Hedging with a Correlated Asset:

An Insurance Approach

by

Jian Wang

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada 2005

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Hedging a contingent claim with an asset which is not perfectly correlated with the underlying asset results in an imperfect hedge. The residual risk from hedging with a correlated asset is priced using an actuarial standard deviation principle in infinitesmal time, which gives rise to a nonlinear partial differential equation (PDE). A fully implicit, monotone discretization method is developed for solving the 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, which show the standard deviation of the profit and loss at the expiry of the option.

Acknowledgements

First, sincere gratitude is extended to my supervisor, Dr. Peter Forsyth, for his considerable guidance in choosing the topic of my thesis and keeping my research on the right track, his great help in solving the core mathematical problems and with writing the academic paper. I have greatly benefited from his experience and creative ideas.

I would like to thank Dr. Heath A. Windcliff for developing the mathematical model which forms the basis of this thesis. I am grateful to Dr. Ken Vetzal for his ideas and guidance, which helped drive and shape this work. I would also like to thank my readers, Dr. Bruce Simpson and Dr. Ken Vetzal (again).

Members of the Scientific Computation Group, both past and present, have always been there when needed. Thank you all for your friendship and help. In particular, I would thank Shannon Kennedy for his great help in both my course work and research.

Finally, I would like to thank my parents, Xiangyang Wang and Xiaofeng Guo, who have always been there to support and encourage me. Numerous words and emotions merge into one sentence: THANK YOU AND I LOVE YOU!

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

A derivative security, simply called a derivative, is a financial security, whose characteristics and value depend on the characteristics and value of other securities (often referred to as underlying securities). Derivatives play a very important role in modern financial markets. Derivatives are usually hedged by taking positions in the underlying assets. The standard Black-Sholes analysis assumes that delta hedging eliminates all risk. However, underlying assets cannot be traded in some situations. It follows that other assets, which are correlated with the underlying assets, must be used to hedge the derivative.

As a motivation, consider the following situation. Segregated funds [39] are guarantees on pension plan investment accounts, offered by Canadian insurance companies. In many cases, the underlying asset is a mutual fund, managed by the insurance company offering the guarantee. Since the insurance company cannot legally short its own funds, these guarantees are hedged using index futures. The index, of course, will not be perfectly correlated with the underlying

mutual fund.

In this case, there is unhedgeable residual risk because of the imperfect correlation. As is well known [19], it is possible to construct a best local hedge, in the sense that the residual risk is orthogonal to the risk which is hedged. If an index is used to construct the hedge, and the residual risk is not correlated with the market index, it could be argued that this residual risk is firm specific, and can be diversified away. However, since it is not easy for insurance policy holders to diversify their risk, insurance companies are mandated to have sufficient reserves to guarantee solvency. As a result, the usual approach in the insurance industry is to build up a reserve to provide for unhedgeable risk. We are then left with the problem of determining an appropriate pricing mechanism for an option with unhedgeable residual risk.

One possible approach is based on utility maximization [24, 15]. However, it is not obvious how to construct a utility function for an insurance company. As discussed in [8], expected utility maximization approaches have had limited acceptance in practice.

1.2 Overview

In this thesis, we will follow along the lines suggested in [7], where the expected return of the hedging portfolio is adjusted to reflect a risk premium due to the unhedgeable risk. More recently, this idea has also been recognized as a common actuarial valuation principle [27, 37]. This approach is also known as a *safety loading*, in the sense that insurance companies must charge premia larger than the expected payoff of the hedging portfolio (in incomplete markets), so that sufficient reserves are built up to ensure solvency. In [37], 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.

We will use the *modified standard deviation principle* [25] in infinitesimal time to derive a nonlinear PDE for the value of a contingent claim in the incomplete market case where the underlying cannot be traded, but the the claim can be partially hedged using a correlated asset. The best hedge is determined by local risk minimization [20, 35, 36]. We then use the actuarial standard deviation valuation principle to account for the residual risk. This provides the hedger with compensation for bearing the unhedgeable risk. The standard deviation principle is used as opposed to the variance principle, since the standard deviation method gives a value which is linear in terms of the number of units bought/sold [25]. Applying this principle in infinitesmal time results in a method which is easily extended to American style contracts with complex path dependent features, which are typical of pension portfolio guarantees offered by insurance firms. In a financial context, this method can also be interpreted as specifying a drift rate which ensures that the hedge portfolio has a desired Sharpe ratio.

The nonlinear PDE gives a different price 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 [2, 38, 31]. The pricing equation resulting from hedging with a correlated asset requires estimation of the objective measure drift rates, which are difficult to ascertain. An uncertain drift rate model results in a nonlinear PDE with the same form as the nonlinear PDE based on the actuarial standard deviation principle. 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.

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 [3, 14] 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 cer-

tain 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, VaR and CVaR of the final profit and loss of the hedging portfolio at the expiry time of the contingent claim.

1.3 Outline

The outline of this thesis is as follows. In Chapter 2 we introduce the financial model and formulate the pricing PDE. Boundary conditions are also studied carefully. In Chapter 3 we give one method used to discretize the pricing PDE. In Chapter 4 we show that the numerical discretizations are stable, consistent and monotone, so that they are guaranteed to converge to the viscosity solution. We also provide a iterative method for solution of the nonlinear discretized algebraic equations derived in Chapter 3. In Chapter 5 we introduce a node insertion algorithm and study

the grid aspect ratio after application of the algorithm. In Chapter 6 we give a number of numerical examples which illustrate the performance and convergence of our iteration scheme. We experiment with a minimum value in the asset grid (S_0) and show that the solution is insensitive to small positive S_0. The node insertion algorithm is stressed by experimenting with a very poor atypical initial grid. In Chapter 7 we experiment with Monte Carlo hedging simulations. Finally, we provide some conclusions in Chapter 8.

Chapter 2

Pricing Model Formulation

In this chapter we use the *modified standard deviation principle* to formulate the pricing PDE. We start with European options, and then extend the model to price American options as well.

2.1 The Nonlinear PDE

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

$$dS = \mu S\,dt + \sigma S\,dZ\,, \tag{2.1.1}$$

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

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

$$dH = \mu' H\,dt + \sigma' H\,dW\,, \tag{2.1.2}$$

where dW is the increment of a Wiener process. In the following we will use the usual Wiener

process properties

$$E[dW^2] = dt \;\; ; \;\; E[dZ^2] = dt \;\; ; \;\; E[dZ\, dW] = \rho\, dt \tag{2.1.3}$$

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

$$\Pi \;\; = \;\; -V + xH + B\,, \tag{2.1.4}$$

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

$$\begin{aligned} d\Pi &= -\left[V_t + \mu S V_S + \frac{\sigma^2 S^2}{2} V_{SS}\right] dt - \sigma S V_S\, dZ \\ &\quad + r(V - xH)\, dt + x(\mu' H\, dt + \sigma' H\, dW) \\ &= -\left[V_t + \mu S V_S + \frac{\sigma^2 S^2}{2} V_{SS} + (xH - V)r - x\mu' H\right] dt \\ &\quad - \sigma S V_S\, dZ + x\sigma' H\, dW\,. \end{aligned} \tag{2.1.5}$$

The variance of $d\Pi$ is given by

$$E^P\left[(x\sigma' H\, dW - \sigma S V_S\, dZ)^2\right] \;\; = \;\; \left[x^2(\sigma')^2 H^2 + \sigma^2 S^2 V_S^2 - 2\sigma S V_S x \sigma' H \rho\right] dt \tag{2.1.6}$$

where $E^P[\cdot]$ is the expectation operator under the objective or P measure, and we have used properties (2.1.3). Choosing x to minimize equation (2.1.6) gives

$$x \;\; = \;\; \left(\frac{\sigma S \rho}{\sigma' H}\right) V_S\,. \tag{2.1.7}$$

Substituting equation (2.1.7) into equation (2.1.5) gives

$$\begin{aligned} d\Pi &= -\left[V_t + \mu S V_S + \frac{\sigma^2 S^2}{2} V_{SS} - rV + \left(\frac{r\sigma S \rho}{\sigma'}\right) V_S \right. \\ &\quad \left. - \left(\frac{\sigma S \rho \mu'}{\sigma'}\right) V_S\right] dt - \sigma S V_S\, dZ + \sigma S V_S \rho\, dW\,. \end{aligned} \tag{2.1.8}$$

Defining

$$r' = \mu - (\mu' - r)\frac{\sigma\rho}{\sigma'} \tag{2.1.9}$$

in equation (2.1.8) gives

$$d\Pi = -\left[V_t + SV_S r' - rV + \frac{\sigma^2 S^2}{2} V_{SS}\right] dt + \sigma S V_S(\rho\, dW - dZ) \,. \tag{2.1.10}$$

Substituting equation (2.1.7) into equation (2.1.6) results in

$$var[d\Pi] = (1-\rho^2)\sigma^2 V_S^2 S^2\, dt \,. \tag{2.1.11}$$

Noting that

$$cov[\rho\, dW - dZ, dW] = 0 \tag{2.1.12}$$

we obtain

$$cov[d\Pi, dW] = 0\,, \tag{2.1.13}$$

so that the residual risk is orthogonal (in this sense) to the hedging instrument.

Define a new Brownian increment

$$dX = \frac{1}{\sqrt{1-\rho^2}}[\rho dW - dZ] \tag{2.1.14}$$

with the property

$$dX^2 = dt \,. \tag{2.1.15}$$

This allows us to write equation (2.1.10) as

$$\begin{aligned} d\Pi = & -\left[V_t + SV_S r' - rV + \frac{\sigma^2 S^2}{2} V_{SS}\right] dt \\ & + SV_S\sqrt{1-\rho^2}\,\sigma\, dX \,. \end{aligned} \tag{2.1.16}$$

We can also write equation (2.1.16) as

$$d\Pi \;=\; (\text{deterministic component})\; dt + SV_S\sqrt{1-\rho^2}\;\sigma\; dX\;. \tag{2.1.17}$$

Based on equation (2.1.17), a possible pricing approach is to require

$$E^P[d\Pi] = 0\;. \tag{2.1.18}$$

However, an insurance company which charged premia based only on equation (2.1.18) would soon have solvency problems [18]. As discussed in [27], insurance companies typically charge a premium for unhedgeable risk. We will use the actuarial standard deviation principle in infinitesmal time. In our notation, this becomes

$$E^P[d\Pi] \;=\; \lambda\sqrt{\frac{var[d\Pi]}{dt}}\;dt \tag{2.1.19}$$

where λ is the *safety loading* parameter. In order words, during each interval $[t,t+dt]$, the portfolio should earn a premium at a rate proportional to the 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.

Alternatively, we can derive equation (2.1.19) using more typical financial reasoning. Recall that $\Pi(t) = 0$, so that any gain in this portfolio must be due to risk. Let $G(t,dt)$ be the expected rate of gain of the portfolio, and $R(t,dt)$ be the risk associated with this portfolio, as measured by the square root of the variance per unit time

$$\begin{aligned} G(t,dt) &= \frac{E^P[d\Pi]}{dt} \\ R(t,dt) &= \sqrt{\frac{var[d\Pi]}{dt}}\;. \end{aligned} \tag{2.1.20}$$

We now seek to impose the condition that the gain lies on the efficient frontier

$$G(t,dt) \;=\; \lambda R(t,dt) \tag{2.1.21}$$

where λ is the safety loading parameter, or Sharpe ratio. We can write equation (2.1.21) in a more familiar form if we note that

$$\Pi = \Pi' + B \; ; \; \Pi' = -V + xH \; ; \; B = V - xH \; , \tag{2.1.22}$$

so that

$$d\Pi = d\Pi' + rB \; dt = d\Pi' - r\Pi' \; dt \; . \tag{2.1.23}$$

Combining equations (2.1.19) and (2.1.23) gives

$$\frac{E^P[d\Pi']}{dt} \;\; = \;\; r\Pi' + \lambda \sqrt{\frac{var[d\Pi']}{dt}} \; , \tag{2.1.24}$$

or

$$\begin{aligned} \frac{E^P[\frac{d\Pi'}{\Pi'}]}{dt} &= r + \lambda \sigma_{eff} \\ \sigma_{eff} &= \sqrt{\frac{var[\frac{d\Pi'}{\Pi'}]}{dt}} \; , \end{aligned} \tag{2.1.25}$$

So, we require that the instantaneous systematic gain in the portfolio compensates for the expected risk. 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 (2.1.11) we have that

$$\sqrt{\frac{var[d\Pi]}{dt}} = \sigma S |V_S| \sqrt{1-\rho^2} \; . \tag{2.1.26}$$

Combining equations (2.1.16, 2.1.19, 2.1.26) gives

$$V_t + SV_S r' + S|V_S| \lambda \sigma \sqrt{1-\rho^2} + \frac{\sigma^2 S^2}{2} V_{SS} - rV = 0 \; , \tag{2.1.27}$$

or equivalently

$$V_t + SV_S\left[r' + \lambda\sigma\sqrt{1-\rho^2}\,\mathrm{sgn}(V_S)\right] + \frac{\sigma^2 S^2}{2}V_{SS} - rV = 0\,. \tag{2.1.28}$$

Note that the definition of Π in equation (2.1.4) assumed that the hedger was short the contingent claim V. Consequently, equation (2.1.28) is valid for a short position in V. Repeating the above arguments for a long position gives

$$V_t + SV_S\left[r' - \lambda\sigma\sqrt{1-\rho^2}\,\mathrm{sgn}(V_S)\right] + \frac{\sigma^2 S^2}{2}V_{SS} - rV = 0\,. \tag{2.1.29}$$

For future reference, note that the two cases are

$$\begin{aligned}
&\text{Short Position:} \quad V_t + SV_S\left[r' + \lambda\sigma\sqrt{1-\rho^2}\,\mathrm{sgn}(V_S)\right] + \frac{\sigma^2 S^2}{2}V_{SS} - rV = 0\\
&\text{Long Position:} \quad V_t + SV_S\left[r' - \lambda\sigma\sqrt{1-\rho^2}\,\mathrm{sgn}(V_S)\right] + \frac{\sigma^2 S^2}{2}V_{SS} - rV = 0\,.
\end{aligned} \tag{2.1.30}$$

From equation (2.1.16) and (2.1.30) we have that

$$\begin{aligned}
&\text{Short Position:} \quad d\Pi = \lambda\sigma\sqrt{1-\rho^2}S|V_S|\,dt + SV_S\sqrt{1-\rho^2}\sigma\,dX\\
&\text{Long Position:} \quad d\Pi = \lambda\sigma\sqrt{1-\rho^2}S|V_S|\,dt - SV_S\sqrt{1-\rho^2}\sigma\,dX\,.
\end{aligned} \tag{2.1.31}$$

Note that the drift term r' can be difficult to estimate since r' is a function of the P measure drift rates of S and H. In [2], an uncertain volatility model was proposed. This model assumed that volatility was bounded within an upper and lower bound. Worst case pricing methods ensure that the hedging portfolio has a nonnegative balance, regardless of the evolution of the uncertain volatility. We can follow a similar approach here. Suppose we estimate bounds for r', i.e.

$$r' \in [r'_{min}, r'_{max}] \tag{2.1.32}$$

then, we can use an uncertain drift rate model [38] to obtain worst case prices for the option. Note that if $|\rho| \neq 1$, then the worst case pricing cannot ensure that the hedging portfolio has a

nonnegative balance. Define

$$\begin{aligned} r^* &= \frac{r'_{max} + r'_{min}}{2} \\ \lambda^* &= \lambda + \frac{r'_{max} - r'_{min}}{2\sigma\sqrt{1-\rho^2}} \end{aligned} \tag{2.1.33}$$

so that equation (2.1.30) becomes (worst case prices)

$$\text{Short Position:} \qquad V_t + SV_S\left[r^* + \lambda^*\sigma\sqrt{1-\rho^2}\,\text{sgn}(V_S)\right] + \frac{\sigma^2S^2}{2}V_{SS} - rV = 0$$

$$\text{Long Position:} \qquad V_t + SV_S\left[r^* - \lambda^*\sigma\sqrt{1-\rho^2}\,\text{sgn}(V_S)\right] + \frac{\sigma^2S^2}{2}V_{SS} - rV = 0\ . \tag{2.1.34}$$

Observe that even if $\lambda = 0$, a nonlinear PDE of the form (2.1.34) is obtained, due to the uncertainty in r'.

Assuming $\lambda \geq 0$, then equations (2.1.30) are equivalent to

$$\text{Short Position:} \qquad V_\tau = \max_{q\in\{-1,+1\}}\left[SV_S\left[r' + q\,\lambda\sigma\sqrt{1-\rho^2}\,\right] + \frac{\sigma^2S^2}{2}V_{SS} - rV\right]$$

$$\text{Long Position:} \qquad V_\tau = \min_{q\in\{-1,+1\}}\left[SV_S\left[(r' + q\,\lambda\sigma\sqrt{1-\rho^2})\right] + \frac{\sigma^2S^2}{2}V_{SS} - rV\right] \tag{2.1.35}$$

where we have defined $\tau = T - t$, where T is the expiry time of the contingent claim. Note that the optimal choice for q in equation (2.1.35) is

$$q = \begin{cases} +\text{sgn}(V_S) & \text{if short} \\ -\text{sgn}(V_S) & \text{if long} \end{cases}\ . \tag{2.1.36}$$

If we write

$$\mathcal{L}V \equiv V_\tau - \left\{SV_S\left[r' + q\,\lambda\sigma\sqrt{1-\rho^2}\,\right] + \frac{\sigma^2S^2}{2}V_{SS} - rV\right\} \tag{2.1.37}$$

with the payoff denoted by $V = V^*$, then the price of a contingent claim with an American early exercise feature would be given by

$$\min(\mathcal{L}V, V - V^*) = 0\ . \tag{2.1.38}$$

We will focus on European options in this thesis, but much of the analysis can be extended to the American case as well.

2.2 Boundary Conditions

At $\tau = 0$, we set $V(S,0)$ to the payoff. As $S \to 0$, equation (2.1.35) reduces to

$$V_\tau = -rV \ . \tag{2.2.1}$$

In fact, in order to ensure certain properties of the discrete equations, we will impose condition (2.2.1) 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 approximation (2.2.1) through numerical tests.

As $S \to \infty$, we make the common assumption that

$$V_{SS} \simeq 0 \ \ ; \ \ S \to \infty \tag{2.2.2}$$

which means that

$$V \simeq A(\tau)S + B(\tau) \ \ ; \ \ S \to \infty. \tag{2.2.3}$$

Assuming equation (2.2.3) holds, then substituting equation (2.2.3) into equation (2.1.35) gives ordinary differential equations for $A(\tau), B(\tau)$, with solution

$$V \ = \ A(0)S \ \exp\left[(r' - r + q\ \lambda\sigma\sqrt{1-\rho^2})\tau\right] + B(0)e^{-r\tau} \tag{2.2.4}$$

where q is given from equation (2.1.36) at $\tau = 0$. The initial conditions for $A(0), B(0)$ are given from the option payoff.

2.3 Relation to Previous Work

We can relate equation (2.1.5) to the work in [35] by noting that for $\lambda = 0$, $d\Pi$ is the incremental profit of hedging. (In [35], 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 back to zero, cash is added or subtracted from the portfolio so that

$$\Pi(t+dt^+) = \Pi(t+dt^-) - d\Pi(t) = 0 \,, \tag{2.3.1}$$

hence this portfolio is not self-financing.

If $\lambda = 0$, then the approach used above is based on local risk minimization [35], i.e. we choose the trading strategy to minimize the square of the incremental hedging profit/loss at each hedging time. Note that if $\lambda = 0$, then from equation (2.1.31) we have that $E^P[d\Pi] = 0$, hence this strategy is mean self-financing.

Given that the payoff of the option is used as an initial condition for equations (2.1.35) 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 [11], using the hedging parameters (2.1.7) given from the solution to equation (2.1.30), then 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 Chapter 7.

The local risk minimization approach can be contrasted with the mean variance hedging or total risk minimization approach [36, 23]. 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 [11], total risk minimization is a dynamic stochastic programming problem which is difficult, in general, to solve. In this thesis, 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 [39].

We have combined the local risk minimization concept with an infinitesimal time modified standard deviation principle [25, 37]. Consequently, the holder of the hedge portfolio is compensated for the unhedgeable risk by receiving an expected return above the risk free rate. Note that this return is a function of the unhedgeable risk, which depends on the current state of the hedging portfolio. Hence, it seems appropriate to apply the standard deviation principle in infinitesimal time, since the capital at risk in the hedging portfolio will change with time.

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 [29].

It is curious to note that if we had specified an actuarial variance principle,

$$E^P[d\Pi] \quad = \quad \lambda^V \left[\frac{var[d\Pi]}{dt} \right] \tag{2.3.2}$$

then we would obtain a nonlinear PDE identical to the PDE derived in [28], which was derived using a utility maximization approach. (Note that the PDE in [28] is written for the case $r = 0$.)

Chapter 3

Discretization

In this chapter we give a brief overview of the discrete equations used in following chapters.

3.1 Summary of Discretization

For discretization purposes, PDEs (2.1.35) can be written as

$$V_\tau = SV_S\left[r' + q\lambda\sigma\sqrt{1-\rho^2}\right] + \frac{\sigma^2 S^2}{2}V_{SS} - rV \tag{3.1.1}$$

where the nonlinear term q is given from equation (2.1.36). Define a grid $\{S_0, S_1, ..., S_p\}$, and let $V_i^n = V(S_i, \tau^n)$.

Equation (3.1.1) can be discretized using forward, backward or central differencing in the S direction, coupled with a fully implicit timestepping to give

$$V_i^{n+1} - V_i^n \quad = \quad \alpha_i^{n+1}V_{i-1}^{n+1} + \beta_i^{n+1}V_{i+1}^{n+1} - (\alpha_i^{n+1} + \beta_i^{n+1} + r\Delta\tau)V_i^{n+1}\ , \tag{3.1.2}$$

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 (2.1.35). Let

$$\begin{aligned} \alpha_i^n &= \alpha_i' - q_{i,cent}^n \gamma_{i,cent}' - q_{i,back}^n \gamma_{i,back}' \\ \beta_i^n &= \beta_i' + q_{i,cent}^n \gamma_{i,cent}' + q_{i,for}^n \gamma_{i,for}' \ . \end{aligned} \tag{3.1.3}$$

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 (3.1.2) in the form

$$\begin{aligned} V_i^{n+1} - V_i^n &= \alpha_i' V_{i-1}^{n+1} + \beta_i' V_{i+1}^{n+1} - (\alpha_i' + \beta_i' + r\Delta\tau) V_i^{n+1} \\ &\quad + \kappa\, \gamma_{i,back}' |V_i^{n+1} - V_{i-1}^{n+1}| + \kappa\, \gamma_{i,for}' |V_{i+1}^{n+1} - V_i^{n+1}| \\ &\quad + \kappa\, \gamma_{i,cent}' |V_{i+1}^{n+1} - V_{i-1}^{n+1}| \ , \end{aligned} \tag{3.1.4}$$

where

$$\kappa = \begin{cases} +1 & \text{if short} \\ -1 & \text{if long} \end{cases} \ . \tag{3.1.5}$$

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 (2.2.4) at $S = S_{imax}$ be given by

$$D_{imax}^{n+1} = A(0) S_{imax} \exp\left[(r' - r + q\, \lambda\sigma\sqrt{1-\rho^2}) \tau^{n+1} \right] + B(0) \exp\left[-r\tau^{n+1} \right] \ , \tag{3.1.6}$$

For further notational convenience, we can write equation (3.1.2) in matrix form. Let

$$\begin{aligned} V^{n+1} &= [V_0^{n+1}, V_1^{n+1}, \ldots, V_{imax}^{n+1}]' \\ V^n &= [V_0^n, V_1^n, \ldots, V_{imax}^n]' \end{aligned} \tag{3.1.7}$$

and

$$\left[\hat{M}^n V^n \right]_i = -\left[(-\alpha_i^n - \beta_i^n - r\Delta\tau) V_i^n + \alpha_i^n V_{i-1}^n + \beta_i^n V_{i+1}^n \right] \ ; \ i < imax \ . \tag{3.1.8}$$

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 (2.2.1)) is enforced by setting $\alpha_i = \beta_i = 0$ at $i = 0$. Let $D^{n+1} = [0, \ldots, D^{n+1}_{imax}]'$, 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 (3.1.8) hold only for $i < imax$, with $(\hat{M})_{imax} \equiv 0$.

$$\hat{M}^n = \begin{bmatrix} r\Delta\tau & 0 & & & & \\ -\alpha_1^n & \alpha_1^n + \beta_1^n + r\Delta\tau & -\beta_1^n & & & \\ & & \ddots & & & \\ & & & -\alpha^n_{imax-1} & \alpha^n_{imax-1} + \beta^n_{imax-1} + r\Delta\tau & -\beta^n_{imax-1} \\ 0 & \ldots & \ldots & & & 0 \end{bmatrix} \tag{3.1.9}$$

The discrete equations (3.1.2) can then be written as

$$\left[I + (1-\theta)\hat{M}^{n+1}\right] V^{n+1} = \left[I - \theta\hat{M}^n\right] V^n + I^*(D^{n+1} - V^n) \tag{3.1.10}$$

where the term $I^*(D^{n+1} - V^n)$ enforces the boundary condition at $S = S_{imax}$, and we have generalized the discretization (3.1.2) to the Crank Nicolson ($\theta = 1/2$) or fully implicit ($\theta = 0$) cases. Note that the discrete equations (3.1.10) are nonlinear since $\hat{M}^{n+1} = \hat{M}(V^{n+1})$.

Chapter 4

Convergence to the Viscosity Solution

In [31], examples were given whereby 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 [3, 14]. Equation (2.1.35) satisfies the Strong Comparison result [4, 5, 9]. Hence from [6, 3], a numerical scheme converges to the viscosity solution if the 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 (3.1.2) is a positive coefficient method, with $\alpha_i^n, \beta_i^n \geq 0$. This can be enforced by selecting a grid, and choosing forward, backward or central differencing so that the following condition is satisfied:

Condition 4.1.1 (Positive Coefficient Condition).

$$\begin{aligned} \beta'_i - \gamma'_{i,\text{cent}} - \gamma'_{i,\text{for}} &\geq 0 \ ; \ i = 0, ..., imax-1 \\ \alpha'_i - \gamma'_{i,\text{cent}} - \gamma'_{i,\text{back}} &\geq 0 \ ; \ i = 0, ..., imax-1 \end{aligned} \tag{4.1.1}$$

Note from the definitions of γ'_i in equations (A.0.17-A.0.19) that at each node, only one of $\gamma'_{i,\text{cent}}, \gamma'_{i,\text{for}}, \gamma'_{i,\text{back}}$ is nonzero, and that $\gamma'_i \geq 0$. Condition (4.1.1) is based on the worst case choice of q_i^n in equation (3.1.3), 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 (4.1.1) 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 [30] which will be used in later Chapters. We will also give an algorithm in Section 5, which, given an arbitrary grid, can satisfy condition (4.1.1) by insertion of a finite number of nodes.

Given condition (4.1.1), we then have the following stability result

Lemma 4.1.1 (Stability of discretization (3.1.2)). *Provided that*

- $r \geq 0$,
- *condition (4.1.1) is satisfied, and*
- *Dirichlet boundary conditions (2.2.1) and (2.2.3) are imposed,*

then the fully implicit discretization (3.1.2) is unconditionally stable in the sense that

$$\|V^{n+1}\|_\infty \leq \max(\|V^n\|_\infty, D_{imax}^{n+1}) \tag{4.1.2}$$

independent of the timestep size.

Proof. If conditions (4.1.1) are satisfied and $r \geq 0$, then it follows from equation (3.1.3) that α_i^n, β_i^n in discretization (3.1.2) 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 [3]. We write equation (3.1.2-3.1.4) as

$$\begin{aligned} g_i(V_i^{n+1},V_{i-1}^{n+1},V_{i+1}^{n+1},V_i^n) &= -(V_i^{n+1}-V_i^n)+\alpha_i^{n+1}V_{i-1}^{n+1}+\beta_i^{n+1}V_{i+1}^{n+1}-(\alpha_i^{n+1}+\beta_i^{n+1}+r\Delta\tau)V_i^{n+1} \\ &= -(V_i^{n+1}-V_i^n)+\alpha_i'V_{i-1}^{n+1}+\beta_i'V_{i+1}^{n+1}-(\alpha_i'+\beta_i'+r\Delta\tau)V_i^{n+1} \\ &\quad +\kappa\,\gamma'_{i,back}|V_i^{n+1}-V_{i-1}^{n+1}|+\kappa\,\gamma'_{i,for}|V_{i+1}^{n+1}-V_i^{n+1}|+\kappa\,\gamma'_{i,cent}|V_{i+1}^{n+1}-V_{i-1}^{n+1}| \\ &= 0\ ,\ \ i=0,\ldots,imax-1\ , \end{aligned} \tag{4.2.1}$$

where κ is defined in equation (3.1.5).

Definition 4.2.1 (Monotonicity). *A discretization of the form (4.2.1) is monotone if the following conditions hold*

$$g_i(V_i^{n+1},V_{i-1}^{n+1}+\varepsilon_1,V_{i+1}^{n+1}+\varepsilon_2,V_i^n+\varepsilon_3) \geq g_i(V_i^{n+1},V_{i-1}^{n+1},V_{i+1}^{n+1},V_i^n)\ ;\ \forall\varepsilon_i\geq 0 \tag{4.2.2}$$

$$g_i(V_i^{n+1}+\varepsilon_4,V_{i-1}^{n+1},V_{i+1}^{n+1},V_i^n) < g_i(V_i^{n+1},V_{i-1}^{n+1},V_{i+1}^{n+1},V_i^n)\ ;\ \forall\varepsilon_4\geq 0\ . \tag{4.2.3}$$

Observe that Definition (4.2.1) includes condition (4.2.3), whereas only condition (4.2.2) is usually specified in the viscosity solution literature [3]. Condition (4.2.3) leads to a more intuitively appealing interpretation, and is a consequence of condition (4.2.2) and consistency [17].

Lemma 4.2.1 (Monotonicity). *If condition (4.1.1) is satisfied, then discretization (4.2.1) is monotone.*

Proof. For $i = imax$, we have that $V_{imax}^{n+1} = D_{imax}^{n+1}$, hence the result holds trivially at $i = imax$. For

$i < imax$, from equation (4.2.1) we have that, for $\varepsilon \geq 0$, and noting that $\gamma'_i \geq 0$ (see Appendix A)

$$\begin{aligned}
& g_i(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1} + \varepsilon, V_i^n) - g_(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n) \\
&= \beta'_i\varepsilon + \kappa\gamma'_{i,cent}\left[|V_{i+1}^{n+1} + \varepsilon - V_{i-1}^{n+1}| - |V_{i+1}^{n+1} - V_{i-1}^{n+1}|\right] + \kappa\gamma'_{i,for}\left[|V_{i+1}^{n+1} + \varepsilon - V_i^{n+1}| - |V_{i+1}^{n+1} - V_i^{n+1}|\right] \\
&\geq \beta'_i\varepsilon - \gamma'_{i,cent}\varepsilon - \gamma'_{i,for}\varepsilon \\
&= \varepsilon(\beta'_i - \gamma'_{i,cent} - \gamma'_{i,for}) \\
&\geq 0
\end{aligned} \tag{4.2.4}$$

which follows from condition (4.1.1). Similarly,

$$\begin{aligned}
& g_i(V_i^{n+1}, V_{i-1}^{n+1} + \varepsilon, V_{i+1}^{n+1}, V_i^n) - g_(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n) \\
&\geq \varepsilon(\alpha'_i - \gamma'_{i,cent} - \gamma'_{i,back}) \\
&\geq 0,
\end{aligned} \tag{4.2.5}$$

and

$$\begin{aligned}
& g_i(V_i^{n+1} + \varepsilon, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n) - g_i(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n) \\
&\leq -\varepsilon - \varepsilon(\alpha'_i + \beta'_i + r\Delta\tau) + \varepsilon\gamma'_{i,back} + \varepsilon\gamma'_{i,for} \\
&= -\varepsilon(1 + r\Delta\tau) - \varepsilon(\alpha'_i - \gamma'_{i,back}) - \varepsilon(\beta'_i - \gamma'_{i,for}) \\
&\leq 0\,.
\end{aligned} \tag{4.2.6}$$

Finally, it is obvious from equation (4.2.1) that

$$\begin{aligned}
& g_i(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n + \varepsilon) - g_i(V_i^{n+1}, V_{i-1}^{n+1}, V_{i+1}^{n+1}, V_i^n) \\
&> 0\,.
\end{aligned} \tag{4.2.7}$$

∎

4.3 Consistency

The discrete scheme (3.1.10) is locally consistent with PDE (3.1.1) if the discrete operator applied to any C^∞ function converges to the equation (3.1.1) as the mesh size and timestep vanishes.

Lemma 4.3.1 (Consistency). *The discrete scheme (3.1.10) is locally consistent.*

Proof. From the definitions of the discrete coefficients α_i, β_i in equation (3.1.2) and Appendix A, a simple Taylor series verifies consistency. ∎

4.4 Convergence

Let $\Delta\tau = \max_n \tau^{n+1} - \tau^n$, $\Delta S = \max_i S_{i+1} - S_i$, then can now state our convergence result.

Theorem 4.4.1 (Convergence of the fully implicit discretization). *Provided that*

- $r \geq 0$,
- *the Dirichlet boundary conditions (2.2.1-2.2.3) are imposed, and condition (2.2.1) is imposed at* $S_{min} \to 0$ *as* $\Delta S \to 0$, *and*
- *the positive coefficient condition (4.1.1) holds,*

then the fully implicit discretization (3.1.2) converges unconditionally to the viscosity solution of the nonlinear PDE (3.1.1) as $\Delta S, \Delta\tau \to 0$.

Proof. Since PDE (3.1.1) satisfies the strong comparison principle, then from [3], we have that a consistent, stable, monotone discretization converges to the viscosity solution of PDE (3.1.1). Hence Theorem 4.4.1 follows directly from the results in [3] and Lemmas 4.1.1, 4.2.1 and 4.3.1. ∎

4.5 Solution of the Nonlinear Algebraic Equations

Although we have proven that the discretization converges to the viscosity solution, it is not at this point clear that scheme (3.1.2) 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 (3.1.10)

Iteration

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

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

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

Solve

$$\begin{aligned} &\left[I + (1-\theta)\hat{M}(\hat{V}^k)\right]\hat{V}^{k+1} \\ &= \left[I - \theta\hat{M}(V^n)\right]V^n + I^*(D^{n+1} - V^n) \end{aligned} \tag{4.5.1}$$

If $\max_i \dfrac{|\hat{V}_i^{k+1} - \hat{V}_i^k|}{\max(\text{scale}, |\hat{V}_i^{k+1}|)} < \textit{tolerance}$ then quit

EndFor

The scale factor in algorithm 4.5.1 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 (4.5.1) results in

$$\left[I + (1-\theta)\hat{M}^k)\right](\hat{V}^{k+1} - \hat{V}^k) \;=\; (1-\theta)\left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k\,, \tag{4.5.2}$$

where $\hat{M}^k = \hat{M}(\hat{V}^k)$. A key property which can be used to establish convergence of algorithm (4.5.1) concerns the sign of the right hand side of equation (4.5.2). We utilize a result obtained in [30].

Lemma 4.5.1 (Single Signed Update). *If $\hat{M}^n V^n$ is given by equation (3.1.8), with nonlinear coefficients determined by a local control problem of form (3.1.3), 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*

$$\textit{Short Position:} \qquad \left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k \geq 0 \tag{4.5.3}$$

$$\textit{Long Position:} \qquad \left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k \leq 0\ . \tag{4.5.4}$$

Proof. For the convenience of the reader, we summarize the proof in [30]. Writing out $\left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k$ in component form gives ($i < imax$)

$$\begin{aligned}\left[\left[\hat{M}^{k-1} - \hat{M}^k\right]\hat{V}^k\right]_i \;=\; & \left(\alpha_i^k \hat{V}_{i-1}^k + \beta_i^k \hat{V}_{i+1}^k - (\alpha_i^k + \beta_i^k + r\Delta\tau)\hat{V}_i^k\right) \\ & - \left(\alpha_i^{k-1} \hat{V}_{i-1}^k + \beta_i^{k-1} \hat{V}_{i+1}^k - (\alpha_i^{k-1} + \beta_i^{k-1} + r\Delta\tau)\hat{V}_i^k\right)\ . \end{aligned} \tag{4.5.5}$$

Consider a short position, so that, in terms of the local control problem (2.1.35), α_i^k, β_i^k are selected so that

$$\alpha_i^k \hat{V}_{i-1}^k + \beta_i^k \hat{V}_{i+1}^k - (\alpha_i^k + \beta_i^k + r\Delta\tau)\hat{V}_i^k \tag{4.5.6}$$

is maximized. Hence any other choice of coefficients, for example $\alpha_i^{k-1}, \beta_i^{k-1}$ cannot exceed the maximum produced by expression (4.5.6), thus

$$\begin{aligned} & \left(\alpha_i^k \hat{V}_{i-1}^k + \beta_i^k \hat{V}_{i+1}^k - (\alpha_i^k + \beta_i^k + r\Delta\tau)\hat{V}_i^k\right) \\ & - \left(\alpha_i^{k-1} \hat{V}_{i-1}^k + \beta_i^{k-1} \hat{V}_{i+1}^k - (\alpha_i^{k-1} + \beta_i^{k-1} + r\Delta\tau)\hat{V}_i^k\right) \\ & \geq 0\ , \end{aligned} \tag{4.5.7}$$

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 (4.5.4). ∎

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

Lemma 4.5.2 (M-matrix). *If the positive coefficient condition (4.1.1) is satisfied, $r \geq 0$, and boundary conditions (2.2.1, 2.2.3) 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 4.1.1, note that condition (4.1.1) implies that α_i^n, β_i^n in equation (3.1.8) are non-negative, hence $\left[I + (1-\theta)\hat{M}^{n+1}\right]$ has positive diagonals, non-positive offdiagonals, and is diagonally dominant, hence $\left[I + (1-\theta)\hat{M}^{n+1}\right]$ is an M-matrix. ∎

Remark 4.5.1 (Properties of M-matrices). *We remind the reader that an M-matrix Q has the important property that $Q^{-1} \geq 0$, and that $diag(Q^{-1}) > 0$.*

We can now state our main result concerning convergence of iteration (4.5.1).

Theorem 4.5.1 (Convergence of Iteration (4.5.1)). *Provided that the conditions required for Lemmas 4.5.1 and 4.5.2 are satisfied, then the nonlinear iteration (4.5.1) converges to the unique solution to equation (3.1.10), 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 4.5.1 and 4.5.2, the proof of this result is similar to the proof of convergence given in [31]. We give a brief outline of the steps in this proof, and refer readers to [31] for details. A straightforward maximum analysis of scheme (4.5.1) can be used to bound $\|\hat{V}^k\|_\infty$ independent of iteration k. Recall equation (4.5.2)

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

From Lemma 4.5.1, we have that the right hand side of equation (4.5.2) 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 4.5.2), 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 (4.5.1) can be regarded as a non-smooth Newton iteration. Since the non-smooth algebraic nonlinear equations (3.1.10) are strongly semi-smooth [33], convergence is quadratic in a sufficiently small neighborhood of the solution [32]. ∎

4.6 Arbitrage Inequality

It is interesting to verify that the discrete equations satisfy discrete arbitrage inequalities [12, 13], 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 (4.2.1), then if $V^0 > W^0$, and $V^k_{imax} > W^k_{imax}, (k = 0, ..., n)$, then $V^n > W^n$.

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

Theorem 4.6.1 (Discrete Comparison Principle). *The fully implicit discretization (3.1.10) satisfies a discrete comparison principle, that is, if $V^n > W^n$, $D_V^{n+1} > D_W^{n+1}$, and V^{n+1}, W^{n+1} satisfy equation (3.1.10), and the conditions for Lemmas 4.5.1 and 4.5.2 are satisfied, then $V^{n+1} > W^{n+1}$.*

Proof. V, W satisfy

$$\begin{aligned} \left[I+\hat{M}(V^{n+1})\right]V^{n+1} &= V^n + I^*(D_V^{n+1} - V^n) \\ \left[I+\hat{M}(W^{n+1})\right]W^{n+1} &= W^n + I^*(D_W^{n+1} - W^n)\ . \end{aligned} \tag{4.6.1}$$

Some manipulation of equation (4.6.1) gives

$$\begin{aligned}\left[I+\hat{M}(W^{n+1})\right](V^{n+1}-W^{n+1}) &= (I-I^*)(V^n-W^n)+\left[\hat{M}(W^{n+1})-\hat{M}(V^{n+1})\right]V^{n+1} \\ &\quad +I^*(D_V^{n+1}-D_W^{n+1}) \qquad (4.6.2)\\ \left[I+\hat{M}(V^{n+1})\right](V^{n+1}-W^{n+1}) &= (I-I^*)(V^n-W^n)-\left[\hat{M}(V^{n+1})-\hat{M}(W^{n+1})\right]W^{n+1} \\ &\quad +I^*(D_V^{n+1}-D_W^{n+1})\,. \qquad (4.6.3)\end{aligned}$$

Consider a short position. From Lemma 4.5.1 (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 4.5.2, and equation (4.6.2)

$$\begin{aligned}\left[I+\hat{M}(W^{n+1})\right]^{-1} \quad [\quad & (I-I^*)(V^n-W^n)+I^*(D_V^{n+1}-D_W^{n+1}) \\ & +\left(\hat{M}(W^{n+1})-\hat{M}(V^{n+1})\right)V^{n+1}]>0\,, \qquad (4.6.4)\end{aligned}$$

hence $V^{n+1}>W^{n+1}$. For a long position, a similar argument using Lemmas 4.5.1 and 4.5.2 and equation (4.6.3) gives the same result. ∎

Remark 4.6.1 (Use of Lemma 4.5.1). *Note that a key property in the above proof is Lemma 4.5.1. This result follows quite generally if we ensure that we solve a discrete version of the control problem (2.1.35), 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.*

Chapter 5

Positive Coefficient Grid Condition

In this chapter we develop an algorithm which ensures that grid condition (4.1.1) can be satisfied by insertion of a finite number of nodes in any arbitrary initial grid.

5.1 Node Insertion Algorithm

Some algebra shows that condition (4.1.1) is satisfied by at least one of forward or backward differencing at node i if

$$\sigma^2 S_i + (S_{i+1} - S_{i-1})(|r'| - \lambda\sigma\sqrt{1-\rho^2}) \quad \geq \quad 0\,. \tag{5.1.1}$$

Equation (5.1.1) is always satisfied if $(|r'| - \lambda\sigma\sqrt{1-\rho^2}) \geq 0$. Consequently, we will examine the case when $(|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0$. Suppose $S_{i+1} - S_i = \Delta S, \forall i$, and $S_i = i\Delta S$, then condition (5.1.1) reduces to

$$\sigma^2 i + 2(|r'| - \lambda\sigma\sqrt{1-\rho^2}) \quad \geq \quad 0\,. \tag{5.1.2}$$

Clearly, for sufficiently large S_i, condition (5.1.2) can be satisfied, since $\sigma^2 > 0$.

Equation (5.1.2) simplifies at $i = 1$ to

$$\sigma^2 + 2(|r'| - \lambda\sigma\sqrt{1-\rho^2}) \;\geq\; 0\,. \tag{5.1.3}$$

Condition (5.1.2) may not be satisfied, no matter how small ΔS is chosen. From equation (5.1.3), 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 (5.1.2) becomes

$$\sigma^2 S_0 + \Delta S(\sigma^2 i + 2|r'|) - 2\Delta S\lambda\sigma\sqrt{1-\rho^2} \;\geq\; 0\,, \tag{5.1.4}$$

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

$$h \;=\; \max_i (S_{i+1} - S_i) \tag{5.1.5}$$

then condition (5.1.1) is always satisfied if

$$h \leq \frac{\sigma^2 S_0^2}{2||r'| - \lambda\sigma\sqrt{1-\rho^2}|}\,. \tag{5.1.6}$$

Note that a grid constructed by enforcing condition (5.1.6) is not required in practice (as we shall see below). Condition (5.1.6) 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 (4.1.1) 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 (4.1.1) is satisfied. For a given grid with $S_0 > 0$, we will apply the boundary condition (2.2.1) 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 (2.2.1) 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 $((|r'| - \lambda\sigma\sqrt{1-\rho^2}) \geq 0)$ Then

Return // *Original grid satisfies condition*

Endif

If $([S_0 = 0]$ and $[\sigma^2 S_1 + min(S_2, 2S_1)(|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0])$ Then

Exit // *Algorithm fails, need* $S_0 > 0$

Endif

$i = 1$

While (S_i is not the largest node)

If $(\sigma^2 S_i + (S_{i+1} - S_{i-1})(|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0)$ Then

If $(\sigma^2 S_i + 2(S_i - S_{i-1})(|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0)$ Then (5.1.7)

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 $S_0 \neq 0$, the algorithm (5.1.7) is guaranteed to produce a fine grid such that condition (5.1.1) holds for all nodes. From equation (5.1.6), the total number of nodes inserted must be finite.

If $S_0 = 0$ and $\sigma^2 S_1 + min(S_2, 2S_1)(|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0$, then a new grid satisfying condition (5.1.1) does not exist. Consequently, in the case that $\sigma^2 + (|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0$, we must have $S_0 > 0$ in order for algorithm (5.1.7) to succeed. In this case, we can set S_0 to be a small number, and apply boundary condition (2.2.1) at S_0. We will verify that this does not cause any significant error at asset values of interest through some numerical experiments that will be reported in subsequent sections. Algorithm 5.1.7 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

$$\begin{aligned} & p_0 \leq \frac{S_{i+1} - S_i}{S_i - S_{i-1}} \leq q_0 \ \ ; \ \ i = 1, ..., n-1 \\ & q_0 \geq p_0 > 0, \end{aligned} \tag{5.1.8}$$

the following result holds.

Theorem 5.1.1 (Grid Aspect Ratio after Application of Algorithm (5.1.7)). *Given an initial grid with n nodes and p_0, q_0 given by equation (5.1.8), then after application of algorithm (5.1.7) with $S_0 > 0$, the new grid (with m nodes, $m \geq n+1$) satisfies*

$$\begin{aligned} p \leq \frac{S_{i+1} - S_i}{S_i - S_{i-1}} \leq q, \qquad & \textit{where } p = min(\frac{1}{3}, p_0) \textit{ and } q = max(5, 2q_0) \\ & 1 \leq i \leq m-1. \end{aligned} \tag{5.1.9}$$

Proof. See Appendix C.

Note that algorithm (5.1.7) 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 (5.1.7).
- Each node i of the new grid is processed, and the discretization coefficients α_i, β_i are constructed (equation(3.1.2)). First, central differencing is tested. If the positive coefficient condition (4.1.1) 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 5.1.7). 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.

Chapter 6

PDE Examples

In this chapter we give 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 5.1.7 is applied. We show that the solution is insensitive to small positive S_0.

6.1 Timestep Selection

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

$$\Delta\tau^{n+2} = \left[min_i \left(\frac{dnorm}{\frac{|V(S_i,\tau^n+\Delta\tau^{n+1})-V(S_i,\tau^n)|}{max(D,|V(V(S_i,\tau^n+\Delta\tau^{n+1})|,|V(S_i,\tau^n)|)}} \right) \right] \Delta\tau^{n+1}, \tag{6.1.1}$$

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).

6.2 Market Parameter Interpolation

Recall from equation (2.1.9) that the drift term in our PDE is,

$$r' = \mu - (\mu' - r)\frac{\sigma\rho}{\sigma'}, \tag{6.2.1}$$

hence,

$$\frac{\mu - r'}{\sigma} = \frac{\rho(\mu' - r)}{\sigma'}. \tag{6.2.2}$$

When $|\rho| = 1$, then the drift term r' must equal the risk free interest rate r [15]. That is, $r' = r$ when $\rho = \pm 1$. Therefore, r, ρ, μ, σ, μ' and σ' cannot be determined independently. We arbitrarily choose μ' as the dependent variable. From equation (6.2.2), we see that if $\rho = 1$, and $r' = r$, we obtain

$$\mu' = r + (\mu - r)\frac{\sigma'}{\sigma}, \tag{6.2.3}$$

and if $\rho = -1$ ($r' = r$), we obtain

$$\mu' = r - (\mu - r)\frac{\sigma'}{\sigma}. \tag{6.2.4}$$

This suggests that we could interpolate μ' as

$$\mu' = r + f(\rho)(\mu - r)\frac{\sigma'}{\sigma}, \tag{6.2.5}$$

where $f(\rho)$ has the properties that $f(-1) = -1, f(1) = 1$. In our numerical examples we choose

$$f(\rho) = \rho. \tag{6.2.6}$$

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

$$\mu' = r + (\mu - r)\frac{\sigma'\rho}{\sigma}\ . \tag{6.2.7}$$

Substituting equation (6.2.5) into equation (6.2.1), gives

$$r' = (1 - \rho f(\rho))\mu + r\rho f(\rho)\ . \tag{6.2.8}$$

Assuming equation (6.2.6) holds, then we obtain

$$r' = (1-\rho^2)\mu + \rho^2 r \; . \tag{6.2.9}$$

6.3 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 data is given in Table 6.1. In this example, we will assume a European straddle, with payoff

$$V(S,\tau=0) \;\; = \;\; \max(K-S,0) + \max(S-K,0) \; . \tag{6.3.1}$$

Since the derivative (V_S) of the payoff changes sign, then the PDE is truly nonlinear. The tolerance in algorithm 4.5.1 is set to 10^{-6}.

Using data in Table 6.1, Table 6.2 shows the convergence results for fully implicit timestepping. Table 6.3 shows the results obtained with Crank-Nicolson timestepping, where we have used the modification suggested in [34]. We use variable timestepping as given in equation (6.1.1). As expected, fully implicit timestepping gives first order convergence and Crank-Nicolson method gives quadratic convergence. Recall from Theorem 4.4.1 that convergence to the viscosity solution is only guaranteed for fully implicit timestepping. In this case, Crank-Nicolson also converges to the viscosity solution. From algorithm 4.5.1, we can see the minimum number of iterations per timestep is two. In Table 6.3, 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.

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	100
Payoff	straddle
T	1 year

TABLE 6.1: *Data used in the straddle option examples. This data gives* $|r'| - \lambda\sigma\sqrt{1-\rho^2} = 0.03636 > 0$*, so no new node is inserted into the asset grid when algorithm (5.1.7) is applied.*

Nodes	dnorm	Timesteps	Nonlinear iterations	Option value	Change	Ratio
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

TABLE 6.2: *Convergence study with fully implicit timestepping, variable timesteps (equation (6.1.1)), data in Table 6.1 used. No new nodes are inserted into the asset grid. Straddle payoff (6.3.1), short position.*

Nodes	dnorm	Timesteps	Nonlinear iterations	Option value	Change	Ratio
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 6.3: *Convergence study with Crank-Nicolson timestepping, Rannacher smoothing [34], variable timestepping, data in Table 6.1 used. No new nodes are inserted into the asset grid. Straddle payoff (6.3.1), short position.*

6.4 Positive S_0 Tests

In Chapter 5, we argued that when $\sigma^2 + (|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0$, we must have $S_0 > 0$ (the minimum value for the asset grid) in order for algorithm 5.1.7 to succeed. Table 6.4 shows data which requires $S_0 > 0$ to ensure that algorithm 5.1.7 completes successfully.

Table 6.5 shows the option prices, deltas and gammas under different S_0 values, for various asset price values. We see that, as the 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 6.4 was used for this test. Observe that this data requires $S_0 > 0$ for algorithm 5.1.7 to succeed.

Table 6.6 gives a convergence study using data in Table 6.4. 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, and the hedger is very risk averse. These parameter values make the option prices extremely high as shown in Table 6.5 and 6.6. From the data in Table 6.5 and 6.6,

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	100
Payoff	Straddle
T	1 year

TABLE 6.4: *Data used for positive S_0 tests. This data gives $\sigma^2 + (|r'| - \lambda\sigma\sqrt{1-\rho^2}) = -0.0181 < 0$, In this case, S_0 has to be positive, in order for algorithm (5.1.7) to succeed. When algorithm (5.1.7) is applied, new nodes may be inserted into the asset grid.*

we can conclude that small positive S_0 has little effect on the solution.

6.5 Stress Test for the Node Insertion Algorithm

In Chapter 5, we have shown that when $S_0 > 0$, algorithm (5.1.7) guarantees construction of a new grid, which satisfies condition (5.1.1) and preserves the grid aspect ratio. If algorithm (5.1.7) adds many nodes to typical initial grids, it is a poor insertion algorithm. Algorithm (5.1.7) has the property that for most original grids, with normal market parameters, only a few (sometimes zero) nodes are inserted. Usually, these nodes are near $S = 0$. However, bad cases do exist. Before giving a bad case example, we analyze condition (5.1.1) to determine the conditions

Asset Price	S_0	Option Price	Delta	Gamma
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	85.393	-0.57618	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	79.736	-0.531765	0.005933

TABLE 6.5: *The effect of positive S_0 on the solution, Crank-Nicolson method, variable timesteps, using data in Table 6.4. Straddle payoff (6.3.1), 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 $S_0 = 2$ and $S_0 = 5$.*

under which it is necessary to insert many nodes. A new insertion is required when condition (5.1.1) fails, which means

$$\sigma^2 S_i + (S_{i+1} - S_{i-1})(|r'| - \lambda\sigma\sqrt{1-\rho^2}) \quad < \quad 0\,, \tag{6.5.1}$$

hence,

$$\frac{S_i}{S_{i+1} - S_{i-1}} < -\frac{|r'| - \lambda\sigma\sqrt{1-\rho^2}}{\sigma^2}\,. \tag{6.5.2}$$

Therefore, S_i should be small to cause node insertion. For sufficiently large S_i, if the original grid has a reasonable aspect ratio (the case in practice), $\frac{S_i}{S_{i+1}-S_{i-1}}$ will be large, and thus equation (6.5.2) will not be satisfied. Hence, most node insertions occur as $S_0 \to 0$.

Substituting equation (6.2.9) into equation (6.5.1), we obtain

$$\sigma^2 S_i + (S_{i+1} - S_{i-1})((1-\rho^2)\mu + \rho^2 r - \lambda\sigma\sqrt{1-\rho^2}) \quad < \quad 0\,. \tag{6.5.3}$$

S_0	Nodes	dnorm	Option value ($S = 100$)	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 6.6: *A case where S_0 has to be positive, in order for node insertion algorithm to succeed. Crank-Nicolson timestepping, Rannacher smoothing [34], variable timesteps. Straddle payoff (6.3.1). The data in Table 6.4 gives $\sigma^2 + (|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0$, so that S_0 has to be positive. When algorithm (5.1.7) 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 nodes.*

Thus, to insert many nodes, we have to set $\sigma^2 + (1-\rho^2)\mu + \rho^2 r - \lambda\sigma\sqrt{1-\rho^2}$ to a large negative value (the absolute value of it is large). As a result, we assign small values to ρ, μ and r, and assign a large value to λ. Table 6.8 shows the node insertion results using data in Table 6.7. We see that many nodes are inserted when S_0 is small. However, as S_0 increases, the number of insertions decreases rapidly. We have already shown that small positive S_0 has little effect on the solution, so we can avoid problems by assigning S_0 a reasonable (e.g. not too small). In order to construct this bad case, we assign unreasonable values to market parameters as in Table 6.7. Normal market parameters do not lead to a poor node insertion case. Finally, we can conclude that algorithm (5.1.7) works very well in practice.

r	0.005
ρ	0
σ	0.5
μ	0.01
σ'	0.25
$\mu' = r + (\mu - r)\frac{\sigma'\rho}{\sigma}$	0.005
λ	0.8
$r' = \mu - (\mu' - r)\frac{\sigma\rho}{\sigma'}$	0.01

TABLE 6.7: *Data used for poor cases for node insertion algorithm (5.1.7)*

S_0	Old Grid Size ($S_{max} = 5000$)	New Grid Size	Insertions Around S_0
0.001	51	101	38
0.01	51	92	29
0.1	51	82	19
1	51	71	8
5	51	65	2

TABLE 6.8: *Poor cases for node insertion algorithm (5.1.7). In the original grid, the first three nodes are S_0, 10 and 20. Grid is constructed assuming prices of interest are near $S = 100$.*

6.6 Long and Short Positions

From equation (2.1.30), it is clear that the option price of a short position should always be higher than the option price of a long position. Figure 6.1 illustrates this fact, for the straddle payoff (6.3.1).

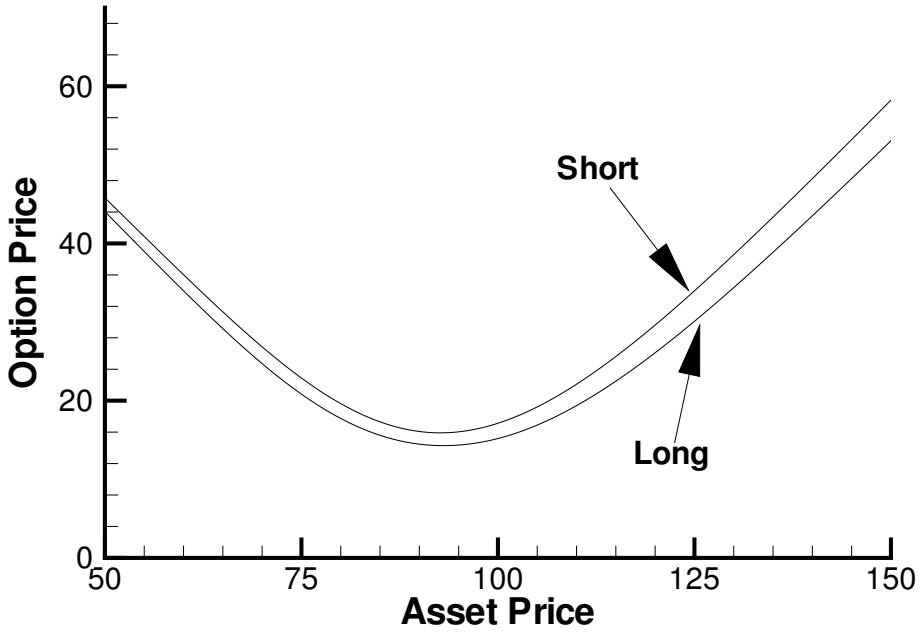


FIGURE 6.1: *The option price of a long position and a short position, straddle payoff, data in Table 6.1.*

Chapter 7

Hedging Simulations

In this chapter we will use a Monte Carlo method to simulate the hedging process. We will illustrate the results by showing histograms of the hedging portfolio (the profit and loss *P&L*) at the expiry time. We also study the mean, *VaR*, *CVaR* and standard deviation of *P&L*. *VaR*, which is know as *Value at Risk*, is the worst case loss with a given probability. *CVaR* is the mean of the worst case losses.

7.1 Algorithm Description

We make a slight change from the description of the hedging portfolio as given in equation (2.1.4). 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 described in Chapter 2, consider the case where we wish to hedge a short position in the claim with value $V = V(S,t)$. Then, a portfolio $\mathcal{P}^i$ at time $t_i = i\Delta t$ has three components:

• A short claim position worth V^i;

• Long x^i shares of asset H, where x^i is the number of units of H held in the portfolio;

• A risk-free bank account $\mathcal{B}^i$.

Hence,

$$\mathcal{P}^i = -V^i + x^i H^i + \mathcal{B}^i. \tag{7.1.1}$$

In contrast to the hedging portfolio Π in Chapter 2, we do not inject any cash into this portfolio $\mathcal{P}$ in order to ensure that $\mathcal{P} = 0$ after the initial time. In the case $\lambda = 0$, this portfolio is then self-financing (on $[0,T)$, where T is option expiry time), but will in general not meet the obligations of the contingent claim exactly at the option expiry. Note that PDE (2.1.30) does not contain B (the risk free bank account), so that use of x^i given by equation (2.1.7) minimizes the local risk, regardless the amount in B. We have denoted the bank account in the portfolio $\mathcal{P}$ by $\mathcal{B}$ to distinguish it from the bank account in equation (2.1.4). In this case, $\mathcal{P}$ will not necessarily be zero after the initial time, since we will not inject cash into this portfolio.

As discussed in [26], for the case $\lambda = 0$, this strategy is self-financing on $[0,T)$, with a single payment at time T. Moller [26] then points out that the 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 show the results of this strategy since it is easy to interpret the resulting *P&L* diagrams. These diagrams will show the distribution of the future value of the incremental profit/loss of the hedge portfolio.

Given the option price at $t = 0$, which is given from the solution of the PDEs (2.1.30), the initial portfolio is given by

$$\mathcal{P}^0 \;=\; -V^0 + x^0 H^0 + \mathcal{B}^0 \;. \tag{7.1.2}$$

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

Let,

$$V_S^i = \frac{\partial V}{\partial S}(S^i, t_i) \tag{7.1.3}$$

According to equation (2.1.7), to minimize the local variance, we choose x^i at time t_i to be

$$x^i = \left(\frac{\sigma S^i \rho}{\sigma' H^i}\right) V_S^i, \tag{7.1.4}$$

Let ϕ_H, ϕ_S be random draws from a normal distribution with mean zero and unit variance. The prices and correlation of S and H at time t_{i+1} are given by

$$\begin{aligned} S^{i+1} &= S^i \exp\left[(\mu - \sigma^2/2)\,\Delta t + \sigma\phi_S\sqrt{\Delta t})\right] \\ H^{i+1} &= H^i \exp\left[(\mu' - (\sigma')^2/2)\,\Delta t + \sigma'\phi_H\sqrt{\Delta t})\right] \\ E^P(\phi_S\ \phi_H) &= \rho, \end{aligned} \tag{7.1.5}$$

where $E^P[\cdot]$ is the expectation operator. Initially, we solve equation (2.1.35) numerically backward in time, from $t = T$ to $t = 0$. At each timestep, the option prices 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 (7.1.6).

Hedging Algorithm

$\mathcal{P}^0 = 0$

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

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

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

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 (7.1.5)

Interpolate V_S^i from the stored tables

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

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

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

Endfor

$\mathcal{P}(T) = -V(T) + x(T)H(T) + \mathcal{B}(T)$ (7.1.6)

Recall that

$$\begin{aligned} &\text{Short Position:} \quad d\Pi = \lambda\sigma\sqrt{1-\rho^2}S|V_S|\,dt + SV_S\sqrt{1-\rho^2}\sigma\,dX \\ &\text{Long Position:} \quad d\Pi = \lambda\sigma\sqrt{1-\rho^2}S|V_S|\,dt - SV_S\sqrt{1-\rho^2}\sigma\,dX \end{aligned} \tag{7.1.7}$$

Considering only the short position, we have that

$$d\Pi \;=\; \lambda\sigma\sqrt{1-\rho^2}S|V_S|\,dt + SV_S\sqrt{1-\rho^2}\sigma\,dX\;. \tag{7.1.8}$$

We will show histograms of $\mathcal{P}(T)$, i.e. the future profit and loss. Since the cash shortfall is only realized at the expiry time in the portfolio $\mathcal{P}$, then the final value of $\mathcal{P}$ can be determined in

terms of the solution V by considering the future value of $d\Pi$ at each instant, so that

$$\mathcal{P}(T) = \lambda \int_0^T e^{r(T-t)} \sigma \sqrt{1-\rho^2} S |V_S| \, dt + \int_0^T e^{r(T-t)} S V_S \sqrt{1-\rho^2} \sigma \, dX \, , \quad (7.1.9)$$

which means that

$$E^P[\mathcal{P}(T)] = E^P \left[\lambda \int_0^T e^{r(T-t)} \sigma \sqrt{1-\rho^2} S |V_S| \, dt \right] , \quad (7.1.10)$$

in the limit as the rebalancing interval tends to zero.

7.2 Hedging Parameters

Hedging simulations are carried out using the Monte Carlo parameters described in Table 7.1. There are many parameters which affect the hedging results, but we are particularly interested in λ and ρ. λ is known as the safety loading parameter, and could be determined on the basis of the firm's beta, or the underlying asset correlation with the firm's existing investment portfolio [22].

When $\lambda = 0$, from equation (7.1.10), $E^P[\mathcal{P}(T)] = 0$. According to equation (2.1.21), increasing λ implies a greater reward for bearing the unhedgeable risk, hence the mean *P&L* ($P\&L = E^P[\mathcal{P}(T)]$) should also increase (when $|\rho| \neq 1$). Table 7.2 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 (the mean *P&L* 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 7.3 shows the case where λ increases and the other parameters are held constant. As λ increases (we require greater reward), the mean, *VaR* and *CVaR* of *P&L* increase, while standard deviation is nearly constant. Those results are also shown in Figure 7.1 (a), (c) and (d).

As stated in Chapter 2, ρ ($0 \leq |\rho| \leq 1$) is the correlation between dW and dZ. When $|\rho| = 1$, hedging with asset H is a perfect hedge, and equation (2.1.30) reverts back to usual Black-Scholes equation. In this case, the hedging simulation should be the same as the standard discrete

Hedging interval	2 days
Number of Simulations	1,000,000

TABLE 7.1: *Data used for Monte Carlo hedging simulations.*

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

TABLE 7.2: *Hedging simulation results ($\mathcal{P}(T)$) with $\lambda = 0$, using data in Table 6.1 and 7.1. Straddle, short position.*

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

TABLE 7.3: *Hedging simulation results ($\mathcal{P}(T)$) with λ varying, and the other parameters held constant. Data in Table 6.1 and 7.1 is used. Straddle, short position.*

delta hedging, and thus the mean of the *P&L* should be zero. Some results for the case $|\rho| = 1$ are given in Table 7.4. Note that the standard deviation is not identically zero in this case due to the finite (two day) rebalancing interval.

Table 7.5 shows the results obtained when ρ is increased. When $|\rho|$ increases, the hedging result becomes closer to the result given by standard delta hedging. The mean shifts closer to zero (the mean decreases, since we take less risk), and the standard deviation of the *P&L* decreases. These results are also shown in Figure 7.1 (b), (e) and (f).

ρ	λ	Mean	VaR (95%)	CVaR (95%)	Stndrd Dev.	$V(S=100,\tau=0)$
1	0.5	-0.001	-1.9365	-2.7062	1.1752	16.0237
-1	0.5	-0.001	-2.2683	-3.0510	1.3583	16.0237

TABLE 7.4: *Hedging simulations with* $|\rho| = 1$*, using data in Table 6.1 and 7.1. Straddle, short position. Note that the standard deviation of the P&L is nonzero due to the finite rebalancing interval (Table 7.1.)*

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

TABLE 7.5: *Hedging simulations with variable* ρ*, other parameters held constant. Data in Table 6.1 and 7.1 is used. Straddle, short position.*

7.3 Convergence of Monte Carlo Hedging

Table 7.6 shows the convergence of mean and variance of the portfolio at time T, as the number of simulations becomes large, using data in Table 6.1 and Table 7.1. The mean and variance are insensitive to the change of the number of simulations, when the number of simulations exceeds 100,000. Therefore, we set the number of simulations to 1,000,000 in the following numerical tests.

As the hedging interval goes to zero and the number of simulations goes to infinity, the variance of the portfolio at time T converges to a positive value due to the unhedgeable residual risk. Table 7.7 shows a numerical example of the convergence of the standard deviation.

# of Simulations	Mean	VaR (95%)	CVaR (95%)	Stndrd Dev.
5000	1.0995	-9.5732	-14.2440	6.2630
20000	1.0308	-9.8677	-14.4589	6.3284
80000	1.0725	-9.7672	-14.3726	6.2729
160000	1.0508	-9.8391	-14.4611	6.3254
640000	1.0493	-9.8383	-14.4437	6.3162
1000000	1.0499	-9.8403	-14.4212	6.3072

TABLE 7.6: *Hedging simulation results ($\mathcal{P}(T)$) with number of simulations varying, and the other parameters held constant. Data in Table 6.1 and 7.1 is used. Straddle, short position.*

Hedging interval (days)	Mean	$VaR(95\%)$	$CVaR(95\%)$	Standrd 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 7.7: *Convergence of standard deviation, as the hedging interval is decreased. Data in Table 6.1 is used. 1,000,000 simulations.*

7.4 An American Example

The price of an American claim is given by equation (2.1.37). We can generalize the numerical methods described in this work to the American case using the penalty method described in [21, 16]. 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 data in Table

6.1. Table 7.8 shows the prices for a long/short American/European straddle.

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

TABLE 7.8: *Straddle example, data in Table 6.1. Results are correct to the number of digits shown.*

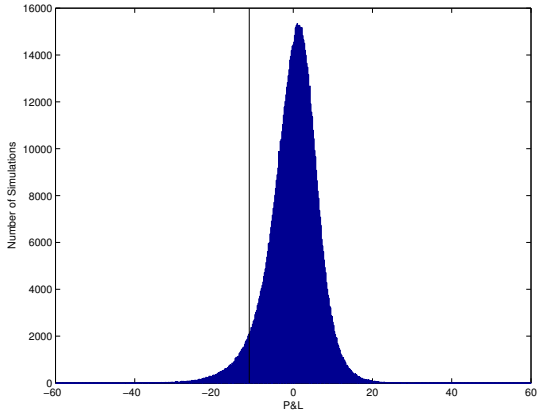
7.5 Nonlinearity and Reinsurance

Suppose there are two firms, A and B, and a reinsurer C. A,B,C all price short positions using the data in Table 6.1. In particular, A,B,C all have the same estimates for drift rates and safety loading parameter.

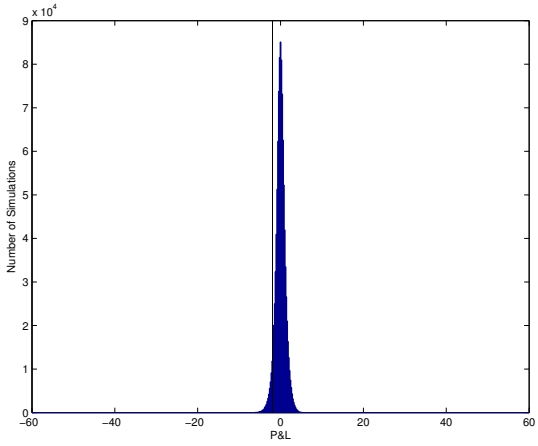
Suppose A needs to hedge a short call, and B needs to hedge a short put. A and B can hedge these positions, or purchase reinsurance from C. C would then have a short straddle position. The prices of individually hedging a call, put and a straddle are given in Table 7.9, computed using PDE (2.1.30), and data in Table 6.1. If A and B individually hedged 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 hedged 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. These effects become even more interesting if different firms have different market prices of risk.

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

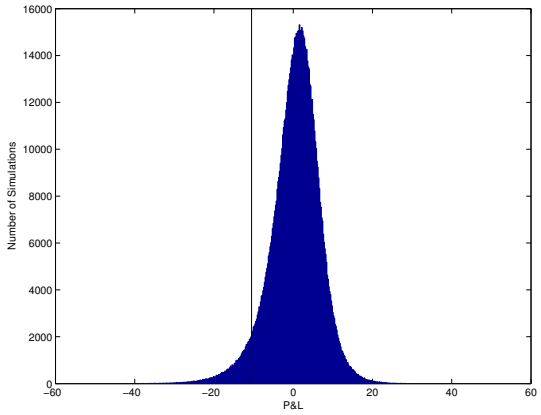
TABLE 7.9: *Call, put and straddle prices, data in Table 6.1. Results are correct to the number of digits shown. Note that the payoff of the straddle is the sum of the call and put payoffs.*



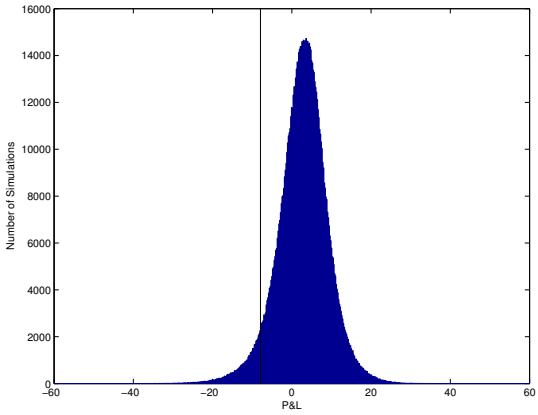
(a) $\lambda = 0$, $\rho = 0.9$



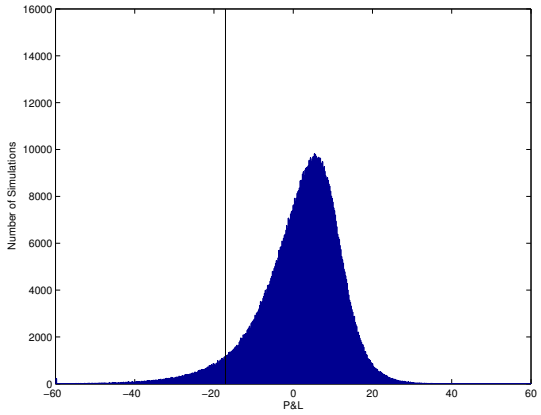
(b) $\lambda = 0.5$, $\rho = 1$



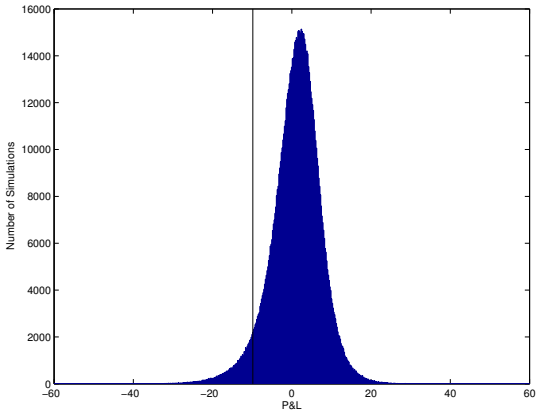
(c) $\lambda = 0.1$, $\rho = 0.9$



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



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



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

FIGURE 7.1: *Histograms of Monte Carlo hedging simulations of a short position for a straddle, using data in Table 6.1 and 7.1. Note that the vertical scale for Figure (b) is different from the vertical scales in the other figures. The vertical line in each figure represents the* 95% *VaR P&L.*

Chapter 8

Conclusions

In this thesis, we have considered the situation where an insurance firm cannot hedge directly with the underlying of a contingent claim.

At each infinitesmal time, the best local hedge is constructed. However, since the hedge is not perfect, the insurance company requires an extra return to compensate for this residual risk. The risk preferences of the insurance company enters into the valuation through a saftey loading parameter, or local Sharpe ratio. For non-zero safety loading parameter, the PDE is nonlinear, with a different price for long or short positions.

The methods used in this thesis are easily extended to American claims. We have included an example of pricing an American option, but the details of the algorithm will be discussed in future work.

An implicit discretization method is developed for the nonlinear pricing PDE. This scheme 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.

Monte Carlo hedging experiments are given which demonstrate the use of the hedging strategy which follows from the numerical solution of the pricing PDE. These examples clearly show that the unhedgeable risk is compensated by a reserve which is built up as time proceeds.

This approach to option pricing in the situation where the option seller hedges with an asset imperfectly correlated with the underlying has the following advantages.

- The holder of the hedging portfolio is rewarded for the unhedgeable risk by specifying a safety loading parameter, or Sharpe ratio. This is a familiar and easily explained parameter.
- The prices are linear in terms of the number of units bought/sold. This is in contrast to many other approaches.
- Long/short prices are different. This is easily explained in terms of risk/reward for long/short positions.
- The pricing PDE is nonlinear, and the market price of risk may be different for reinsurers and primary insurers. This makes reinsurance a risk and cost-reducing strategy under some circumstances.
- The nonlinear pricing PDE is easily solved using an implicit method. Existing PDE option pricing software can be modified in straightforward fashion to price options using this model, simply by adding an updating step to the American pricing iteration.

8.1 Future Work

There could be several extensions for future research:

- It would be interesting to develop techniques which guarantee quadratic convergence of nonlinear option pricing PDEs to the viscosity solution as the timestep and mesh spacing are reduced, or to prove that quadratic convergence cannot always be guaranteed. However, such a proof would be difficult to obtain.

- It would be desirable to prove convergence of iterative schemes without the use of M-matrices. For both single factor options and two factor options, the M-matrix property may not be preserved. Some advances in the basic theory of viscosity solutions would be required.

- The model could be extended to include uncertain parameters, for example, uncertain volatility.

Appendix A

Discrete Equation Coefficients

The detailed form of the discrete equation coefficients used in equation (3.1.3) are given here.

In the case of a central discretization

$$\begin{aligned} \alpha^n_{i,cent} &= \alpha'_{i,cent} - \gamma_{i,cent}\, q^n_{i,cent} \\ \beta^n_{i,central} &= \beta'_{i,cent} + \gamma_{i,cent}\, q^n_{i,cent}\ , \end{aligned} \tag{A.0.1}$$

where

$$q^n_{i,cent} = \begin{cases} \mathrm{sgn}(V_S)^n_{i,cent} & \text{if short} \\ -\mathrm{sgn}(V_S)^n_{i,cent} & \text{if long} \end{cases} . \tag{A.0.2}$$

and

$$\begin{aligned} \gamma_{i,cent} &= \left(\frac{S_i \lambda \sigma \sqrt{1-\rho^2}}{S_{i+1} - S_{i-1}} \right) \Delta\tau \\ (V_S)^n_{i,cent} &= \frac{V^n_{i+1} - V^n_{i-1}}{S_{i+1} - S_{i-1}} \end{aligned} \tag{A.0.3}$$

and

$$\alpha'_{i,cent} = \left[\frac{\sigma^2 S_i^2}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} - \frac{r' S_i}{S_{i+1} - S_{i-1}} \right] \Delta\tau$$
$$\beta'_{i,cent} = \left[\frac{\sigma^2 S_i^2}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \frac{r' S_i}{S_{i+1} - S_{i-1}} \right] \Delta\tau. \tag{A.0.4}$$

Note that the above definitions ensure that we are solving a discrete version of the local control problem (2.1.36).

In the case of forward differencing, we obtain

$$\alpha^n_{i,for} = \alpha'_{i,for}$$
$$\beta^n_{i,for} = \beta'_{i,for} + \gamma_{i,for}\, q^n_{i,for} \tag{A.0.5}$$

where

$$q^n_{i,for} = \begin{cases} \text{sgn}(V_S)^n_{i,for} & \text{if short} \\ -\text{sgn}(V_S)^n_{i,for} & \text{if long} \end{cases}, \tag{A.0.6}$$

and

$$\gamma_{i,for} = \left(\frac{S_i \lambda \sigma \sqrt{1-\rho^2}}{S_{i+1} - S_i} \right) \Delta\tau$$
$$(V_S)^n_{i,for} = \frac{V^n_{i+1} - V^n_i}{S_{i+1} - S_i} \tag{A.0.7}$$

and

$$\alpha'_{i,for} = \frac{\sigma^2 S_i^2}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} \Delta\tau$$
$$\beta'_{i,for} = \left[\frac{\sigma^2 S_i^2}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \frac{r' S_i}{S_{i+1} - S_i} \right] \Delta\tau. \tag{A.0.8}$$

Again, note that we have used definition (A.0.6), so that we solve a discrete version of the local control problem (2.1.36).

In the case of backward differencing we have

$$\begin{aligned} \alpha^n_{i,back} &= \alpha'_{i,back} - \gamma_{i,back}\, q^n_{i,back} \\ \beta^n_{i,back} &= \beta'_{i,back} \end{aligned} \tag{A.0.9}$$

where

$$q^n_{i,back} = \begin{cases} \mathrm{sgn}(V_S)^n_{i,back} & \text{if short} \\ -\mathrm{sgn}(V_S)^n_{i,back} & \text{if long} \end{cases}, \tag{A.0.10}$$

and

$$\begin{aligned} \gamma_{i,back} &= \left(\frac{S_i \lambda \sigma \sqrt{1-\rho^2}}{S_i - S_{i-1}} \right) \Delta\tau \\ (V_S)^n_{i,back} &= \frac{V^n_i - V^n_{i-1}}{S_i - S_{i-1}} \end{aligned} \tag{A.0.11}$$

and

$$\begin{aligned} \alpha'_{i,back} &= \left[\frac{\sigma^2 S_i^2}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} - \frac{r' S_i}{S_i - S_{i-1}} \right] \Delta\tau \\ \beta'_{i,back} &= \left[\frac{\sigma^2 S_i^2}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} \right] \Delta\tau. \end{aligned} \tag{A.0.12}$$

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

$$\alpha^n_i = \begin{cases} \alpha^n_{i,cent} & \text{if central differencing} \\ \alpha^n_{i,for} & \text{if forward differencing} \\ \alpha^n_{i,back} & \text{if backward differencing} , \end{cases} \tag{A.0.13}$$

$$\beta^n_i = \begin{cases} \beta^n_{i,cent} & \text{if central differencing} \\ \beta^n_{i,for} & \text{if forward differencing} \\ \beta^n_{i,back} & \text{if backward differencing} , \end{cases} \tag{A.0.14}$$

$$\alpha'_i = \begin{cases} \alpha'_{i,cent} & \text{if central differencing} \\ \alpha'_{i,for} & \text{if forward differencing} \\ \alpha'_{i,back} & \text{if backward differencing ,} \end{cases} \tag{A.0.15}$$

$$\beta'_i = \begin{cases} \beta'_{i,cent} & \text{if central differencing} \\ \beta'_{i,for} & \text{if forward differencing} \\ \beta'_{i,back} & \text{if backward differencing .} \end{cases} \tag{A.0.16}$$

We also define

$$\gamma'_{i,cent} = \begin{cases} \gamma_{i,cent} & \text{if central differencing} \\ 0 & \text{otherwise ,} \end{cases} \tag{A.0.17}$$

$$\gamma'_{i,for} = \begin{cases} \gamma_{i,for} & \text{if forward differencing} \\ 0 & \text{otherwise ,} \end{cases} \tag{A.0.18}$$

$$\gamma'_{i,back} = \begin{cases} \gamma_{i,back} & \text{if backward differencing} \\ 0 & \text{otherwise .} \end{cases} \tag{A.0.19}$$

Recalling equation (3.1.2)

$$V_i^{n+1} - V_i^n = \alpha_i^{n+1} V_{i-1}^{n+1} + \beta_i^{n+1} V_{i+1}^{n+1} - (\alpha_i^{n+1} + \beta_i^{n+1} + r\Delta\tau) V_i^{n+1} , \tag{A.0.20}$$

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

$$\begin{aligned} \alpha_i^n &= \alpha'_i - \gamma'_{i,cent} q_{i,cent}^n - \gamma'_{i,back} q_{i,back}^n \\ \beta_i^n &= \beta'_i + \gamma'_{i,cent} q_{i,cent}^n + \gamma'_{i,for} q_{i,for}^n . \end{aligned} \tag{A.0.21}$$

Appendix 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 [14].

Consider a short position, so that we can write equation (2.1.35) as

$$\begin{aligned} g(V,V_S,V_{SS},V_\tau) &= -V_\tau + \max_{q\in\{-1,+1\}} \left[SV_S \left[r' + q\,\lambda\sigma\sqrt{1-\rho^2}\,\right] + \frac{\sigma^2 S^2}{2} V_{SS} - rV \right] \\ &= 0\,. \end{aligned} \tag{B.0.1}$$

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

$$g(x,y,z+\varepsilon,w) \geq g(x,y,z,w) \;\;;\;\; \forall \varepsilon \geq 0\,, \tag{B.0.2}$$

which is our case simply means that $\sigma^2 \geq 0$. Suppose for the moment that smooth solutions to equation (B.0.1) 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 that $\phi(S_0,\tau_0) = V(S_0,\tau_0)$ at the single point (S_0,τ_0). Then, the single point (S_0,τ_0) is a global minimum of $(\phi - V)$,

$$\begin{aligned} \phi - V &\geq 0 \\ \min(\phi - V) &= \phi(S_0,\tau_0) - V(S_0,\tau_0) = 0\,. \end{aligned} \tag{B.0.3}$$

Consequently,

$$\begin{aligned} \phi_\tau &= V_\tau \\ \phi_S &= V_S \\ \phi_{SS} &\geq V_{SS} \,. \end{aligned} \tag{B.0.4}$$

Hence, from equations (B.0.2, B.0.4), we have

$$\begin{aligned} g(V(S_0,\tau_0),\phi_S(S_0,\tau_0),\phi_{SS}(S_0,\tau_0),\phi_\tau(S_0,\tau_0)) &= g(V(S_0,\tau_0),V_S(S_0,\tau_0),\phi_{SS}(S_0,\tau_0),V_\tau(\S_0,\tau_0)) \\ &\geq g(V(S_0,\tau_0),V_S(S_0,\tau_0),V_{SS}(S_0,\tau_0),V_\tau(S_0,\tau_0)) \\ &= 0 \,, \end{aligned} \tag{B.0.5}$$

or, to summarize,

$$\begin{aligned} g(V(S_0,\tau_0),\phi_S(S_0,\tau_0),\phi_{SS}(S_0,\tau_0),\phi_\tau(S_0,\tau_0)) &\geq 0 \\ \phi - V &\geq 0 \\ \min(\phi - V) &= \phi(S_0,\tau_0) - V(S_0,\tau_0) = 0 \,. \end{aligned} \tag{B.0.6}$$

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 $V-\chi$,

$$\begin{aligned} V - \chi &\geq 0 \\ \min(V-\chi) &= V(S_0,\tau_0) - \chi(S_0,\tau_0) \\ &= 0 \,. \end{aligned} \tag{B.0.7}$$

Repeating the above arguments we have

$$\begin{aligned} g(V(S_0,\tau_0),\chi_S(S_0,\tau_0),\chi_{SS}(S_0,\tau_0),\chi_\tau(S_0,\tau_0)) &\leq 0 \\ V - \chi &\geq 0 \\ \min(V-\chi) &= V(S_0,\tau_0) - \chi(S_0,\tau_0) = 0 \,. \end{aligned} \tag{B.0.8}$$

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.0.6) and (B.0.8) to define a viscosity solution to equation (B.0.1), since all derivatives are applied to smooth test functions. Informally, a viscosity solution V to equation (B.0.1) is defined such that

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

$$\phi - V \geq 0 \ ; \ \phi(S_0, \tau_0) = V(S_0, \tau_0) \ , \tag{B.0.9}$$

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

$$g(V(S_0, \tau_0), \phi_S(S_0, \tau_0), \phi_{SS}(S_0, \tau_0), \phi_\tau(S_0, \tau_0)) \geq 0 \ . \tag{B.0.10}$$

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

$$V - \chi \geq 0 \ ; \ V(S_0, \tau_0) = \chi(S_0, \tau_0) \ , \tag{B.0.11}$$

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

$$g(V(S_0, \tau_0), \chi_S(S_0, \tau_0), \chi_{SS}(S_0, \tau_0), \chi_\tau(S_0, \tau_0)) \leq 0 \ . \tag{B.0.12}$$

This definition is illustrated in Figure B.1.

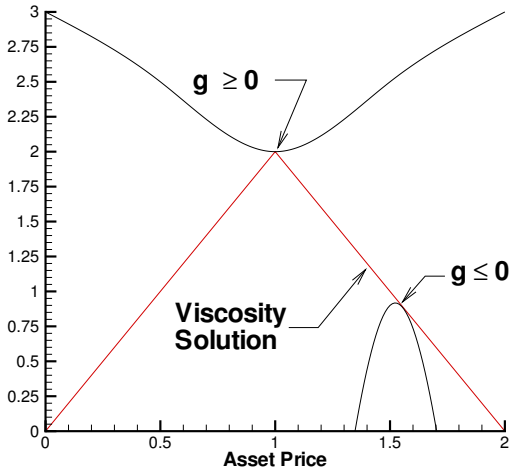


FIGURE B.1: *Illustration of viscosity solution definition. The upper and lower curves represent smooth test functions. The differential operator (B.0.1) can be applied to these test functions with the results given by equation (B.0.10) (upper curve), and equation (B.0.12) (lower curve). When a smooth test function* χ *touches the viscosity solution from below at* (S_0,τ_0)*, then* $g(V(S_0,\tau_0),\chi_S(S_0,\tau_0),\chi_{SS}(S_0,\tau_0),\chi_\tau(S_0,\tau_0)) \leq 0$*. When a smooth test function* ϕ *touches the viscosity solution from above at* $(S_0.\tau_0)$*, then* $g(V(S_0,\tau_0),\phi_S(S_0,\tau_0),\phi_{SS}(S_0,\tau_0),\phi_\tau(S_0,\tau_0)) \geq 0$*. 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.*

Appendix C

Grid Aspect Ratio Proof

In this appendix, we will prove Theorem (5.1.1). For convenience, we call nodes in the original grid old nodes and call nodes added by our algorithm (5.1.7) 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

$$(|r'| - \lambda\sigma\sqrt{1-\rho^2}) \geq 0, \tag{C.0.1}$$

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

$$(|r'| - \lambda\sigma\sqrt{1-\rho^2}) < 0. \tag{C.0.2}$$

Let

$$\mathcal{K} = -\frac{\sigma^2}{(|r'| - \lambda\sigma\sqrt{1-\rho^2})} > 0, \tag{C.0.3}$$

then in the new grid for $1 \leq i \leq m-1$, we have

$$S_{i+1} - S_{i-1} \leq \mathcal{K} S_i. \tag{C.0.4}$$

Now, we prove Theorem (5.1.1).

Proof. Suppose Theorem (5.1.1) is not true, then in the new grid $\exists i, 1 \leq i \leq m-1$ such that

$$\frac{S_{i+1}-S_i}{S_i-S_{i-1}} = t, \text{ where } t > q = max(5, 2q_0) \text{ or } t < p = min(\frac{1}{3}, p_0). \tag{C.0.5}$$

Let

$$a_i = S_i - S_{i-1}, \tag{C.0.6}$$

then

$$\frac{a_i}{a_{i-1}} = t. \tag{C.0.7}$$

Now suppose

$$\frac{a_i}{a_{i-1}} = t > q = max(5, 2q_0). \tag{C.0.8}$$

We prove the following observations first.

FIGURE C.1: *Condition (5.1.9) failed in a new grid.*

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

Proof. See Figure C.1. Suppose S_{i-1} is an old node, then

if S_i is also an old node, we will have

$$\frac{a_i}{a_{i-1}} \leq q_0; \tag{C.0.9}$$

if S_i is a new node, we will have

$$\frac{a_i}{a_{i-1}} \leq 1. \tag{C.0.10}$$

Both cases contradict with equation (C.0.8). 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 C.1. Otherwise we will have equation (C.0.10) ∎

By Observation 1 and Observation 2, we can know that S_{i-1} must be inserted in the middle of S_i and a node S_j, where $i > j$, as Figure C.2 shows below.

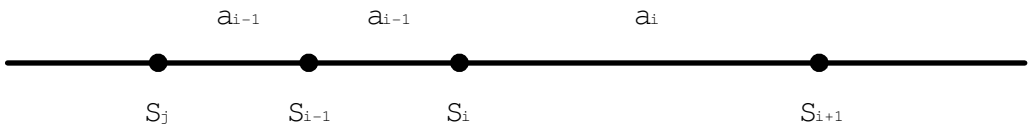


FIGURE C.2: *S_{i-1} is inserted at $\frac{S_j+S_i}{2}$.*

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

Proof. See Figure C.2. Suppose S_j is an old node. Then,

if S_i is also an old node, we will have

$$\frac{a_i}{a_{i-1}} \leq 2q_0; \tag{C.0.11}$$

if S_i is a new node, we will have

$$\frac{a_i}{a_{i-1}} \leq 2. \tag{C.0.12}$$

Both cases contradict with equation (C.0.8). Observation 3 follows.

∎

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

Proof. See Figure C.1. Otherwise we will have equation (C.0.12). ∎

By Observation 3 and 4, we can know that S_j must be inserted in the middle of S_i and a node S_h,where $h < j < i$, as Figure C.3 shows below.

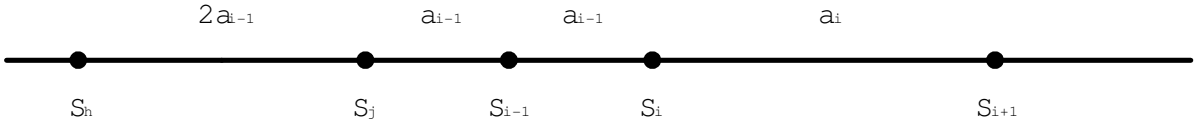


FIGURE C.3: *S_j is inserted at $\frac{S_h+S_i}{2}$.*

Observation 5. $2S_j \geq S_i$

Proof. See Figure C.3. Note that

$$S_j - S_h = S_i - S_j = 2a_{i-1} \tag{C.0.13}$$

and

$$S_h \geq 0, \tag{C.0.14}$$

so,

$$\frac{S_j}{S_i} = \frac{S_h + 2a_{i-1}}{S_h + 4a_{i-1}} \geq \frac{2a_{i-1}}{4a_{i-1}} = \frac{1}{2}, \tag{C.0.15}$$

hence,

$$2S_j \geq S_i. \tag{C.0.16}$$

∎

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

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

$$6a_{i-1} < (t+1)a_{i-1} = S_{i+1} - S_{i-1} \leq \mathcal{K} S_i. \tag{C.0.17}$$

∎

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

$$\begin{aligned} S_f - S_j &> \mathcal{K} S_i \\ S_i - S_j &\geq \frac{\mathcal{K} S_i}{2}, \end{aligned} \tag{C.0.18}$$

where S_f is the right neighbour of S_i when S_{i-1} is added, as Figure C.4 shows below. Note that $S_f \geq S_{i+1}$.

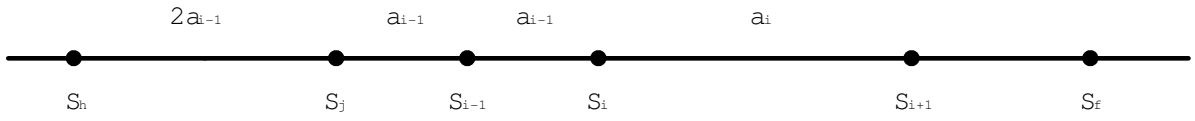


FIGURE C.4: *S_{i-1} is added because condition (C.0.18) is true.*

Then,

$$2a_{i-1} = S_i - S_j \geq \frac{\mathcal{K} S_i}{2}, \tag{C.0.19}$$

so,

$$4a_{i-1} \geq \mathcal{K} S_i, \tag{C.0.20}$$

which is a contradiction with Observation 6.

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

$$\begin{aligned} S_i - S_e &> \mathcal{K} S_j \\ S_j - S_e &< \tfrac{\mathcal{K} S_j}{2}, \end{aligned} \tag{C.0.21}$$

where S_e is the left neighbour of S_j when S_{i-1} is added, as Figure C.5 shows below.

Note that

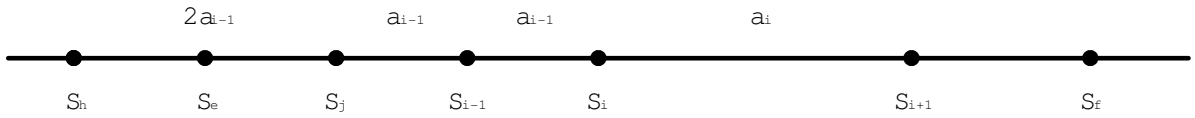


FIGURE C.5: *S_{i-1} is added because condition (C.0.21) is true.*

$$S_e \geq S_h + a_{i-1}. \tag{C.0.22}$$

Since suppose otherwise, we have

$$S_e = S_h, \tag{C.0.23}$$

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

$$2a_{i-1} = S_j - S_h = S_j - S_e < \frac{\mathcal{K} S_j}{2}, \tag{C.0.24}$$

so,

$$a_{i-1} < \frac{\mathcal{K} S_j}{4}. \tag{C.0.25}$$

But,

$$4a_{i-1} = S_i - S_h > \mathcal{K} S_j, \tag{C.0.26}$$

so,

$$a_{i-1} > \frac{\mathcal{K} S_j}{4}. \tag{C.0.27}$$

We get a contradiction. Hence, equation (C.0.24) is true.

Then we have

$$3a_{i-1} \geq S_i - S_e > \mathcal{K} S_j, \tag{C.0.28}$$

so,

$$a_{i-1} > \frac{\mathcal{K} S_j}{3}. \tag{C.0.29}$$

By observation 6 and equation (C.0.29), we can get

$$6a_{i-1} \leq \mathcal{K} S_i, \tag{C.0.30}$$

so,

$$\frac{6\mathcal{K} S_j}{3} < \mathcal{K} S_i, \tag{C.0.31}$$

hence,

$$2S_j < S_i, \tag{C.0.32}$$

which is a contradiction with observation 5.

Hence, we get contradictions in both cases. Therefore,

$$\frac{a_i}{a_{i-1}} \leq q = max(5, 2q_0). \tag{C.0.33}$$

Now suppose

$$\frac{a_i}{a_{i-1}} = t < p = min(\frac{1}{3}, p_0). \tag{C.0.34}$$

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.

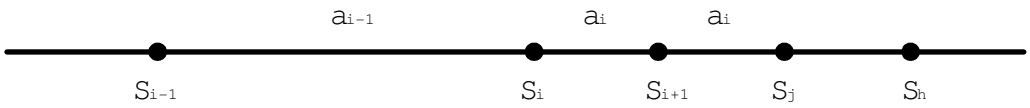


FIGURE C.6: *S_{i+1} is added because condition (C.0.35) is true.*

Since now $a_{i-1} > a_i$, the only reason to add S_{i+1} is

$$\begin{aligned} S_h - S_i &> \mathcal{K} S_j \\ S_j - S_i &\geq \frac{\mathcal{K} S_j}{2}, \end{aligned} \tag{C.0.35}$$

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

Hence,

$$2a_i = S_j - S_i \geq \frac{\mathcal{K} S_j}{2} \tag{C.0.36}$$

so,

$$a_i \geq \frac{\mathcal{K} S_j}{4}. \tag{C.0.37}$$

then,

$$\mathcal{K} S_j = \frac{3+1}{4} \mathcal{K} S_j < \frac{(\frac{1}{t}+1)\mathcal{K} S_j}{4} \leq (\frac{1}{t}+1) a_i = S_{i+1} - S_{i-1} < \mathcal{K} S_i. \tag{C.0.38}$$

so we get,

$$S_j < S_i \tag{C.0.39}$$

which is certainly false. Hence,

$$t \geq p = min(\frac{1}{3}, p_0). \tag{C.0.40}$$

The claim follows. ∎

Bibliography

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

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

[3] 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.

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

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

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

[7] J.P. Bouchaud, G. Iori, and D. Sornette. Real world options: Smile and residual risk. *RISK*, 9:61–65, March 1996.

[8] P. Carr, H. Geman, and D. Madan. Pricing and hedging in incomplete markets. *Journal of Financial Economics*, 61:131–167, 2001.

[9] 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, Universite Nancy I, 2004.

[10] F.H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, New York, 1983.

[11] T.F. Coleman, Yuying Li, and M.C. Patron. Hedging gurantees in variable annuities. Working paper, Cornell Theory Center, 2004.

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

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

[14] 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, July 1992.

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

[16] 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.

[17] 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*. to appear.

[18] P. Embrechts. Acturarial versus finanical pricing of insurance. *Risk Finance*, 1:4:17–26, 2000.

[19] H. Follmer and M. Schweitzer. Hedging by sequential regression. *ASTIN Bulletin*, 18:147–160, 1988.

[20] H. Follmerr and M. Schweizer. Hedging derivative securities in incomplete markets: An ε-arbitrage approach. *Operations Research*, 49:372–397, 2001.

[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] K. Froot and J. Stein. Risk management, capital budgeting and capital structure for financial institutions. *Journal of Financial Economics*, 47:55–82, 1998.

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

[24] V. Henderson. Valuation of claims on nontraded assets using utility maximization. *Mathematical Finance*, 12:351–373, 2002.

[25] T. Moller. On valuation and risk management at the interface of insurance and finance. Working paper, Laboratory of Actuarial Mathematics, University of Copenhagen, 2001.

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

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

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

[29] M. Otaka and Y. Kawaguchi. Hedging and pricing of real estate securities. Quantitative methods in Finance, Cairns, 2002.

[30] D. Pooley. Numerical methods for nonlinear equations in option pricing. PhD Thesis, School of Computer Science, University of Waterloo, 2003.

[31] 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.

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

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

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

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

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

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

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

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