi - V \geq 0 \qquad \phi(S_0,\tau_0) = V(S_0,\tau_0), \tag{B.9}$$

  (ϕ touches V at the single point (S_0,τ_0)), then

$$g(V(S_0,\tau_0),\phi_S(S_0,\tau_0),\phi_{SS}(S_0,\tau_0),\phi_\tau(S_0,\tau_0)) \geq 0. \tag{B.10}$$

- As well, for any $C^{2,1}$ test function χ such that

$$V-\chi \geq 0 \qquad V(S_0,\tau_0) = \chi(S_0,\tau_0), \tag{B.11}$$

  (χ touches V at the single point (S_0,τ_0)) then

$$g(V(S_0,\tau_0),\chi_S(S_0,\tau_0),\chi_{SS}(S_0,\tau_0),\chi_\tau(S_0,\tau_0)) \leq 0. \tag{B.12}$$

This definition is illustrated in Figure 2.

C Grid Aspect Ratio Proof

In this appendix we will prove Theorem (7.1). For convenience, we call nodes in the original grid old nodes and we call nodes added by algorithm (7.7) new nodes. We assume there are n nodes in the original grid and m $(m \geq n)$ nodes in the new grid. For $i \geq 0$, let S_i be the $(i+1)th$ node in a grid. If

$$\left(|r'| - \lambda\sigma\sqrt{1-\rho^2}\right) \geq 0, \tag{C.1}$$

the new grid will be the same as the original one. Hence, the non-trivial case is when

$$\left(|r'| - \lambda\sigma\sqrt{1-\rho^2}\right) < 0. \tag{C.2}$$

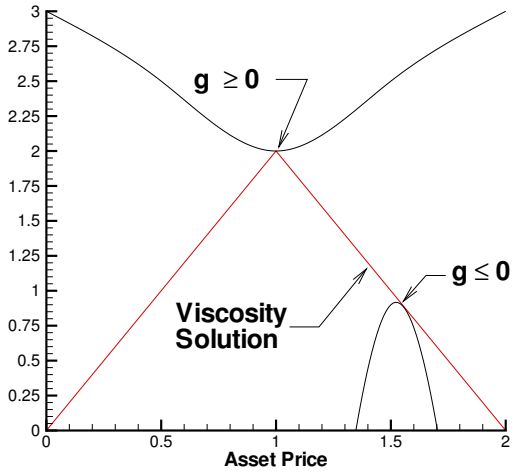


FIGURE 2: *Illustration of viscosity solution definition. The upper and lower curves represent smooth test functions. The differential operator (B.1) can be applied to these test functions with the results given by equation (B.10) (upper curve) and equation (B.12) (lower curve). When a smooth test function χ touches the viscosity solution from below at (S_0,τ_0), then $g(V(S_0,\tau_0),\chi_S(S_0,\tau_0),\chi_{SS}(S_0,\tau_0),\chi_\tau(S_0,\tau_0)) \le 0$. Similarly, when a smooth test function ϕ touches the viscosity solution from above at (S_0,τ_0), then $g(V(S_0,\tau_0),\phi_S(S_0,\tau_0),\phi_{SS}(S_0,\tau_0),\phi_\tau(S_0,\tau_0)) \ge 0$. Note that there may be some points where a smooth test function can touch the viscosity solution only from above or below, but not both. The kink at $S=1$ is an example of such a point.*

Let

$$\mathscr{K} = -\frac{\sigma^2}{\left(|r'| - \lambda\sigma\sqrt{1-\rho^2}\right)} > 0. \tag{C.3}$$

Then in the new grid for $1 \le i \le m-1$, we have (from equation (7.1))

$$S_{i+1} - S_{i-1} \le \mathscr{K} S_i. \tag{C.4}$$

Now, we prove Theorem (7.1).

Proof. Suppose Theorem (7.1) is not true. Then in the new grid $\exists i, 1 \le i \le m-1$ such that

$$\frac{S_{i+1}-S_i}{S_i - S_{i-1}} = t, \text{ where } t > q = max(5, 2q_0) \text{ or } t < p = min(1/3, p_0). \tag{C.5}$$

Let

$$a_i = S_{i+1} - S_i, \tag{C.6}$$

so

$$\frac{a_i}{a_{i-1}} = t. \tag{C.7}$$

Now suppose

$$\frac{a_i}{a_{i-1}} = t > q = max(5, 2q_0). \tag{C.8}$$

We prove the following observations first.

FIGURE 3: *Condition (7.9) failed in a new grid.*

FIGURE 4: S_{i-1} *is inserted at* $\frac{S_j+S_i}{2}$.

Observation 1. S_{i-1} *has to be a new node.*

Proof. See Figure 3. Suppose S_{i-1} is an old node. Then if S_i is also an old node, we have

$$\frac{a_i}{a_{i-1}} \leq q_0, \tag{C.9}$$

while if S_i is a new node, we have

$$\frac{a_i}{a_{i-1}} \leq 1. \tag{C.10}$$

Both cases contradict equation (C.8). Observation 1 follows. ∎

Observation 2. *When* S_{i-1} *is added into the grid,* S_i *has already been in the grid.*

Proof. See Figure 3. Otherwise we will have equation (C.10). ∎

By Observations 1 and 2, S_{i-1} must be inserted in the middle of S_i and a node S_j, where $i > j$, as depicted in Figure 4.

Observation 3. S_j *has to be a new node.*

Proof. See Figure 4. Suppose S_j is an old node. Then if S_i is also an old node, we have

$$\frac{a_i}{a_{i-1}} \leq 2q_0, \tag{C.11}$$

while if S_i is a new node, we have

$$\frac{a_i}{a_{i-1}} \leq 2. \tag{C.12}$$

Both cases contradict with equation (C.8). Observation 3 follows. ∎

Observation 4. *When* S_j *is added,* S_i *has already been in the grid.*

Proof. See Figure 3. Otherwise we will have equation (C.12). ∎

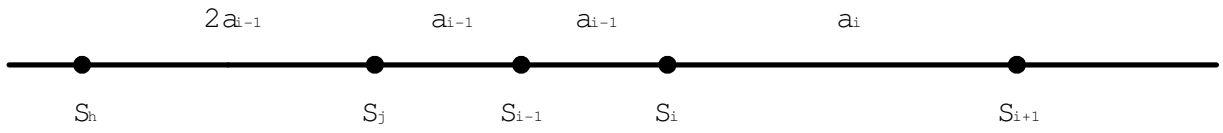


FIGURE 5: S_j *is inserted at* $\frac{S_h+S_i}{2}$.

By Observations 3 and 4, S_j must be inserted in the middle of S_i and a node S_h, where $h < j < i$, as shown in Figure 5.

Observation 5. $2S_j \geq S_i$

Proof. See Figure 5. Note that

$$S_j - S_h = S_i - S_j = 2a_{i-1} \tag{C.13}$$

and

$$S_h \geq 0. \tag{C.14}$$

This implies

$$\frac{S_j}{S_i} = \frac{S_h + 2a_{i-1}}{S_h + 4a_{i-1}} \geq \frac{2a_{i-1}}{4a_{i-1}} = \frac{1}{2}, \tag{C.15}$$

hence,

$$2S_j \geq S_i. \tag{C.16}$$

∎

Observation 6. $6a_{i-1} < \mathscr{K} S_i$

Proof. See Figure 3. Since S_{i-1}, S_i, and S_{i+1} are three consecutive nodes in the new grid and $t > 5$, we have

$$6a_{i-1} < (t+1)a_{i-1} = S_{i+1} - S_{i-1} \leq \mathscr{K} S_i. \tag{C.17}$$

∎

We now show that $\frac{a_i}{a_{i-1}} > q$ is false. Suppose it is true. By Observation 1, we know S_{i-1} is a new node.

Case 1: S_{i-1} is added because

$$\begin{aligned} S_f - S_j &> \mathscr{K} S_i \\ S_i - S_j &\geq \frac{\mathscr{K} S_i}{2}, \end{aligned} \tag{C.18}$$

where S_f is the right neighbour of S_i when S_{i-1} is added. This is shown in Figure 6. Note that $S_f \geq S_{i+1}$. Then

$$2a_{i-1} = S_i - S_j \geq \frac{\mathscr{K} S_i}{2}, \tag{C.19}$$

so

$$4a_{i-1} \geq \mathscr{K} S_i, \tag{C.20}$$

which is a contradiction with Observation 6.

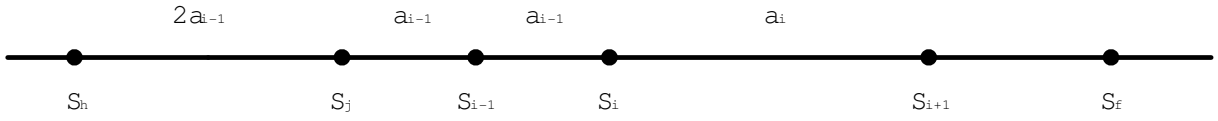


FIGURE 6: *S_{i-1} is added because condition (C.18) is true.*

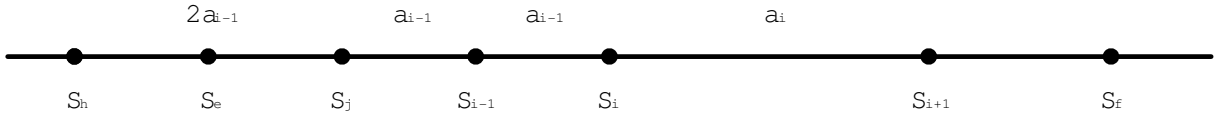


FIGURE 7: *S_{i-1} is added because condition (C.21) is true.*

Case 2: S_{i-1} is added because

$$\begin{aligned} S_i - S_e &> \mathscr{K} S_j \\ S_j - S_e &< \frac{\mathscr{K} S_j}{2}, \end{aligned} \tag{C.21}$$

where S_e is the left neighbour of S_j when S_{i-1} is added. This is illustrated in Figure 7.
Note that

$$S_e \geq S_h + a_{i-1}. \tag{C.22}$$

If this does not hold, we have

$$S_e = S_h, \tag{C.23}$$

since there can be no node between S_h and S_j when S_{i-1} is added. This gives

$$2a_{i-1} = S_j - S_h = S_j - S_e < \frac{\mathscr{K} S_j}{2}, \tag{C.24}$$

so that

$$a_{i-1} < \frac{\mathscr{K} S_j}{4}. \tag{C.25}$$

But

$$4a_{i-1} = S_i - S_h > \mathscr{K} S_j, \tag{C.26}$$

so

$$a_{i-1} > \frac{\mathscr{K} S_j}{4} \tag{C.27}$$

which is a contradiction. Hence, equation (C.22) is true. Then we have

$$3a_{i-1} \geq S_i - S_e > \mathscr{K} S_j, \tag{C.28}$$

so that

$$a_{i-1} > \frac{\mathscr{K} S_j}{3}. \tag{C.29}$$

By Observation 6 and equation (C.29), we have

$$6a_{i-1} \leq \mathscr{K} S_i, \tag{C.30}$$

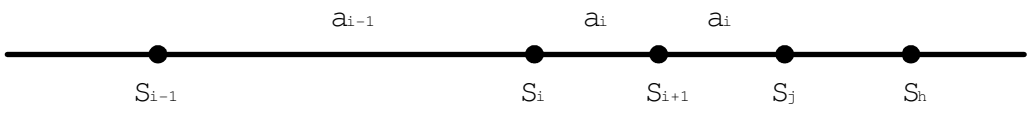


FIGURE 8: *S_{i+1} is added because condition (C.35) is true.*

implying

$$\frac{6\mathscr{K}S_j}{3} < \mathscr{K}S_i, \tag{C.31}$$

so that

$$2S_j < S_i, \tag{C.32}$$

contradicting Observation 5.

Hence, we get contradictions in both cases. Therefore,

$$\frac{a_i}{a_{i-1}} \leq q = max(5, 2q_0). \tag{C.33}$$

Now suppose

$$\frac{a_i}{a_{i-1}} = t < p = min(1/3, p_0). \tag{C.34}$$

As before, S_{i+1} is a new node, and when it is inserted, S_i has already been in the grid. Suppose S_{i+1} is in the middle of S_i and S_j, as shown in Figure 8.
Since now $a_{i-1} > a_i$, the only reason to add S_{i+1} is

$$\begin{aligned} S_h - S_i &> \mathscr{K}S_j \\ S_j - S_i &\geq \frac{\mathscr{K}S_j}{2}, \end{aligned} \tag{C.35}$$

where S_h is the right neighbour of S_j when S_{i+1} is added. Hence,

$$2a_i = S_j - S_i \geq \frac{\mathscr{K}S_j}{2} \tag{C.36}$$

so that

$$a_i \geq \frac{\mathscr{K}S_j}{4}, \tag{C.37}$$

and

$$\mathscr{K}S_j = \frac{3+1}{4}\mathscr{K}S_j < \frac{(\frac{1}{t}+1)\mathscr{K}S_j}{4} \leq \left(\frac{1}{t}+1\right)a_i = S_{i+1} - S_{i-1} < \mathscr{K}S_i. \tag{C.38}$$

This means

$$S_j < S_i, \tag{C.39}$$

which is certainly false. Hence,

$$t \geq p = \min(1/3, p_0). \tag{C.40}$$

The result follows. ∎

References

[1] C. Albanese and S. Tompaidis. Small transaction cost asymptotics and dynamic hedging. Working paper, Imperial College, Department of Mathematics, to appear in *European Journal of Operational Research*, 2004.

[2] A. L. Amadori. Nonlinear integro-differential evolution problems arising in option pricing: A viscosity solution approach. *Differential and Integral Equations: An International Journal for Theory and Applications*, 16:787–811, 2003.

[3] L. B. G. Andersen, J. Andreasen, and R. Brotherton-Ratcliffe. The passport option. *Journal of Computational Finance*, 1(3):15–36, 1998.

[4] M. Avellaneda, A. Levy, and A. Parás. Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2:73–88, 1995.

[5] G. Barles. Convergence of numerical schemes for degenerate parabolic equations arising in finance. In L. C. G. Rogers and D. Talay, editors, *Numerical Methods in Finance*, pages 1–21. Cambridge University Press, Cambridge, 1997.

[6] G. Barles and J. Burdeau. The Dirichlet problem for semilinear second-order degenerate elliptic equations and applications to stochastic exit time control problems. *Communications in Partial Differential Equations*, 20:129–178, 1995.

[7] G. Barles and E. Rouy. A strong comparison result for the Bellman equation arising in stochastic exit time control problems and applications. *Communications in Partial Differential Equations*, 23:1945–2033, 1998.

[8] G. Barles and P.E. Souganidis. Convergence of approximation schemes for fully nonlinear equations. *Asymptotic Analysis*, 4:271–283, 1991.

[9] D. Bertsimas, L. Kogan, and A. W. Lo. Hedging derivative securities in incomplete markets: An ε-arbitrage approach. *Operations Research*, 49:372–397, 2001.

[10] J. P. Bouchaud, G. Iori, and D. Sornette. Real world options: Smile and residual risk. *Risk*, 9(3):61–65, 1996.

[11] S. Chaumont. A strong comparison result for viscosity solutions to Hamilton-Jacobi-Bellman equations with Dirichlet conditions on a non-smooth boundary. Working paper, Institute Elie Cartan, Université Nancy I, 2004.

[12] T. F. Coleman, Y. Li, and M. C. Patron. Hedging guarantees in variable annuities. Working paper, Cornell Theory Center, 2004.

[13] R. Cont and P. Tankov. *Financial Modelling with Jump Processes*. Chapman & Hall, London, 2004.

[14] R. Cont and E. Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential Lévy models. Internal Report 513, CMAP, École Polytechnique, 2003.

[15] M. G. Crandall, H. Ishii, and P. L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 27:1–67, 1992.

[16] M. A. H. Davis. Optimal hedging with basis risk. Working Paper, Vienna University of Technology, 2000.

[17] Y. d'Halluin, P. A. Forsyth, and G. Labahn. A penalty method for American options with jump diffusion processes. *Numerische Mathematik*, 97:321–352, 2004.

[18] Y. d'Halluin, P. A. Forsyth, and G. Labahn. A semi-Lagrangian approach for American Asian options under jump diffusion. *SIAM Journal on Scientific Computing*, forthcoming, 2005.

[19] P. Embrechts. Actuarial versus financial pricing of insurance. *Risk Finance*, 1(4):17–26, 2000.

[20] H. Föllmer and M. Schweizer. Hedging by sequential regression: An introduction to the mathematics of option trading. *ASTIN Bulletin*, 18:147–160, 1988.

[21] P. A. Forsyth and K. R. Vetzal. Quadratic convergence of a penalty method for valuing American options. *SIAM Journal on Scientific Computation*, 23:2096–2123, 2002.

[22] D. Heath, E. Platen, and M. Schweizer. A comparison of two quadratic approaches to hedging in incomplete markets. *Mathematical Finance*, 11:385–413, 2001.

[23] H. E. Leland. Option pricing and replication with transaction costs. *Journal of Finance*, 40:1283–1301, 1985.

[24] T. Møller. Risk minimizing strategies for unit linked life insurance contracts. *ASTIN Bulletin*, 28:17–47, 1998.

[25] T. Møller. On transformations of actuarial valuation principles. *Insurance: Mathematics and Economics*, 28:281–303, 2001.

[26] T. Møller. On valuation and risk management at the interface of insurance and finance. *British Actuarial Journal*, 8:787–827, 2002.

[27] A. Oberman and T. Zariphopoulou. Pricing early exercise contracts in incomplete markets. *Computational Management Science*, 1:75–107, 2003.

[28] M. Otaka and Y. Kawaguchi. Hedging and pricing of real estate securities under market incompleteness. Quantitative Methods in Finance Conference, Cairns, Australia, 2002.

[29] D. M. Pooley. *Numerical Methods for Nonlinear Equations in Option Pricing*. PhD thesis, School of Computer Science, University of Waterloo, 2003.

[30] D. M. Pooley, P. A. Forsyth, and K. R. Vetzal. Numerical convergence properties of option pricing PDEs with uncertain volatility. *IMA Journal of Numerical Analysis*, 23:241–267, 2003.

[31] L. Qi and J. Sun. A nonsmooth version of Newton's method. *Mathematical Programming*, 58:353–367, 1993.

[32] L. Qi and G. Zhou. A smoothing Newton method for minimizing a sum of Euclidean norms. *SIAM Journal on Optimization*, 11:389–410, 2000.

[33] R. Rannacher. Finite element solution of diffusion problems with irregular data. *Numerische Mathematik*, 43:309–327, 1984.

[34] M. Schal. On quadratic cost criteria for option hedging. *Mathematics of Operations Research*, 19:121–131, 1994.

[35] M. Schweizer. Variance optimal hedging in discrete time. *Mathematics of Operations Research*, 20:1–31, 1995.

[36] M. Schweizer. From actuarial to financial valuation principles. *Insurance: Mathematics and Economics*, 28:31–47, 2001.

[37] P. Wilmott. *Derivatives: The Theory and Practice of Financial Engineering*. John Wiley & Sons Ltd., West Sussex, England, 1998.

[38] H. Windcliff, P. A. Forsyth, and K. R. Vetzal. Segregated funds: Shout options with maturity extensions. *Insurance, Mathematics and Economics*, 29:1–21, 2001.