# MAXIMAL USE OF CENTRAL DIFFERENCING FOR HAMILTON-JACOBI-BELLMAN PDES IN FINANCE

J. WANG * AND P.A. FORSYTH †

**Abstract.** In order to ensure convergence to the viscosity solution, the standard method for discretizing HJB PDEs uses forward/backward differencing for the drift term. In this paper, we devise a monotone method which uses central weighting as much as possible. In order to solve the discretized algebraic equations, we have to maximize a possibly discontinuous objective function at each node. Nevertheless, convergence of the overall iteration can be guaranteed. Numerical experiments on two examples from the finance literature show higher rates of convergence for this approach compared to the use of forward/backward differencing only.

**Key words.** Stochastic control, nonlinear HJB PDE, central differencing, monotone scheme, finance

**AMS subject classifications.** 65M12, 65M60, 91B28

**Version:** November 17, 2006

## 1. Introduction.

There are a number of financial models which result in nonlinear Hamilton-Jacobi-Bellman (HJB) partial differential equations (PDEs). These problems usually arise in the context of optimal stochastic control. Some examples of these HJB type equations include: transaction cost/uncertain volatility models [22, 3, 27], passport options [2, 30], unequal borrowing/lending costs [13], large investor effects [1], risk control in reinsurance [24], pricing options and insurance in incomplete markets using an instantaneous Sharpe ratio [32, 23, 11], minimizing ruin probability in insurance [29, 14], and optimal consumption [12, 16]. A recent survey article on the theoretical aspects of this topic is given in [26].

In many cases, classical solutions to these PDEs do not exist, and we seek to find the viscosity solution of the HJB equation [18]. It is important to ensure that the discrete solution converges to the viscosity solution as the mesh size and timestep are reduced (see [27] for an example where seemingly reasonable discretization methods can converge to non-viscosity solutions).

In order to guarantee convergence to the viscosity solution, the discrete scheme must be be pointwise consistent, $l_\infty$ stable and monotone [10, 4].

A monotone scheme is usually constructed by using a positive coefficient method [21, 25, 7, 19]. Typically, a positive coefficient method is developed using forward or backward differencing for the drift term. The choice of forward or backward differencing depends on the control variable. This has the disadvantage that the truncation error in the space-like direction is only first order.

If implicit timestepping is used, then the nonlinear discretized algebraic equations are solved by an iterative method. The usual iterative approach [21, 19] requires solution of a local optimization problem for the optimal control at each grid node, at every iteration. Since the discretization at each node is a function of the control variable at that node, the type of discretization (i.e. forward or backward differencing) may change at each iteration. Use of forward/backward differencing means that the local objective function at each node is a continuous function of the control variable, but non-smooth.

* David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 e-mail: `j27wang@uwaterloo.ca`,

† David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 e-mail: `paforsyt@uwaterloo.ca`

In this paper, we take a slightly different approach compared to the standard technique. We will use a combination of central/forward/backward differencing at each node. Given a value of the the control variable at a node, we use the following criteria to select the differencing method

- Central differencing is used if the discretization is a positive coefficient method (for this particular choice of control).
- Forward/backward differencing is used only if central differencing does not result in a positive coefficient method. One of forward or backward differencing must satisfy the positive coefficient condition.

This method has the advantage that central differencing is used as much as possible, so that use of a locally second order method is maximized. However, in general, the local objective function at each grid node is now a discontinuous function of the control. It would appear that iterative solution of the discretized equations would be problematic in this case.

In this paper we note that the proof of convergence of the iterative scheme for solution of the discretized algebraic equations does not, in fact, require continuity of the local objective function. Hence, convergence of the iterative method for solution of the discrete equations can be guaranteed, even if the local objective function is a discontinuous function of the control. Nevertheless, it is not clear that, in practice

- the use of a locally second order method as much as possible will result in improved convergence as the mesh is refined, for practical parameter values;
- that the rate of convergence of the nonlinear iteration will be acceptable, if the local objective function is a discontinuous function of the control.

We report the results of several numerical experiments, for passport options [2] and optimal asset allocation[15]. These experiments show that we can often obtain higher rates of convergence using central differencing as much as possible, although at some additional cost compared to the standard approach.

**2. General Form for the Example Problems.** To avoid repetition, we will carry out our analysis for a general form for the example problems.

As is typically the case with finance problems, we solve backwards in time from the expiry date of the contract $t = T$ to $t = 0$ by use of the variable $\tau = T - t$. Set

$$\mathcal{L}^Q V \equiv a(S,\tau,Q)V_{SS} + b(S,\tau,Q)V_S - c(S,\tau,Q)V \ , \tag{2.1}$$

where the control parameter $Q$ is in general a vector, that is, $Q = (q_1, q_2, \ldots)'$. We write our problems in the general form

$$\begin{aligned} V_\tau &= \sup_{Q \in \hat{Q}} \left\{ \mathcal{L}^Q V + d(S,\tau,Q) \right\} \\ & S \in [S_{\min}, S_{\max}] \ \ , 0 \leq \tau \leq T \ . \end{aligned} \tag{2.2}$$

Here we include the $d(S,t,Q)$ term in equation (2.2) for generality, although in the examples in this paper, we will always have $d \equiv 0$. We can also replace the $\sup$ in equation (2.2) by an $\inf$, and all the results of this paper hold in this case as well.

**2.1. Boundary Conditions.** At $\tau = 0$, we set $V(S,0)$ to the specified contract payoff. As $S \to S_{\min}, S \to S_{\max}$, we assume that either

- a Dirichlet boundary condition is specified;
- the coefficient $a(S,\tau,Q)$ vanishes, and the sign of $b(S,\tau,Q)$ is such that no boundary condition is required [9].

Note that it may be the case that the original problem has $S_{\max} = +\infty$ or $S_{\min} = -\infty$. In these cases, we will use a finite computational domain, and we assume that financial reasoning can be used to determine an appropriate Dirichlet condition. This is clearly an approximation, and introduces a *localization error*. However, As pointed out in [6], we can expect any errors incurred by imposing approximate boundary conditions at finite values of $S_{\min}, S_{\max}$ to be small in areas of interest if $|S_{\min}|$, $|S_{\max}|$ are selected sufficiently large. We will assume in the following that the original problem has been localized to a finite domain.

ASSUMPTION 2.1 (Properties of the HJB PDE.). *We make the assumption that the coefficients $a,\ b,\ c,\ d$ are continuous functions of $(S, \tau, Q)$, with $a \geq 0$, and $c \geq 0$, and that $a, b, c, d$, are bounded on $S_{\min} \leq S \leq S_{max}, Q \in \hat{Q}$. Since we restrict ourselves to a finite computational domain $S_{\min} \leq S \leq S_{max}$, we avoid difficulties associated with coefficients that grow with $S$ as $|S| \to \infty$. We also assume that the controls $Q$ are bounded. It follows from [17, 8] that solutions to equation (2.2) along with suitable boundary conditions satisfy the strong comparison property. Hence, we make the assumption that a unique viscosity solution exists for equation (2.2).*

**3. Implicit Controls.** Define a grid $\{S_0, S_1, \ldots, S_p\}$ with $S_0 = S_{\min}$, $S_p = S_{\max}$ and let $V_i^n$ be a discrete approximation to $V(S_i, \tau^n)$. Let $V^n = [V_0^n, \ldots, V_p^n]'$, and let $(\mathcal{L}_h^Q V^n)_i$ denote the discrete form of the differential operator (2.2) at node $(S_i, \tau^n)$. The operator (2.2) can be discretized using forward, backward or central differencing in the $S$ direction to give

$$(\mathcal{L}_h^Q V^{n+1})_i = \alpha_i^{n+1} V_{i-1}^{n+1} + \beta_i^{n+1} V_{i+1}^{n+1} - (\alpha_i^{n+1} + \beta_i^{n+1} + c_i^{n+1}) V_i^{n+1} \ . \tag{3.1}$$

Here $\alpha_i,\ \beta_i$ are defined in Appendix A.

It is important that central, forward or backward discretizations be used to ensure that (3.3) is a positive coefficient discretization. To be more precise, this condition is

CONDITION 3.1. *Positive Coefficient Condition*

$$\alpha_i^{n+1} \geq 0,\ \ \beta_i^{n+1} \geq 0,\ \ c_i^{n+1} \geq 0.\ \ i = 0, .., p\ \ . \tag{3.2}$$

We will assume that all models have $c_i^{n+1} \geq 0$. Consequently, we choose central, forward or backward differencing at each node to ensure that $\alpha_i^{n+1},\ \beta_i^{n+1} \geq 0$. Note that different nodes can have different discretization schemes.

Equation (2.2) can now be discretized using fully implicit timestepping along with the discretization (3.1) to give

$$\frac{V_i^{n+1} - V_i^n}{\Delta \tau} = \max_{Q^{n+1} \in \hat{Q}} \left\{ (\mathcal{L}_h^{Q^{n+1}} V^{n+1})_i + d_i^{n+1} \right\} \ . \tag{3.3}$$

Note that $\alpha_i^{n+1} = \alpha_i^{n+1}(Q_i^{n+1})$, $\beta_i^{n+1} = \beta_i^{n+1}(Q_i^{n+1})$, $c_i^{n+1} = c_i^{n+1}(Q_i^{n+1})$ and $d_i^{n+1} = d_i^{n+1}(Q_i^{n+1})$, that is, the discrete equation coefficients are functions of the local optimal control $Q_i^{n+1}$. This makes equations (3.3) highly nonlinear in general. We refer to methods which use an implicit timestepping method where the control is handled implicitly as an *implicit control* method in the following.

**3.1. Matrix Form of the Discrete Equations.** It will be convenient to use matrix notation for equations (3.3), coupled with boundary conditions.

If a Dirichlet condition is specified at $S = S_{\min}, \tau = \tau^n$ $(i = 0)$, then we denote this value by $G_0^n$. If a Dirichlet boundary condition is specified at $S = S_{\max}, \tau = \tau^n$ $(i = p)$, then we denote this value by $G_p^n$. Set $Q^n = [Q_0^n, Q_1^n, \ldots, Q_p^n]'$, with each $Q_i^n$ a local optimal control. We can write the discrete operator $(\mathcal{L}_h^Q V^n)_i$ as

$$\begin{aligned}(\mathcal{L}_h^Q V^n)_i &= [A^n V^n]_i \\ &= \left[\alpha_i^n V_{i-1}^n + \beta_i^n V_{i+1}^n - (\alpha_i^n + \beta_i^n + c_i^n) V_i^n\right]; \qquad 1 < i < p.\end{aligned} \tag{3.4}$$

The first and last rows of $A$ are modified as needed to handle the boundary conditions. The boundary conditions at $S = S_{\min}, S_{\max}$ can be enforced by specifying a boundary condition vector $G^n = [G^n_0, 0, \ldots, 0, G^n_p]'$. If a Dirichlet condition is specified at $i = p$, we set $G^n_p$ to the appropriate value, and set the last row in $A^n$ to be zero. With a slight abuse of notation, we denote this last row in this case as $(A^n)_p \equiv 0$. Conversely, if no boundary condition is required at $i = p$, then we use backward differencing at node $i = p$ (which means that $\beta_p = 0$), and set $G^n_p = 0$. The boundary condition at $S = S_{\min}$, $i = 0$, is handled in a similar fashion. Let $D^n$ be the diagonal matrix with entries

$$[D]^n_{ii} = \begin{cases} d^n_i \ , \ \ 1 < i < p \\ \ 0 \ \ , \ \ i = 0, p \ ; \ \ \text{if a Dirichlet condition is specified} \\ d^n_i \ \ , \ \ i = 0, p \ ; \ \ \text{if no boundary condition is required} \end{cases}$$

Recalling that $A^n = A(Q^n)$, then the discrete equations (3.3) can be written as

$$\begin{aligned} &\left[I - \Delta\tau A^{n+1}\right] V^{n+1} = V^n + \Delta\tau D^{n+1} + (G^{n+1} - G^n) \ , \\ &\qquad \text{where } Q^{n+1}_i = \underset{Q^{n+1}_i \in \hat{Q}}{\arg\max} \left\{ \left[A^{n+1}(Q^{n+1}) V^{n+1} + D^{n+1}(Q^{n+1})\right]_i \right\} \ . \end{aligned} \tag{3.5}$$

Here the term $(G^{n+1} - G^n)$ enforces possible Dirichlet boundary conditions at $S = S_0, S_p$. Note also that the discrete equations (3.5) are nonlinear since $A^{n+1} = A(Q^{n+1})$ and $Q^{n+1} = Q^{n+1}(V^{n+1})$.

**4. Convergence to the Viscosity Solution.** In [27], examples were given in which seemingly reasonable discretizations of nonlinear option pricing PDEs were unstable or converged to the incorrect solution. It is important to ensure that we can generate discretizations which are guaranteed to converge to the viscosity solution [4, 18]. Assuming that equation (2.2) satisfies the strong comparison property [5, 8, 17], then, from [10, 4], a numerical scheme converges to the viscosity solution if the method is pointwise consistent, stable (in the $l_\infty$ norm) and monotone.

It is straightforward, using the methods in [7, 19] to show that scheme (3.3) is monotone, pointwise consistent, and stable.

THEOREM 4.1 (Convergence to the Viscosity Solution). *Provided that the original HJB satisfies Assumption 2.1 and discretization (3.5) satisfies the positive coefficient condition (3.2) with suitable boundary conditions then scheme (3.5) converges to the viscosity solution of equation (2.2).*

*Proof.* Using the methods in [7, 19], this can be shown to follow from results in [10, 4]. □

REMARK 4.1 (Rate of Convergence). *If $\Delta S = \max_i (S_{i+1} - S_i)$ and $\Delta\tau = C_1 h$, $\Delta S = C_2 h$, where $C_1, C_2$ are positive constants, then there has been considerable effort in recent years in attempts to determine rates of convergence for monotone finite difference schemes for HJB equations. Typically, one obtains estimates of the error of the form $O(h^\rho)$ where $\rho$ varies from $1/27$ to $1/2$ depending on assumptions about regularity of the solution and the PDE coefficients. See [7] for an overview of recent work along these lines. These results seem generally pessimistic when compared with numerical experiments.*

It is also useful to note the following property of the matrix $[I - \Delta\tau A^n]$.

LEMMA 4.2 (M-matrix). *If the positive coefficient condition (3.1) is satisfied, and either Dirichlet boundary conditions are specified, or no boundary condition is required, then $[I - \Delta\tau A^n]$ is an M-matrix.*

*Proof.* Condition (3.1) implies that $\alpha_i^n, \beta_i^n, c_i^n$ in equation (3.4) are non-negative. Hence $[I - \Delta\tau A^n]$ has positive diagonals, non-positive off diagonals, and is diagonally dominant, hence it is an M-matrix. □

**4.1. Discretization of the Control.** Suppose we have a single control $q \in \hat{Q}$ where $\hat{Q} = [q_{\min}, q_{\max}]$, where $(q_{\min}, q_{\max})$ are bounded. It is sometimes convenient to discretize the control, i.e. we replace $\hat{Q}$ by $\hat{Y}$ where $\hat{Y} = [y_0, y_1, y_2, ..., y_k]$, with $y_0 = q_{\min}, y_k = q_{\max}$. Let $\max_i(y_{i+1} - y_i) = C_3 h$, where $C_3$ is a positive constant. Then we have the following Lemma.

LEMMA 4.3 (Consistency of Discrete Control Approximation). *If the HJB equation satisfies Assumption 2.1, then the discretized control problem with* $\max_i(y_{i+1} - y_i) = C_3 h$

$$V_\tau = \sup_{Q \in \hat{Y}} \left\{ \mathcal{L}^Q V + d(S, \tau, Q) \right\} \tag{4.1}$$

*is consistent with equation (2.2).*

*Proof.* Let $\phi(S,\tau)$ be a smooth test function possessing bounded derivatives of appropriate order, then, in view of the fact that the coefficients of equation (2.2) are assumed to be continuous, bounded functions of $Q$, then

$$\begin{aligned} &\left| \phi_\tau - \sup_{Q \in \hat{Q}} \left\{ \mathcal{L}^Q \phi + d(S,\tau,Q) \right\} - \left( \phi_\tau - \sup_{Q \in \hat{Y}} \left\{ \mathcal{L}^Q \phi + d(S,\tau,Q) \right\} \right) \right| \\ &= O(h) \ . \end{aligned} \tag{4.2}$$

□

LEMMA 4.4 (Discrete Control Approximation: Convergence to the Viscosity Solution). *Let $\Delta\tau = C_1 h$, $\max_i(S_{i+1} - S_i) = C_2 h$, $\max_i(y_{i+1} - y_i) = C_3 h$, with $C_i$ being positive constants. Provided the conditions for Theorem 4.1 and Lemma 4.3 are satisfied, then the discretization (3.5) with $\hat{Q}$ replaced by the discrete control set $\hat{Y}$ converges to the viscosity solution of equation (2.2) as $h \to 0$.*

*Proof.* This follows immediately from Theorem 4.1 and Lemma 4.3. □

**5. Solution of Algebraic Discrete Equations .** Although we have established that discretization (3.5) is consistent, stable and monotone, it is not obvious that this is a practical scheme, since the implicit timestepping method requires solution of highly nonlinear algebraic equations at each timestep.

**5.1. Iterative Method.** Consider the following iteration scheme:

**Iterative Solution of the Discrete Equations**

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

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

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

Solve

$$\begin{aligned} \big[I - \Delta\tau A(Q^k)\big]\, \hat{V}^{k+1} &= V^n + (G^{n+1} - G^n) + \Delta\tau D^k(Q^k) \\ Q_i^k &= \operatorname*{arg\,max}_{Q_i^k \in \hat{Q}} \left\{ \Big[ A^k(Q^k)\hat{V}^k + D^k(Q^k) \Big]_i \right\} \end{aligned} \tag{5.1}$$

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

EndFor

The term $scale$ in scheme (5.1) is used to ensure that unrealistic levels of accuracy are not required when the value is very small. Typically, $scale = 1$ for options priced in dollars.

Some manipulation of Algorithm (5.1) results in

$$\big[I - \Delta\tau A^k\big]\, (\hat{V}^{k+1} - \hat{V}^k) = \Delta\tau \left[ (A^k\hat{V}^k + D^k) - (A^{k-1}\hat{V}^k + D^{k-1}) \right] . \tag{5.2}$$

The proof of convergence of the iteration scheme (5.1) is given in [19]. In [28, 32], a similar proof was given, but only for the case where the discretization did not depend on the control. Since scheme (5.1) can be regarded as a variant of Policy iteration for infinite horizon Markov chains, the convergence proof is similar to the proof of convergence for Policy iteration [21]. For the convenience of the reader, we sketch the proof below.

In order to prove the convergence of Algorithm (5.1), we first need an intermediate result.

LEMMA 5.1 (Sign of RHS of Equation (5.2)). *If $A^k(Q^k)\hat{V}^k$ is given by equation (3.4), with the control parameter determined by*

$$Q_i^k = \operatorname*{arg\,max}_{Q_i^k \in \hat{Q}} \left\{ \Big[ A^k(Q^k)\hat{V}^k + D^k(Q^k) \Big]_i \right\} , \tag{5.3}$$

*then every element of the right hand side of equation (5.2) is nonnegative, that is,*

$$\Big[ (A^k\hat{V}^k + D^k) - (A^{k-1}\hat{V}^k + D^{k-1}) \Big]_i \geq 0 . \tag{5.4}$$

*Proof.* Recall that $Q^k$ is selected so as to maximize $A^k\hat{V}^k + D^k$, for given $\hat{V}^k$. Hence, any other choice of coefficients, for example $A^{k-1}\hat{V}^k + D^{k-1}$ cannot exceed $A^k\hat{V}^k + D^k$. □

It is now easy to show that iteration (5.1) always converges.

THEOREM 5.2 (Convergence of Iteration (5.1)). *Provided that the conditions for Lemmas 4.2 and 5.1 are satisfied, then the iteration (5.1) converges to the unique solution of equation (3.5) for any initial iterate $\hat{V}^0$. Moreover, the iterates converge monotonically.*

*Proof.* Given Lemmas 5.1 and 4.2, the proof of this result is similar to the proof of convergence given in [27]. We give a brief outline of the steps in this proof, and refer readers to [27] for details. A straightforward maximum analysis of scheme (5.1) can be used to bound $\|\hat{V}^k\|_\infty$ independent of iteration $k$. From Lemma 5.1, we have that the right hand side of equation (5.2) is non-negative. Noting that $\left[I - \Delta\tau A^k\right]$ is an M-matrix (from Lemma 4.2) and hence $\left[I - \Delta\tau A^k\right]^{-1} \geq 0$, it is easily seen that the iterates form a bounded non-decreasing 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 $\left[I - \Delta\tau A^k\right]$ that the solution is unique.

□

REMARK 5.1 (Q Dependent Discretizations). *Note that we obtain convergence of iteration (5.1) if the discretization depends on the control Q, and, in particular, even if the discrete equations, regarded as a function of the control Q, are discontinuous. We do, however, require that the coefficients $a, b, c, d$ in equations (2.1-2.2) are bounded functions of the control $Q$, in order to ensure that the maximum in equation (5.3) exists.*

REMARK 5.2 (Uniqueness of Solution). *The above argument shows that the solution for $V^{n+1}$ is unique. However, this does not imply that the controls $Q^{n+1}$ are unique. As a simple counterexample, consider the case where $a, b, c, d$ in equations (2.1-2.2) are independent of $Q$, in which case the solution for $V$ is unique for any choice of $Q \in \hat{Q}$.*

**6. Passport Options.** Passport options are financial derivative contracts which allow the holder to take profit from a trading account while obligating the writer to cover losses [2, 30]. The holder is allowed to trade an underlying asset $S$ at any time during the option life time, say $T$. Let $q$ denote the number of shares of the underlying the holder holds at time $t$, $0 \leq t \leq T$. $q$ is limited to an amount $C$, i.e. $|q| \leq C$. At the maturity $T$, the holder keeps any net gain, while any loss is covered by the writer.

In [2], this problem is solved using central weighting. While the results appear to converge to the correct solution, convergence to the viscosity solution cannot be guaranteed. In [28], it is shown that, for passport options, it is not possible to pre-select central, forward or backward differencing at a node, independent of the control, and guarantee a positive coefficient scheme. In other words, the scheme must depend on the control.

**6.1. The Pricing Model for Passport Options.** Let $S$ be the underlying asset price which follows the stochastic process

$$dS = \mu S \, dt + \sigma S \, dZ \ , \tag{6.1}$$

where $dZ$ is the increment of a Wiener process, $\sigma$ is volatility, $\mu$ is the drift rate. Let $V(S, W, t)$ denote the option value at time $t$ with underlying price $S$ and wealth $W$. Under the process (6.1), the pricing PDE for passport options can be written as [31]

$$\begin{aligned} -V_t = &- rV + (r-\gamma) S V_S \\ &+ \sup_{|q| \leq 1} \left[ -((\gamma - r + r_c) qS - r_t W) V_W + \frac{\sigma^2 S^2}{2} (V_{SS} + 2qV_{SW} + q^2 V_{WW}) \right] \ , \end{aligned} \tag{6.2}$$

where

$W$ is the accumulated wealth of the underlying trading account.

$r$ is the risk-free interest rate.

$\gamma$ is the dividend rate on the underlying asset S.

$r_c$ is a cost of carry rate.

$r_t$ is an interest rate for the trading account.

$q$ is the number of shares of $S$ that an investor holds, which is also termed the trading strategy. $q$ is limited to $|q| \leq 1$ in equation (6.2). Different position limits can be handled by scaling [20].

We consider two types of payoff at $t = T$. The standard payoff is

$$V(S, W, t = T) = \max(W, 0) \ , \tag{6.3}$$

and the *asset or nothing* payoff [2]

$$V(S, W, t = T) = \begin{cases} S & \text{if } W \geq 0 \\ 0 & \text{otherwise} \end{cases} \ . \tag{6.4}$$

The above payoff can be generalized to specify a non-zero strike, but we assume the form (6.4) in this paper.

For both these payoffs, we can reduce the problem to solving for $V(S, W, t) = Su(x, \tau)$, where $x = W/S$ and $\tau = T - t$ [2], so that equation (6.2) can be reduced to a one-dimensional problem for $u$

$$u_\tau = -\gamma u + \sup_{|q| \leq 1} \left[ ((r - \gamma - r_c)q - (r - \gamma - r_t)x)u_x + \frac{\sigma^2}{2}(x - q)^2 u_{xx} \right] \ , \tag{6.5}$$

where $x \in [-\infty, +\infty]$. The standard payoff becomes

$$u(x, \tau = 0) = \max(x, 0), \tag{6.6}$$

with boundary conditions

$$u(x \to -\infty, \tau) = 0 \ ; \ u(x \to \infty, \tau) = x \ , \tag{6.7}$$

while the asset or nothing payoff is

$$u(x, \tau = 0) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.8}$$

with boundary conditions

$$u(x \to -\infty, \tau) = 0 \ ; \ u(x \to \infty, \tau) = \exp(-\gamma\tau) \ . \tag{6.9}$$

For computational purposes, we truncate the domain to $x \in [x_{\min}, x_{\max}]$, and apply the boundary conditions (6.7) and (6.9) at $x_{\min}, x_{\max}$.

**6.2. Discretization.** Passport option valuation is a special case of the general HJB equation (2.2), if we note that in this case

$$\begin{aligned} &Q = (q) \ , \ \ \hat{Q} = [-1, +1] \ , \ \ a(x, \tau, Q) = \frac{\sigma^2}{2}(x - q)^2 \ , \\ &b(x, \tau, Q) = (r - \gamma - r_c)q - (r - \gamma - r_t)x \ , \\ &c(x, \tau, Q) = \gamma \ , \ \ d(x, \tau, Q) = 0 \ , \end{aligned} \tag{6.10}$$

where $Q, \ \hat{Q}, \ a, \ b, \ c$ are defined in Section 2. Let the discrete approximation for $u(x_i, \tau^n)$ be denoted by $u_i^n$ with $U^n = [u_0^n, u_1^n, ..., u_{imax}^n]'$. Let $\hat{U}^k$ be the $k'th$ estimate for $U^n$, then the local objective function which must be maximized in Algorithm 5.1 is

$$\begin{aligned} [F(q)]_i &= [A(Q)\hat{U}^k + D(Q)]_i \\ &= ((r - \gamma - r_c)q - (r - \gamma - r_t)x_i)[(\hat{u}^k)_x]_i + \frac{\sigma^2}{2}(x_i - q)^2[(\hat{u}^k)_{xx}]_i \ , \end{aligned} \tag{6.11}$$

where $[\cdot]_i$ refers to the discrete form for $[\cdot]$. We need a positive coefficient scheme to solve the pricing PDE (6.5). Given node $x = x_i$, with current solution estimate $\hat{U}^k$, suppose the sets of $q$'s, which give a positive coefficient scheme for central, forward and backward differencing respectively, are $P_i^{cent},\ P_i^{fwd},\ P_i^{bwd}$. Since central differencing is the most accurate, it should be used as much as possible. Consequently, given a $q$, if central differencing satisfies positive coefficient conditions, central differencing will be used for that $q$. In other words, the proper ranges of $q$ for various differencings are $Range_i^{cent} = P_i^{cent}$, $Range_i^{fwd} = P_i^{fwd} - (P_i^{cent} \cap P_i^{fwd})$ and $Range_i^{bwd} = P_i^{bwd} - (P_i^{cent} \cap P_i^{bwd})$. Note that

$$\begin{aligned} &Range_i^j \cap Range_i^k = \emptyset \text{ , where } j,k \in \{cent, bwd, fwd\} \text{ and } j \neq k \\ &Range_i^{bwd} \cup Range_i^{fwd} \cup Range_i^{cent} = \hat{Q} \ , \end{aligned} \tag{6.12}$$

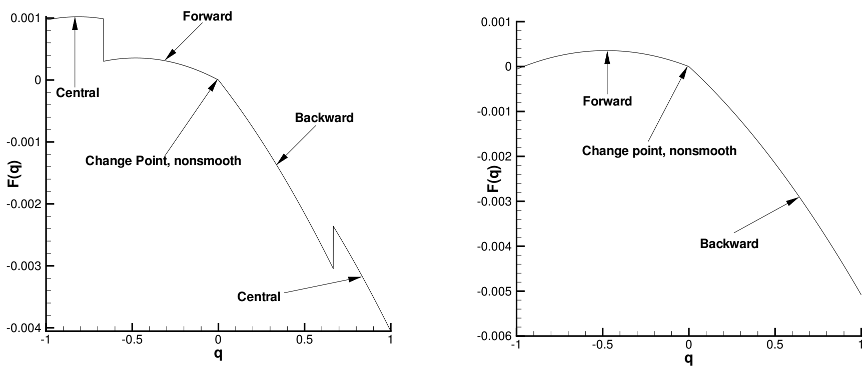
so that $[F(q)]_i$ is a well-defined function of $q$. For a given asset grid, and the option values at the current iteration, Algorithm 6.13 is used to determine the optimal control and to decide which differencing should be applied. For a given differencing method, the range of possible values of the control is divided into segments where the objective function is smooth. Standard methods are then used to determine the maximum within each segment. If an analytic form for the local objective function is not available, then an alternate approach is discussed in Section 7.4.

**Determining the Optimal Control and the Differencing Method**

Apply boundary conditions at the first node $x_0$ and the last node $x_p$
For each $x_0 < x_i < x_p$
  Compute the positive coefficient sets $Range_i^{cent},\ Range_i^{fwd},\ Range_i^{bwd}$
  diff = cent, $q^* = 0,\ F_{max} = -\infty$
  For j = cent, fwd, bwd
    Solve $q_j^* = \underset{q \in Range_i^j}{\arg\max} \{[F(q)]_i\}$ (6.13)
    If $[F(q_j^*)]_i > F_{max}$
      diff $= j,\ q^* = q_j^*,\ F_{max} = [F(q_j^*)]_i$
    EndIf
  EndFor
EndFor

**6.3. Discontinuity of The Objective Function.** When we use central differencing as much as possible, the local objective function at each node is in general a discontinuous function of the control $q$. However, the proof of convergence of the iterative scheme for solution of the discretized algebraic equations (Theorem 5.2) does not require continuity of the local objective function. If forward/backward differencing are applied, the local objective function is continuous but not smooth. Figure 6.1 shows these features, for a given grid and solution values.

Note that Algorithm 6.13 requires maximization of the local objective function for each set of points where central, forward, or backward differencing is used. For example, as shown in Figure 6.1, central differencing can be used on two disjoint intervals of the control space. On each of the subintervals, the objective function is a smooth function of the control, hence we can use standard methods to maximize the objective function. Determination of the range of controls where central, forward and backward differencing gives rise to a positive coefficient method is generally only possible if we have an analytic expression for the objective function. If this is not available, we can discretize the control, and use a linear search to find the maximum as described in Section 7.4. This is, of course, much more computationally expensive compared to analytic maximization.




(a) (b)

FIG. 6.1: *Local objective function (6.11) for the Passport option at* ($x_i = 0$). *(a) using central differencing as much as possible; (b) using forward/backward differencing only. Parameters:* $r = 0.05,\ \sigma = 0.03,\ dividend = 0.04,\ r_c = 0.07,\ r_t = 0.03$. *Nodes:* $x_{i-1} = -0.01,\ u_{i-1} = 0.173298$*;* $x_i = 0.,\ u_i = 0.173888$*;* $x_{i+1} = 0.01,\ u_{i+1} = 0.174135$.

**6.4. Numerical Results.** In this section, we will examine the convergence as the grid and timesteps are refined for various differencing methods.

We use a convergence tolerance of $10^{-7}$ in Algorithm 5.1, and we truncate the computational domain to $[x_{\min}, x_{\max}] = [-3, 4]$. Numerical experiments show that increasing the size of the computational domain does not affect solution values to six digits. The input parameters are given in Table 6.1.

Table 6.2 presents a convergence study, which also reports the actual initial option values, i.e. $V = S_0\, u(x = W_0/S_0, \tau = T)$.

An unequally spaced grid in the $x$ direction is used, and new fine grid nodes are added between each two coarse grid nodes at each level of refinement. In Table 6.2, *ratio* refers to the ratio of successive changes in the solution as the grid is refined by a factor of two, and the timestep sizes are reduced by a factor of four. Since fully implicit timestepping is used, this allows us to isolate the effect of the use of central weighting as much as possible, compared to forward/backward differencing only. Local second order convergence (in terms of $x$ node spacing) would be consistent with a ratio of four, while first order convergence would be consistent with a ratio of two.

The payoff type is a call (convex payoff). As expected, quadratic convergence is obtained by using central differencing as much as possible, and first order convergence is obtained by using forward/backward differencing.

For a convex payoff, it is always optimal to choose $q = -1$ or $1$ [28]. But for a non-convex payoff, $q$ can be any value in $[-1, 1]$. Table 6.3 presents a convergence study using the parameters in Table 6.1, but the payoff type is an asset or nothing with strike $K = 0$, i.e. a non-convex payoff (this is a digital call in terms of $u$, see equation (6.8) ). For the asset

**Table 6.1** Parameters for the convex payoff, passport option.

| | | | |
|---|---|---|---|
| $r$ | 0.08 | $\sigma$ | 0.2 |
| dividend rate $\gamma$ | 0.03 | $r_c$ | 0.12 |
| $r_t$ | 0.05 | $S_0$ | \$100 |
| Payoff | Call | Strike | \$10 |
| Initial Wealth | 0 | Time to expiry $T$ | 1 year |

**Table 6.2** Convergence study, passport option, convex payoff. Fully implicit timestepping is applied, using constant timesteps. On each refinement, new nodes are inserted between each coarse grid node, and the timestep is divided by four. Parameters are given in Table 6.1. Ratio is the ratio of successive changes in computed solution as the discretization parameters are reduced.

| Nodes | Timesteps | Nonlinear iterations | CPU Time (Sec) | Option value | Ratio |
|---|---|---|---|---|---|
| Central Differencing as much as possible | | | | | |
| 133 | 100 | 223 | 0.05 | 6.75321 | |
| 265 | 400 | 847 | 0.46 | 6.75951 | |
| 529 | 1600 | 3237 | 3.51 | 6.76110 | 3.962 |
| 1057 | 6400 | 12801 | 27.93 | 6.76150 | 3.995 |
| 2113 | 25600 | 51201 | 217.53 | 6.76160 | 3.980 |
| Forward/backward differencing only | | | | | |
| 133 | 100 | 223 | 0.04 | 6.79012 | |
| 265 | 400 | 846 | 0.28 | 6.77809 | |
| 529 | 1600 | 3241 | 2.08 | 6.77042 | 1.569 |
| 1057 | 6400 | 12801 | 16.53 | 6.76617 | 1.803 |
| 2113 | 25600 | 51201 | 130.93 | 6.76393 | 1.905 |

or nothing payoff, if $W_0 \geq 0$, the option value will be very high (close to the initial stock value) and insensitive to the grid refinement, so it is difficult to carry out convergence study. In this example, option values are reported at $W_0 = -25$ (initial wealth is $-\$25$). When central differencing as much as possible is applied, the convergence rate is close to second order. First order convergence is obtained by using forward/backward differencing.

From the numerical results, we can conclude that generally we can obtain higher rates of convergence using central weighting as much as possible, compared to forward/backward differencing only. Of course, we cannot guarantee that this will always occur, but we can rarely obtain second order convergence using forward/backward differencing. In all our numerical experiments, we have never seen a case where central weighting as much as possible converges at a slower rate compared to forward/backward differencing only.

Note that in both these examples, forward/backward differencing only requires about 60% of the CPU time compared to central differencing as much as possible. This is simply because of the additional tests required to determine the ranges of possible central weighting in Algorithm 6.13.

Both Table 6.2 and Table 6.3 show that the number of nonlinear iterations per timestep is about two, indicating that Algorithm 5.1 converges rapidly, in spite of the discontinuous objective function that is maximized at each node.

**7. Defined Contribution Pension Plan.** The second example in this paper concerns an optimal dynamic asset allocation strategy for a defined contribution pension plan. A

**Table 6.3** Convergence study, passport option, non-convex payoff. Fully implicit timestepping is applied, using constant timesteps. On each refinement, new nodes are inserted between each coarse grid node, and the timestep is divided by four. Parameters are given in Table 6.1 except that $W_0 = -25$, $K = 0$, and the payoff is an asset or nothing. Ratio is the ratio of successive changes in computed solution as the discretization parameters are reduced.

| Nodes | Timesteps | Nonlinear iterations | CPU Time (Sec) | Option value | Ratio |
|---|---|---|---|---|---|
| Central Differencing as much as possible, $W_0 = -25$ | | | | | |
| 133 | 100 | 274 | 0.06 | 26.6543 | |
| 265 | 400 | 873 | 0.44 | 26.9001 | |
| 529 | 1600 | 3307 | 3.26 | 26.9650 | 3.786 |
| 1057 | 6400 | 12900 | 25.46 | 26.9819 | 3.852 |
| 2113 | 25600 | 51290 | 199.83 | 26.9865 | 3.695 |
| Forward/backward differencing only, $W_0 = -25$ | | | | | |
| 133 | 100 | 274 | 0.04 | 27.2442 | |
| 265 | 400 | 873 | 0.27 | 27.1997 | |
| 529 | 1600 | 3308 | 1.95 | 27.1158 | 0.5306 |
| 1057 | 6400 | 12917 | 15.28 | 27.0576 | 1.440 |
| 2113 | 25600 | 51283 | 119.69 | 27.0244 | 1.753 |

traditional asset allocation strategy for a defined contribution pension plan is *deterministic lifestyling*. Initially, the contributions of the plan are invested entirely in equities. Beginning on a predetermined date, say $N$ years prior to retirement, the contributions are switched into bonds at a rate of $1/N$ per year. Then, all assets are invested in bonds by the date of retirement. Deterministic lifestyling can reduce the losses of the plan in case of a sudden fall in the stock market just before the date of retirement. This strategy is simple and widely used. However, obviously it is not the optimal strategy.

We will follow the approach given in [15] to find the optimal dynamic asset allocation strategy for a defined contribution pension plan. The objective of the strategy is to maximize the plan member's utility at retirement. It is assumed that the utility is a function of the plan member's wealth to yearly income ratio [15].

We give a brief derivation of the model equations, for details we refer the reader to [15].

**7.1. Stochastic Model.** Suppose there are two underlying assets in the pension plan: one is risk free (e.g. a government bond) and the other is risky (e.g. a stock market fund). The risky asset $S$ follows the stochastic process

$$dS = (r + \xi_1\sigma_1)S\ dt + \sigma_1 S\ dZ_1\ , \tag{7.1}$$

where $dZ_1$ is the increment of a Wiener process, $\sigma_1$ is volatility, $r$ is the interest rate, $\xi_1$ is the market price of risk. Suppose that the plan member continuously pays into the pension plan at a fraction $\pi$ of her yearly salary $Y$, which follows the process

$$dY = (r + \mu_Y)Y\ dt + \sigma_{Y_0} Y\ dZ_0 + \sigma_{Y_1} Y\ dZ_1\ , \tag{7.2}$$

where $\mu_Y$, $\sigma_{Y_0}$ and $\sigma_{Y_1}$ are constants, and $dZ_0$ is another increment of a Wiener process, which is independent of $dZ_1$. Let $W(t)$ denote the wealth accumulated in the pension plan, and let $p$ denote the proportion of this wealth invested in the risky asset $S$, and let $1-p$ denote the fraction of wealth invested in the risk-free asset. Then

$$dW = (r + p\xi_1\sigma_1)W\ dt + p\sigma_1 W dZ_1 + \pi Y dt\ . \tag{7.3}$$

Assume the plan member has a power utility function $u(W(T), Y(T))$ at retirement time $T$, which is defined as a function of the wealth-income ratio,

$$u(W(T), Y(T)) = \begin{cases} \frac{1}{\gamma}(\frac{W(T)}{Y(T)})^\gamma & \text{where } \gamma < 1 \text{ and } \gamma \neq 0 \\ \log(\frac{W(T)}{Y(T)}) & \text{when } \gamma = 0 \ . \end{cases} \tag{7.4}$$

Our goal is to find the optimal asset allocation strategy to maximize the expected terminal utility. Define a new state variable $X(t) = W(t)/Y(t)$, then by Ito's Lemma, we obtain

$$\begin{aligned} dX = [\pi + X(-\mu_Y + p\sigma_1(\xi_1 - \sigma_{Y_1}) + \sigma_{Y_0}^2 + \sigma_{Y_1}^2)]dt \\ -\sigma_{Y_0} X dZ_0 + X(p\sigma_1 - \sigma_{Y_1})dZ_1 \ . \end{aligned} \tag{7.5}$$

Let $J(t, x, p) = E[u(X_p(T))|X(t) = x]$, where $X(t)$ is the path of $X$ given the asset allocation strategy $p = p(t, x)$, and $E[\cdot]$ is the expectation operator. We define

$$V(x, \tau) = \sup_{p \in \hat{P}} E[u(X_p(T))|X(T - \tau) = x] = \sup_{p \in \hat{P}} J(T - \tau, x, p) \ . \tag{7.6}$$

where $\hat{P}$ is the set of all admissible asset allocation strategies, and $\tau = T - t$. Then $V(x, \tau)$ satisfies the HJB equation

$$V_\tau = \sup_{p \in \hat{P}} \{\mu_X^p V_x + \frac{1}{2}(\sigma_X^p)^2 V_{xx}\} \ ; \ x \in [0, \infty] \ , \tag{7.7}$$

with terminal condition

$$V(x, \tau = 0) = \begin{cases} \gamma^{-1} x^\gamma & \text{where } \gamma < 1 \text{ and } \gamma \neq 0 \\ \log(x) & \text{when } \gamma = 0 \ , \end{cases} \tag{7.8}$$

and where

$$\begin{aligned} \mu_X^p &= \pi + x(-\mu_Y + p\sigma_1(\xi_1 - \sigma_{Y_1}) + \sigma_{Y_0}^2 + \sigma_{Y_1}^2) \\ (\sigma_X^p)^2 &= x^2(\sigma_{Y_0}^2 + (p\sigma_1 - \sigma_{Y_1})^2) \ . \end{aligned} \tag{7.9}$$

with boundary conditions

$$V_\tau(x = 0, \tau) = \pi V_x \ ; \ V(x \to \infty, \tau) = 0 \ . \tag{7.10}$$

For computational purposes, we truncate the domain to $[0, x_{\max}]$, and impose the boundary conditions (7.10) on this finite domain. In order to ensure that Assumption 2.1 holds, we define the range of controls to be

$$\hat{P} = [0, p_{\max}] \ . \tag{7.11}$$

Note that in the original problem in [15], $\hat{P} = [0, \infty]$. A value of $p > 1$ indicates that the holder borrows to invest in risky assets. As a practical matter, it is unlikely that anyone could borrow an unlimited amount relative to her wealth to invest in risky assets. We will choose $p_{\max}$ sufficiently large so that the computed solution is insensitive to $p_{\max}$.

The terminal condition (7.4) is undefined for $x = 0$, if, for example, $\gamma < 0$. We adopt the simple expedient of replacing condition (7.4) by

$$V(x, \tau = 0) = \begin{cases} \frac{1}{\gamma} \max(x, \epsilon)^\gamma & \text{if } \gamma < 0 \\ \frac{1}{\gamma} \log(\max(x, \epsilon)) & \text{if } \gamma = 0 \ , \end{cases} \tag{7.12}$$

**Table 7.1** Computational parameters, pension plan. Convergence tolerance used in Algorithm 5.1. $\epsilon$ is used to adjust the terminal condition (7.12). $p_{\max}$ is the maximum value of the equity proportion (7.11). $x_{\max}$ is the maximum $x$ value in the finite computational domain.

| | |
|---|---|
| Convergence Tolerance | $10^{-7}$ |
| $\epsilon$ | $10^{-3}$ |
| $p_{\max}$ | 200 |
| $x_{\max}$ | 80 |

where $\epsilon > 0, \epsilon \ll 1$. We choose $\epsilon$ sufficently small so that the computed results are insensitive to this value. Table 7.1 shows the computational parameters used in our numerical tests.

Note that the HJB equation (7.7) becomes independent of $p$ at $x = 0$. This nonuniqueness of $p$ does not affect the solution value $u$. It will be understood in the following that if we refer to a value of, say, $(p\,x)$ at $x = 0$, then we are really referring to

$$\lim_{x \to 0} (p\,x) \ . \tag{7.13}$$

**7.2. Discretization.** The pension plan asset allocation model is a special case, of the general HJB equation (2.2), if we make the identification

$$\begin{aligned} &Q = (p) \ , \ \ \hat{Q} = [0, p_{\max}] \ , \ \ a(x, \tau, Q) = \frac{1}{2}(\sigma_X^p)^2 \ , \\ &b(x, \tau, Q) = \mu_X^p \ , \ \ c(x, \tau, Q) = 0 \ , \ \ d((x, \tau, Q) = 0 \ , \end{aligned} \tag{7.14}$$

where $Q, \ a, \ b, \ c$ are defined in equation (2.2). Given node $x = x_i$, with specified solution estimate $\hat{V}^k = [\hat{v}_0^k, ..., \hat{v}_{imax}^k]'$, the objective function which is maximized at each node in Algorithm 5.1 is

$$\begin{aligned} [F(q)]_i &= [A(Q)\hat{V}^k + D(Q)]_i \\ &= [\mu_X^p]_i [(\hat{v}^k)_x]_i + \frac{1}{2}([\sigma_X^p]_i)^2 [(\hat{v}^k)_{xx}]_i \ , \end{aligned} \tag{7.15}$$

where $\hat{V}^k$ is the vector containing the current estimate of the discrete solution values. Similar to the passport option case, if we want to apply central differencing as much as possible, Algorithm 6.13 is used to decide which differencing scheme is used (which depends on $Q$).

**7.3. Numerical Results.** Given parameters in Table 7.2, Table 7.3 shows the numerical results. Recall that as we refine the grid, by inserting a fine grid node between two coarse grid nodes, we reduce the timestep size by four. Since fully implicit timestepping is used (which guarantees a monotone scheme), then the ratio of successive changes in the solution, as the grid is refined, should be four for quadratic convergence, and two for linear convergence. As expected, Table 7.3 shows that quadratic convergence is obtained by using central differencing as much as possible, and first order convergence is obtained by using forward/backward differencing. As for the passport option case, convergence of Algorithm 5.1 is rapid.

Numerical tests with the parameters in Table 7.1, indicated that increasing the truncated domain size $x_{\max}$, increasing the maximum value of the control $p_{\max}$, and decreasing the convergence tolerance and $\epsilon$, resulted in no change to the results in Table 7.3 to six figures.

In the passport option case, particularly with a convex payoff, numerical experiments indicate that using central differencing only does converge to the viscosity solution [2, 28]. However, this cannot be guaranteed. In contrast, in the pension plan case, the numerical scheme does not appear to converge at all using central differencing only. In this respect, the pension plan problem appears to be more challenging than passport option valuation.

**Table 7.2** Parameters used in the pension plan examples. The time units in this problem are years, so that the ratio of wealth to salary $x$ has the units of years.

| | | | |
|---|---|---|---|
| $\mu_y$ | 0. | $\xi_1$ | 0.2 |
| $\sigma_1$ | 0.2 | $\sigma_{Y1}$ | 0.05 |
| $\sigma_{Y0}$ | 0.05 | $\pi$ | 0.1 |
| $T$ | 20 years | $\gamma$ | -5 |

**Table 7.3** Convergence study, pension plan example. Fully implicit timestepping is applied, using constant timesteps. On each refinement, new fine grid nodes are inserted between each coarse grid node, and the timestep size is reduced by four. Parameters are given in Table 7.2. The utility values are given at $x = 1$ and $x = 0$. Ratio is the ratio of successive changes in computed solution as the discretization parameters are reduced.

| Nodes | Timesteps | Nonlinear iterations | CPU Time (Sec) | Utility | Ratio |
|---|---|---|---|---|---|
| Central Differencing as much as possible, $x = 0$ | | | | | |
| 87 | 160 | 331 | 0.04 | $-4.06482 \times 10^{-3}$ | |
| 173 | 640 | 1280 | 0.36 | $-3.65131 \times 10^{-3}$ | |
| 345 | 2560 | 5120 | 2.75 | $-3.58063 \times 10^{-3}$ | 5.851 |
| 689 | 10240 | 20480 | 21.31 | $-3.56354 \times 10^{-3}$ | 4.134 |
| 1377 | 40960 | 81920 | 168.07 | $-3.55922 \times 10^{-3}$ | 3.961 |
| Forward/backward differencing only, $x = 0$ | | | | | |
| 87 | 160 | 399 | 0.03 | $-6.73472 \times 10^{-3}$ | |
| 173 | 640 | 1296 | 0.22 | $-4.68055 \times 10^{-3}$ | |
| 345 | 2560 | 5135 | 1.68 | $-4.04828 \times 10^{-3}$ | 3.249 |
| 689 | 10240 | 20480 | 13.06 | $-3.79150 \times 10^{-3}$ | 2.462 |
| 1377 | 40960 | 81920 | 103.09 | $-3.67543 \times 10^{-3}$ | 2.213 |
| Central differencing as much as possible, $x = 1$ | | | | | |
| 87 | 160 | 331 | 0.04 | $-4.68742 \times 10^{-4}$ | |
| 173 | 640 | 1280 | 0.36 | $-4.31528 \times 10^{-4}$ | |
| 345 | 2560 | 5120 | 2.75 | $-4.26814 \times 10^{-4}$ | 7.894 |
| 689 | 10240 | 20480 | 21.31 | $-4.25611 \times 10^{-4}$ | 3.921 |
| 1377 | 40960 | 81920 | 168.07 | $-4.25305 \times 10^{-4}$ | 3.920 |
| Forward/backward differencing only, $x = 1$ | | | | | |
| 87 | 160 | 399 | 0.03 | $-7.14415 \times 10^{-4}$ | |
| 173 | 640 | 1296 | 0.22 | $-5.55931 \times 10^{-4}$ | |
| 345 | 2560 | 5135 | 1.68 | $-4.87660 \times 10^{-4}$ | 2.321 |
| 689 | 10240 | 20480 | 13.06 | $-4.55786 \times 10^{-4}$ | 2.142 |
| 1377 | 40960 | 81920 | 103.09 | $-4.40348 \times 10^{-4}$ | 2.064 |

Given Parameters in Table 7.2, Figure 7.1 shows the expected terminal utility at $t = 0$ and the corresponding optimal asset allocation strategy $p$ as a function of the salary to wealth $x$ ratio. Note that as $x \to 0$, the proportion invested in the risky asset becomes very large. However, as noted in [15], the amount actually invested in the risky asset $(p\, x)$, tends to zero as $x \to 0$. This is clearly illustrated Figure 7.1 (d). The results in Figure 7.1 (d) are similar to the results in [15].

**7.4. Discretization of the Control.** In some cases, if the form of the HJB equation is complex, then it may be difficult to implement Algorithm 6.13. In this case, a simpler

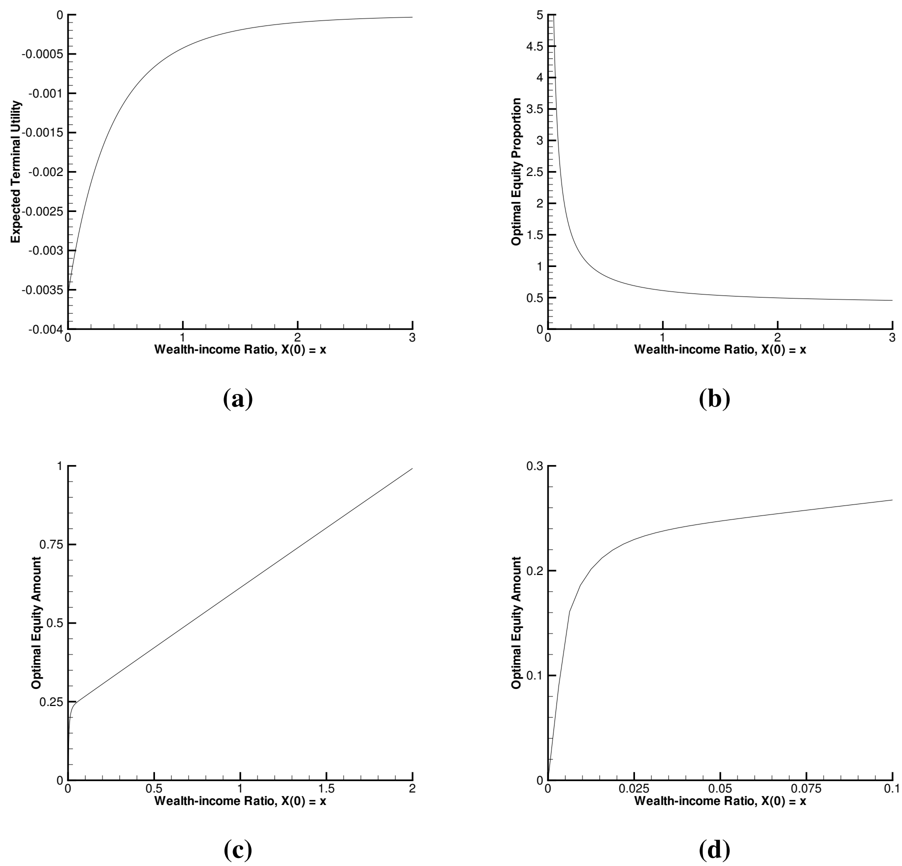



FIG. 7.1: *Utility and optimal asset allocation strategy at $t = 0$, pension plan example. Parameters are given in Table 7.2. (a) Expected terminal utility; (b) Optimal asset allocation strategy; (c) Optimal equity amount $(p\, x)$; (d) Magnified graph of figure (c).*

approach is desirable. Suppose there is one control $q$ at each node, and we discretize the possible control values as described in Section 4.1. From Lemma 4.3, we have that a scheme using discrete controls will converge to the viscosity solution of the original HJB equation. To determine the optimal control at each node, as required in Algorithm 5.1, then we simply perform a linear search of the discrete control values. For a given $q$, we use central weighting if this results in a positive coefficient method, otherwise, forward/backward differencing is used.

Note that since we cannot assume that the objective function is a continuous function of the control, linear search is the only way to find the optimal value of $q$. This method has the obvious advantage that it is very easy to implement, especially in the case where central differencing is used as much as possible.

The numerical results obtained using this method for the pension plan problem are given in Table 7.4. The results are very close to the results reported in Table 7.3. Of course, this method requires much more CPU time compared to Algorithm 6.13. This is simply due to the comparatively crude method used to find the optimal control at each grid node.

In an effort to do better than linear search, we experimented with various approximate methods for finding the optimal control (assuming a discrete set of controls). Seemingly reasonable methods based on smooth approximations to the objective function were very unreliable, and Algorithm 5.1 typically failed to converge. This is simply because the smooth approximation may not maximize the local objective function, and hence the argument used to prove the convergence of the iteration (Theorem 5.2) breaks down.

**Table 7.4** Convergence study, pension plan example, discretized control. $x-$nodes refers to the number of nodes in the $x$ grid. $p-$nodes refers to the number of nodes in the discretization of the range of control values. Fully implicit timestepping is used with constant timesteps. On each refinement, new fine grid nodes are inserted between each coarse grid node, and the timestep is reduced by four. Central differencing is used as much as possible. Ratio is the ratio of successive changes in computed solution as the discretization parameters are reduced. Problem data given in Table 7.2.

| x-Nodes | p-Nodes | Timesteps | Nonlinear iterations | CPU (Sec) | Utility | Ratio |
|---|---|---|---|---|---|---|
| | | | $x = 0$ | | | |
| 173 | 113 | 640 | 1317 | 1.9 | $-3.65307 \times 10^{-3}$ | |
| 345 | 225 | 2560 | 5146 | 29.4 | $-3.58083 \times 10^{-3}$ | |
| 689 | 449 | 10240 | 20511 | 457 | $-3.56358 \times 10^{-3}$ | 4.187 |
| 1377 | 897 | 40961 | 82016 | 7240 | $-3.55923 \times 10^{-3}$ | 3.965 |
| | | | $x = 1.0$ | | | |
| 173 | 113 | 640 | 1317 | 1.9 | $-4.31662 \times 10^{-4}$ | |
| 345 | 225 | 2560 | 5146 | 29.4 | $-4.26845 \times 10^{-4}$ | |
| 689 | 449 | 10240 | 20511 | 457 | $-4.25619 \times 10^{-4}$ | 3.929 |
| 1377 | 897 | 40961 | 82016 | 7240 | $-4.25306 \times 10^{-4}$ | 3.916 |

**8. Conclusions.** Many financial models result in nonlinear HJB PDEs. Classical solutions to these PDEs do not usually exist. In order to ensure convergence to the viscosity solution, monotone difference methods must be used. The standard approach simply uses forward/backward differencing to ensure monotonicity. Clearly, this method suffers from low accuracy.

However, in many financial applications, it is often the case that central differencing can be used at many nodes. This possibility seems to have been ignored in previous work. In this paper, we use central differencing as much as possible. When we use central differencing as much as possible, the local objective function at each node is a discontinuous function of the control. However, the proof of convergence of the iterative scheme for solution of the fully implicit discretized algebraic equations does not require continuity of the local objective function. Hence convergence of the iterative algorithm for solving the nonlinear discretized equations is guaranteed.

We have reported numerical experiments for pricing passport options and optimal asset allocation for defined contribution pension plans. In all cases, use of central differencing as much as possible converges at a higher rate than use of forward/backward differencing only. We have seen this same effect in many numerical experiments. Use of central differencing as much as possible is never worse (in terms of convergence rate) and almost always superior to forward/backward differencing only. Note that these higher rates cannot be guaranteed, but convergence to the viscosity solution is guaranteed.

However, use of central differencing as much as possible is more costly than simply using forward/backward differencing. This is due to the fact that the possible discontinuity points of the objective function must be identified, so that standard methods can be used to maximize the objective function on intervals where the objective function is smooth. This also adds somewhat to implementation complexity.

In cases where it is not possible to analytically maximize the objective function, a very straightforward approach is to discretize the control, and to maximize the objective function using a linear search. This method is very easy to implement, and is also guaranteed to converge to the viscosity solution, but this technique is much more computationally expensive. In these cases, it is clearly advantageous to use central differencing as much as possible, so that higher rates of convergence may compensate for the increased computational cost.

**Appendix A. Discrete Equation Coefficients.** Let $Q_i^n$ denote the vector of optimal controls at node $i$, time level $n$ and set

$$a_i^{n+1} = a(S_i, \tau^n, Q_i^n),\ b_i^{n+1} = b(S_i, \tau^n, Q_i^n),\ \ c_i^{n+1} = c(S_i, \tau^n, Q_i^n)\ . \tag{A.1}$$

Then, we can use central, forward or backward differencing at any node.

Central Differencing:

$$\begin{aligned} \alpha_{i,central}^n &= \left[ \frac{2a_i^n}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} - \frac{b_i^n}{S_{i+1} - S_{i-1}} \right] \\ \beta_{i,central}^n &= \left[ \frac{2a_i^n}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \frac{b_i^n}{S_{i+1} - S_{i-1}} \right]\ . \end{aligned} \tag{A.2}$$

Forward/backward Differencing: ($b_i^n > 0$/ $b_i^n < 0$)

$$\begin{aligned} \alpha_{i,forward/backward}^n &= \left[ \frac{2a_i^n}{(S_i - S_{i-1})(S_{i+1} - S_{i-1})} + \max(0, \frac{-b_i^n}{S_i - S_{i-1}}) \right] \\ \beta_{i,forward/backward}^n &= \left[ \frac{2a_i^n}{(S_{i+1} - S_i)(S_{i+1} - S_{i-1})} + \max(0, \frac{b_i^n}{S_{i+1} - S_i}) \right]\ . \end{aligned} \tag{A.3}$$

REFERENCES


[1] A. L. Amadori, *Nonlinear integro-differential evolution problems arising in option pricing: A viscosity solution approach*, Journal of Differential and Inegral Equations, 16 (2003), pp. 787–811.

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

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

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

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

[6] G. Barles, C. Daher, and M. Romano, *Convergence of numerical shemes for parabolic eqations arising in finance theory*, Mathematical Models and Methods in Applied Sciences, 5 (1995), pp. 125–143.

[7] G. Barles and E. Jakobsen, *Error bounds for monotone approximation schemes for parabolic Hamilton-Jacobi-Bellman equations*. Working Paper, Norwegian Univesity of Science and Technology, 2005.

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

[9] ———, *A strong comparison result for the Bellman equation arising in stochastic exit time control problems and its applications*, Communications in Partial Differential Equations, 23 (1998), pp. 1945–2033.

[10] G. Barles and P. Souganidis, *Convergence of approximation schemes for fully nonlinear equations*, Asymptotic Analysis, 4 (1991), pp. 271–283.

[11] E. Bayraktar and V. Young, *Pricing options in incomplete markets via the instantaneous Sharpe Ratio*. Working paper, University of Michigan, 2005.

[12] F. E. Benth, K. Karlsen, and K. Reikvam, *Optimal portfolio management management rules in a nonguaussian market with durability and intertemporal substitution*, Finance and Stochastics, 5 (2001), pp. 447–467.

[13] Y. Bergman, *Option pricing with differential interest rates*, Review of Financial Studies, 8 (1995), pp. 475–500.

[14] J. Cai and H. Yang, *Ruin in the perturbed compound Poisson risk process under interest force*. Working paper, University of Waterloo, 2005.

[15] A. Cairns, D. Blake, and K. Dowd, *Stochastic lifestyling: optimal dynamic asset allocation for defined contribution pension plans*, Journal of Economic Dynamics and Control, 30 (2006), pp. 843–877.

[16] J. P. Chancellor, B. Oksendal, and A. Sulem, *Combined stochastic control and optimal stopping and application to numerical approximation of combined stochastic and impulse control*, Stochastic Financial Mathematics, 237 (2002), pp. 149–173. Proceedings of the Steklov Institute, Moscow, Editor: A.N. Shiryaev.

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

[18] 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 (1992), pp. 1–67.

[19] P. Forsyth and G. Labahn, *Numerical methods for controlled Hamilton-Jacobi-Bellman pdes in finance*. Working paper, University of Waterloo, 2006.

[20] T. Hyer, A. Lipton-Lifschitz, and D. Pugachevsky, *Passport to success*, Risk Magazine, 10(9) (1997), pp. 127–131.

[21] H. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, Springer-Verlag, New York, 1991.

[22] H. E. Leland, *Option pricing and replication with transaction costs*, Journal of Finance, 40 (1985), pp. 1283–1301.

[23] M. Milevsky, S. Promislow, and V. Young, *Financial valuation of mortality risk via the instantaneous Sharpe ratio*. Working paper, York University, 2005.

[24] M. Mnif and A. Sulem, *Optimal risk control under excess of loss reinsurance*. Working paper, Université Paris 6, 2001.

[25] B. Oksendal and A. Sulem, *Applied Stochastic Control of Jump Diffusions*, Springer Verlag, Berlin, 2005.

[26] H. Pham, *On some recent aspects of stochastic control and their applications*. Working paper, University of Paris 7, 2005.

[27] D. Pooley, P. Forsyth, and K. Vetzal, *Numerical convergence properties of option pricing PDEs with uncertain volatility*, IMA Journal of Numerical Analysis, 23 (2003), pp. 241–267.

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

[29] H. Schmidli, *On minimizing the ruin probability by investment and reinsurance*, Annals of Applied Probability, 12 (2002), pp. 890–907.

[30] S. Shreve and J. Vecer, *Options on a traded account: vacation calls, vacation puts, and passport options*, Finance and Stochastics, 4 (2000), pp. 255–274.

[31] J. Topper, *A finite element implementation of passport options*. MSc. Thesis, Oxford, 2003.

[32] H. Windcliff, J. Wang, P. Forsyth, and K. Vetzal, *Hedging with a correlated asset: Solution of a nonlinear pricing PDE*. Journal of Computational and Applied Mathematics, to appear.