# Continuous Time Mean-Variance Optimal Portfolio Allocation Under Jump Diffusion: An Numerical Impulse Control Approach *

D.M. Dang † P.A. Forsyth ‡

April 1, 2013

**Abstract**

We present efficient partial differential equation (PDE) methods for continuous time mean-variance portfolio allocation problems when the underlying risky asset follows a jump-diffusion. The standard formulation of mean-variance optimal portfolio allocation problems, where the total wealth is the underlying stochastic process, gives rise to a one-dimensional (1-D) non-linear Hamilton-Jacobi- Bellman (HJB) partial integro-differential equation (PIDE) with the control present in the integrand of the jump term, and thus is difficult to solve efficiently. In order to preserve the efficient handling of the jump term, we formulate the asset allocation problem as a 2-D impulse control problem, one dimension for each asset in the portfolio, namely the bond and the stock. We then develop a numerical scheme based on a semi-Lagrangian timestepping method, which we show to be monotone, consistent, and stable. Hence, assuming a strong comparison property holds, the numerical solution is guaranteed to converge to the unique viscosity solution of the corresponding HJB PIDE. The correctness of the proposed numerical framework is verified by numerical examples. We also discuss the effects on the efficient frontier of realistic financial modeling, such as different borrowing and lending interest rates, transaction costs and constraints on the portfolio, such as maximum limits on borrowing and solvency.

**Keywords:** mean-variance, impulse control, HJB equation, finite difference, viscosity solution

**AMS Classification** 65N06, 93C20

## 1 Introduction

In an asset allocation problem, an investor (i) can choose to invest in a risk-free asset, e.g. a bond, or a risky asset, e.g. a stock, and (ii) can dynamically transfer wealth between the two assets, to achieve a pre-determined criteria for the portfolio over a long time horizon, typically 10 years or more. In the mean-variance approach, risk is quantified by variance, so that investors aim to maximize the expected return of their portfolios, given a risk level. Alternatively, they aim to minimize the risk level, given an expected return. As a result, mean-variance strategies are appealing due to their

*This work was supported by Credit Suisse, New York and the Natural Sciences and Engineering Research Council (NSERC) of Canada

†Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada dm2dang@uwaterloo.ca

‡Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1 paforsyt@uwaterloo.ca

intuitive nature, since the results can be easily interpreted in terms of the trade-off between the risk and the expected return.

In the case where the asset follows a pure diffusion process, such as Geometric Brownian Motion (GBM) without jumps, there is considerable literature on the topic. See, for example, [1, 2, 3, 4, 5, 6]. In particular, we note that the optimal strategy adopted in these papers is of the *pre-commitment* type, which is not *time-consistent*, as noted in [7, 8]. A comparison between time-consistent and pre-commitment strategies is given in [9].

Although there is some controversy surrounding pre-commitment strategies, e.g. see [7, 8, 10], it has been shown in [6] that pre-commitment strategies can also be viewed as a target-based optimization which involves minimizing a quadratic loss function. It is suggested in [6] that this is intuitive, adaptable to investor preferences, and is also mean-variance efficient. This view of pre-commitment mean-variance strategies perhaps explains why this criteria has found its way in many insurance applications, where it is natural to consider that the aim of an insurance company is to minimize the risk of a terminal reserve, given an expected terminal reserve constraint [11, 12, 13, 14].

Virtually all the previous work on pre-commitment mean-variance optimal asset allocation has been based on analytic (closed-form) techniques. See, for example, [1, 2, 4, 15]. However, in general, if realistic constraints on portfolio selection are imposed (e.g. no trading if insolvent, maximum borrowing limits), then a fully numerical approach is required. It is important to note that, as shown in [5], in the case where the risky asset follows a GBM without jumps, the inclusion of realistic portfolio constraints has a profound effect on the efficient frontier.

Another modeling deficiency in the previous work on pre-commitment mean-variance optimal asset allocation is the common assumption that the risky asset follows a Geometric Brownian Motion (GBM) without jumps. However, there is increasing empirical evidence that stocks often exhibit jumps. As a result, it is highly desirable to augment the usual GBM with discontinuous jump processes. In this case, the standard formulation of mean-variance optimal asset allocation problems, where the total wealth is the underlying stochastic process (e.g. see [5]), gives rise to a one-dimensional (1-D) non-linear Hamilton-Jacobi- Bellman (HJB) partial integro-differential equation (PIDE) with the control present in the integrand of the jump term. Solving this HJB PIDE is very computationally challenging, since, at each timestep, the presence of the control in the integrand requires the repeated computation of the integral when searching for the control in the control space. In addition, it is not obvious how a fast computational method, such as the FFT, can be utilized for the evaluation of the integral in this case. These shortcomings and challenges motivated our work.

The objective of this article is to develop a fully numerical partial differential equation (PDE) method for solution of the pre-commitment mean-variance portfolio selection problem when the underlying risky asset follows a jump diffusion process. The major contribution of the paper are:

- We formulate the investment problem as the solution to a 2-D impulse control problem, in the form of a non-linear HJB PIDE.

- We include (i) realistic constraints on the portfolio (e.g. maximum limits on borrowing), and (ii) more realistic financial modeling (different interest rates for borrowing and lending, and transaction costs) than previous work.

- We develop a numerical scheme based on a semi-Lagrangian type method, which decouples each PIDE for each discrete value of the riskless asset in the portfolio, and hence, results in solving a a sequence of 1-D non-controlled PIDEs at each timestep. We show that our

numerical scheme is monotone, consistent and $l_\infty$-stable. Assuming a strong comparison property holds, then from [16] and [17], we can be assured that the numerical scheme will converge to the unique viscosity solution of the HJB PIDE.

- Semi-Lagrangian timestepping requires an interpolation at the foot of the characteristic at each timestep. Use of a monotone (linear) interpolation results in poor accuracy for small values of the standard deviation. Use of the exact solution value at a single point dramatically increases accuracy.

- We include several numerical examples, illustrating the convergence of the numerical scheme, as well as the effect of modeling parameters on the efficient frontier.

It is straightforward to incorporate into the numerical scheme developed in this paper additional modeling features, such as non-linear price impact for large transactions. However, we leave this extension to future work.

The remainder of this paper is organized as follows. Section 2 describes the underlying processes and the impulse control framework, and gives a formulation of an associated HJB equation and a linear PIDE . In Section 3, we introduce the concepts of viscosity solution for the HJB equation. Discretization of the relevant equations is given in Section 4. In Section 5, we discuss the convergence of the discrete solutions of the HJB equations to the unique viscosity solution of the HJB PIDE. In Section 6, we highlight some important implementation details of the numerical methods. Numerical results are presented and discussed in Section 7. Section 8 concludes the paper and outlines possible future work.

# 2 Formulation

## 2.1 Underlying Processes

We consider the investor's portfolio to consist of two assets, namely a risky asset and a risk-free asset. We denote by $S$ and $B$ the *amounts* invested in the risky and the risk-free assets, respectively. In general, $S$ would be the amount invested in a broad stock index. For brevity, let $S(t) = S_t$ and $B(t) = B_t$. For use later in the paper, define

$$t^- = t - \epsilon, \quad t^+ = t + \epsilon, \text{ where } \epsilon \to 0^+,$$

i.e. $t^-$ and $t^+$ respectively are instants of time just before and after the (forward) time $t$.

We denote by $\xi$ the random number representing the jump amplitude. We assume that $\xi$ follows a log-normal distribution $p(\xi)$ given by [18]

$$p(\xi) = \frac{1}{\sqrt{2\pi}\zeta\xi} \exp\Big(-\frac{(\log(\xi) - \nu)^2}{2\zeta^2}\Big) \ , \tag{2.1}$$

with parameters $\zeta$ and $\nu$. We have $E[\xi] = \exp(\nu + \zeta^2/2)$, where $E[\cdot]$ denotes the expectation operator. Let $\kappa = E[\xi] - 1$. Under the objective measure, assume that $S$ follows the process

$$\frac{dS_t}{S_{t^-}} \ = (\mu - \lambda\kappa)dt + \sigma dZ + d\bigg(\sum_{i=1}^{\pi_t}(\xi_i - 1)\bigg), \tag{2.2}$$

where $dZ$ is the increment of a Wiener process, $\mu$ is the real world drift rate, and $\sigma$ is the volatility. In addition, $\pi_t$ is a Poisson process with positive intensity parameter $\lambda$, and $\xi_i$ are independent and identically distributed positive random variables having distribution (2.1). When a jump occurs, we have $S_{t^+} = \xi_i S_{t^-}$. To be precise , we consider the process (2.2) to be right continuous with left limits, so that $S_{t^+} = S_t$. However, we will frequently use the notation $S_{t^+}$ and $S_{t^-}$ in the following. This will be especially convenient when considering impulse controls, which can be considered to be left continuous [19].

We assume that the dynamics of the risk-free asset $B$ follow [20]

$$\begin{aligned} dB_t &= \mathcal{R}(B_t)B_t \\ &\mathcal{R}(B_t) = \Big( r_\ell + (r_b - r_\ell)H(B_t) \Big) , \end{aligned} \tag{2.3}$$

where $H(x)$ denotes the Heaviside function

$$H(x) = \begin{cases} 0 & x \geq 0 \\ 1 & x < 0 \end{cases} . \tag{2.4}$$

That is, the investor can be viewed as (i) earning the rate $r_\ell$ for the cash deposit, and (ii) being charged at a rate $r_b > r_\ell$ for borrowing.

In this paper, we assume that $\mu > r_\ell$, hence, it is never optimal (in a mean-variance setting) to short stock. As a result, the amount invested in the risky asset is always nonnegative, i.e. $S_t \geq 0$. However, we allow short positions in the risk-free asset, i.e. it is possible that $B_t < 0$.

## 2.2 Impulse Control

We follow along the lines of [15] and give a brief description of an impulse control problem for dynamic portfolio selection. For more rigor and generality, we refer the reader to [15].

We suppose that, at any time $t$, and any state $(S_t, B_t)$ of the system, the investor can give the system an impulse $\eta \in \mathcal{Z}$, where $\mathcal{Z}$ is the set of admissible impulses. The set of impulse controls for this problem is then the set

$$\mathcal{C} = \{\{t_0, \eta_0\}, \{t_1, \eta_1\}, \ldots, \{t_j, \eta_j\}, \ldots\}_{j \leq M} , \tag{2.5}$$

where $M$ can be finite or infinite, and $t_0 < t_1 < \ldots$. Here, the time $t_j$, $j = 0, \ldots, M$, is referred to as an *intervention* time, with $\eta_j$ being the corresponding *impulse*. Our definitions for the intervention times and impulses for the asset allocation problem are given in Section 2.4. In general, $\mathcal{C} \in \mathcal{A}$, where $\mathcal{A}$ denotes the set of admissible controls.

Given the control $\mathcal{C}$, we denote by $x = (S_t^{\mathcal{C}}, B_t^{\mathcal{C}})$ a controlled state of the system. Between the intervention times $t_{j-1}^+$ and $t_j^-$, $j = 1, \ldots, M$, the controlled system follows the processes (2.2) and (2.3), i.e.

$$\frac{dS_t^{\mathcal{C}}}{S_{t^-}^{\mathcal{C}}} = (\mu - \lambda\kappa)dt + \sigma dZ + d\Bigg( \sum_{i=1}^{\pi_{[t_{j-1}^+, t_j^-]}} (\xi_i - 1) \Bigg) , \tag{2.6}$$

$$dB_t^{\mathcal{C}} = \mathcal{R}(B_t^{\mathcal{C}})B_t^{\mathcal{C}}\ dt \ \ ; \ \ t_{j-1}^+ \leq t \leq t_j^- \ . \tag{2.7}$$

As a result of applying a control $\eta = \eta_j \in \mathcal{Z}$, the state of the system moves instantaneously as follows:

$$x = (S^{\mathcal{C}}_{t^-}, B^{\mathcal{C}}_{t^-}) \quad \rightarrow \quad (S^{\mathcal{C}}_{t^+}, B^{\mathcal{C}}_{t^+}) = (S^{\mathcal{C}}_{t^+}(x,\eta), B^{\mathcal{C}}_{t^+}(x,\eta)). \ , \tag{2.8}$$

In (2.8), we use $(S^{\mathcal{C}}_{t^+}(x,\eta), B^{\mathcal{C}}_{t^+}(x,\eta))$ to emphasize the change in the state of the system after the impulse $\eta = \eta_j$ has been applied.

From now on, to avoid notational clutter, we will generally drop the subscript $t$ and superscript $\mathcal{C}$ from $(S, B)$, with these sub and superscripts understood. We also denote by $x = (s, b) = (S(t^-), B(t^-))$ the state of the system at time $t^-$, and by $(S^+(x,\eta), B^+(x,\eta))$ the state of the system after the impulse $\eta$ has been applied.

## 2.3 Pareto Optimal Points and Efficient Frontier

We denote by $W_{liq}(t) \equiv W_{liq}(S(t), B(t))$, $t \leq T$, the total liquidation value at time $t$ of the investor's portfolio, where $T$ is the time horizon of the investment. Note that $W_{liq}(t)$ may include liquidation costs (see (2.20)). We respectively denote by $E^{x,t}_{\mathcal{C}}[W_{liq}(T)]$ and $Var^{x,t}_{\mathcal{C}}[W_{liq}(T)]$ the expectation and the variance of the terminal liquidation value conditional on the state $(x,t)$ and the impulse control $\mathcal{C}$.

**Definition 2.1.** *We denote by*

$$\mathcal{Y} = \{(Var^{x,t}_{\mathcal{C}}[W_{liq}(T)], E^{x,t}_{\mathcal{C}}[W_{liq}(T)]) : \mathcal{C} \in \mathcal{A}\} \tag{2.9}$$

*the achievable mean-variance objective set, and by $\bar{\mathcal{Y}}$ its closure.*

**Definition 2.2.** *A point $(Var^{x,t}_{\mathcal{C}}[W_{liq}(T)], E^{x,t}_{\mathcal{C}}[W_{liq}(T)]) = (\mathcal{V}, \mathcal{E}) \in \bar{\mathcal{Y}}$ is a Pareto mean-variance optimal point if there exists no admissible impulse control set $\mathcal{C}^* \in \mathcal{A}$ such that*

$$\begin{aligned} Var^{x,t}_{\mathcal{C}^*}[W_{liq}(T)] &\leq \mathcal{V} \ , \\ E^{x,t}_{\mathcal{C}^*}[W_{liq}(T)] &\geq \mathcal{E} \ , \end{aligned} \tag{2.10}$$

*where at least one of the inequalities in equation (2.10) is strict. We denote by $\mathcal{P}$ the set of Pareto mean-variance optimal points. Note that $\mathcal{P} \subseteq \bar{\mathcal{Y}}$.*

Although the above definitions are intuitive, determining the points in $\mathcal{P}$ requires solution of a difficult multi-objective optimization problem, which involves two conflicting criteria. A standard scalarization method can be used to combine the two criteria into an optimization problem with a single objective, from which a point on the efficient frontier can be derived. More specifically, for each point $(\mathcal{V}, \mathcal{E}) \in \bar{\mathcal{Y}}$, and for an arbitrary scaler $\rho > 0$, we first define the set of points $\mathcal{Y}_{P(\rho)}$ to be

$$\mathcal{Y}_{P(\rho)} \quad = \quad \{(\mathcal{V}, \mathcal{E}) \in \bar{\mathcal{Y}} : (\mathcal{V}, \mathcal{E}) = \sup_{(\mathcal{V}_*, \mathcal{E}_*) \in \mathcal{Y}} (\mathcal{E}_* - \rho \mathcal{V}_*)\} \ , \tag{2.11}$$

which involves solving a single-objective optimization problem. We then define the set of points on the efficient frontier, denoted by $\mathcal{Y}_P$, as follows.

**Definition 2.3** (Efficient Frontier)**.** *The set of points on the efficient frontier are defined as*

$$\mathcal{Y}_P = \bigcup_{\rho > 0} \mathcal{Y}_{P(\rho)}. \tag{2.12}$$

**Remark 2.1** (Relationship between $\mathcal{P}$ and $\mathcal{Y}_P$.)**.** *We emphasize the difference between the set of all Pareto mean-variance optimal points $\mathcal{P}$ and the efficient frontier $\mathcal{Y}_P$ defined in equation (2.12). In general, $\mathcal{Y}_P \subseteq \mathcal{P}$. However, the converse may not hold, if the achievable mean-variance objective set $\mathcal{Y}$ is not convex. In this paper, we restrict our attention to determining $\mathcal{Y}_P$.*

As noted in [1, 2], the presence of the variance term in equation (2.11) causes difficulty, if we attempt to determine $\mathcal{Y}_{P(\rho)}$ by solving for the associated value function using dynamic programming. To overcome this difficulty, we make use of the main result in [1, 2, 21] which essentially involves the embedding technique. This main result is summarized in Theorem 2.1.

**Theorem 2.1** (Embedding Result)**.** *Let $\mathcal{Y}$ be a bounded nonempty subset of the set*

$$\{(\mathcal{V},\mathcal{E}) \in \mathbf{R}^2 : \mathcal{V} \geq 0,\ \mathcal{E} \leq C_1\},$$

*where $C_1$ is some positive constant. We define*

$$\mathcal{Y}_{Q(\gamma)} = \{(\mathcal{V},\mathcal{E}) \in \bar{\mathcal{Y}} : \mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E} = \inf_{(\mathcal{V}_*,\mathcal{E}_*)\in\mathcal{Y}} (\mathcal{V}_* + \mathcal{E}_*^2 - \gamma\mathcal{E}_*)\}\ , \tag{2.13}$$

$$\mathcal{Y}_Q = \bigcup_{-\infty\leq\gamma\leq+\infty} \mathcal{Y}_{Q(\gamma)}\ . \tag{2.14}$$

*Then $\mathcal{Y}_P \subseteq \mathcal{Y}_Q$.*

Note that, in Theorem 2.1, the mean and variance $(\mathcal{V},\mathcal{E})$ of $W_{liq}(T)$ are embedded in a scalarization optimization problem with the objective being $\mathcal{V} + \mathcal{E}^2 - \gamma\mathcal{E}$ [21]. Define the value function $\bar{V}(x,t)$ as

$$\bar{V}(x,t) = \inf_{\mathcal{C}\in\mathcal{A}} \left\{ E_{\mathcal{C}}^{x,t}[(W_{liq}(T) - \gamma/2)^2] \right\}\ . \tag{2.15}$$

Theorem 2.1 implies that there exists a $\gamma \equiv \gamma(x,t,\rho)$, such that, for a given positive $\rho$, a control $\mathcal{C}^*$ which maximizes equation (2.11) also minimizes equation (2.15). The benefit of the formulation (2.15) is that dynamic programming can be applied to equation (2.15) to determine the optimal control $\mathcal{C}^*$.

**Remark 2.2** (Construction of efficient frontier)**.** *Our algorithm for determining the points on the efficient frontier is as follows. For a given value of $\gamma$, the optimal strategy $\mathcal{C}^*$ is determined by solving for the value function (2.15). Once this optimal policy $\mathcal{C}^*$ is known, it is then straightforward to determine $(Var_{\mathcal{C}^*}^{x,t}[W_{liq}(T)], E_{\mathcal{C}^*}^{x,t}[W_{liq}(T)])$ and hence, a point on the efficient frontier (see discussions in Subsection 6.2). Repeating this for many values of $\gamma$ traces out a curve in the $(\mathcal{V},\mathcal{E})$ plane. Consequently, the numerical challenge is to solve for the value function (2.15).*

Essentially, the above procedure for constructing the efficient frontier generates points that are in the set $\mathcal{Y}_Q$. As noted in [21], the set $\mathcal{Y}_Q$ may contain spurious points, i.e. points which are not in $\mathcal{Y}_P$. For example, when the controls are not unique, spurious points can be generated. In this case, the set of points in $\mathcal{Y}_Q$ with the spurious points removed generates all the points in $\mathcal{Y}_P$. This is important in the context of a numerical algorithm. An algorithm for removing spurious points is discussed in [21].

**Remark 2.3** (Generality of Theorem 2.1)**.** *Note that it is not guaranteed that problem (2.11) together with (i) the SDEs (2.6-2.7), which can be non-linear, (ii) the impulse (2.8), and (iii) the set of admissible controls $\mathcal{C}$ is a convex optimization problem. See [21] for more discussion of this issue.*

**Remark 2.4** (Time inconsistency of the control)**.** *Although we will use dynamic programming to solve for the optimal control $\mathcal{C}^*$, this control is time inconsistent, as noted in [7, 8], since $\gamma(x,t,\rho)$ depends on the initial state $(x,t)$.*

**Remark 2.5** (Interpretation of 2.15 as a target based optimization)**.** *Equation (2.15) has the obvious interpretation as a target based strategy [6], where the target wealth is $\gamma/2$. This target value can be shown to be [1, 2, 21]*

$$\frac{\gamma}{2} \quad = \quad E^{x,t}_{\mathcal{C}^*}[W_{liq}(T)] + \frac{1}{2\rho^*} \ , \tag{2.16}$$

*where $\mathcal{C}^*$ is control which maximizes equation (2.11) for given $\rho^*$. In the case of a pure diffusion (no transaction costs), for some special cases where analytic solutions are known, it can be shown that the optimal strategy always has $W_{liq}(t)$ less than the discounted value of $\gamma/2$ [6]. As pointed out in [6], we can thus interpret the precommitment mean-variance optimal strategy as a strategy which minimizes the quadratic loss measured relative to the wealth target $\gamma/2$. As long as $W_{liq}(t)$ is less than the discounted target, then equation (2.15) can be interpreted as a quadratic utility. In the case of jumps, again for the special case where analytic solutions are known, the optimal strategy also has $W_{liq}(t)$ less than the discounted target, unless a jump occurs (see Appendix A). This situation is similar to the pure diffusion case if reallocation can only occur at discrete times [10]. In this case, [10] advocates* taking money off the table *to produce a superior efficient frontier. In our tests, we use a jump process where the mean jump size is negative, hence the probability that a jump will exceed the discounted target is extremely small. Alternatively, it is possible to consider an objective function of the form*

$$\inf_{\mathcal{C}\in\mathcal{A}} \left\{ E^{x,t}_{\mathcal{C}}[g(W_{liq}(T))] - \rho Var^{x,t}_{\mathcal{C}}[g(W_{liq}(T))] \right\} \tag{2.17}$$

*where $g(u) = \min(u-L,0)$, i.e. a measure of shortfall for $u < L$. We leave this for future work.*

## 2.4 Intervention Operator

We now give a precise definition the optimal impulse control. The intervention times $t_i$ correspond to the rebalancing times of the portfolio, and the impulse $\eta_i$ corresponds to readjusting the amounts of the stock and bond in the investor's portfolio at time $t_i$. Let $(s,b) = (S(t_i^-), B(t_i^-))$ denote the state of the system at $t_i^-$, and $(S^+(s,b,\eta), B^+(s,b,\eta))$ denote the state after an impulse $\eta$ is applied. More specifically, we assume that fixed and proportional transaction costs, respectively denoted by $c_1 > 0$ and $c_2$, where $c_2 \in [0,1)$, may be imposed on each rebalancing of the portfolio. We then have that

$$\begin{aligned} S^+(s,b,\eta) &= (s+b) - \eta - c_1 - c_2|S^+ - s| \ , \\ B^+(s,b,\eta) &= \eta \ . \end{aligned} \tag{2.18}$$

We now define the intervention operator, denoted by $\mathcal{M}(\eta)\ \bar{V}(s,b,t)$, as

$$\mathcal{M}(\eta)\ \bar{V}(s,b,t) \quad = \quad \bar{V}(S^+(s,b,\eta), B^+(s,b,\eta), t) \ . \tag{2.19}$$

## 2.5 Allowable Portfolios

In general, we assume that trading must cease if the investor is insolvent. To include transaction costs, we define the liquidation value $W_{liq}(s,b)$ to be

$$W_{liq}(s,b) \;=\; b + \max\big[s(1-c_2) - c_1, 0\big] \;. \tag{2.20}$$

As in [22], we assume that shares will be discarded if sale of these shares results in a negative cash flow. We define the solvency region, denoted by $\mathcal{S}$, as

$$\mathcal{S} \;=\; \{(s,b) \in [0,\infty) \times (-\infty,+\infty) : W_{liq}(s,b) > 0\} \;. \tag{2.21}$$

The bankruptcy (insolvency) region, denoted by $\mathcal{B}$, is defined as

$$\mathcal{B} \;=\; \{(s,b) \in [0,\infty) \times (-\infty,+\infty) : W_{liq}(s,b) \leq 0\} \;. \tag{2.22}$$

In the case of a pure diffusion without transaction costs, it is possible to enforce the condition that the stochastic process for the asset value remains in the solvency regions by applying certain boundary conditions to the HJB PDE [5]. However, in the case of a jump process, the asset value may change discontinuously, and move into the insolvency region. This possible movement cannot be prevented by continuous trading, even if the transaction costs are zero. Hence, we must specify the action to be taken in case the process ends up in the bankruptcy region. In the event that insolvency (bankruptcy) occurs, we require that the investor immediately liquidate all investments in the risky asset, and cease trading. That is,

$$S^+ = 0 \quad ; \quad B^+ = W_{liq}(s,b) \quad ; \quad \text{if } (s,b) \in \mathcal{B} \;. \tag{2.23}$$

The investors net debt then accumulates at the borrowing rate.

We will also assume that there is a maximum leverage condition, i.e. the investor must select an asset allocation satisfying

$$\frac{S^+}{S^+ + B^+} < q_{max} \;, \tag{2.24}$$

where $q_{\max}$ is a known positive constant with typical value in $[1.5, 2.0]$. In the event that the asset allocation violates the maximum leverage condition (2.24), we require that the investor choose a different allocation in a region in which (2.24) is satisfied.

## 2.6 Value Function Problem

### 2.6.1 HJB equation formulation

Define $\tau = T - t$, $V(s,b,\tau) = \bar{V}(s,b,t)$, and

$$\begin{aligned} \mathcal{L}V \;&\equiv\; \frac{\sigma^2 s^2}{2} V_{ss} + (\mu - \lambda\kappa) s V_s + \mathcal{R}(b) b V_b - \lambda V \\ \mathcal{J}V \;&\equiv\; \int_0^\infty p(\xi) V(\xi s, b, \tau)\; d\xi \;. \end{aligned} \tag{2.25}$$

Following standard arguments (see [15, 23]), the value function is the viscosity solution of the HJB equation

$$\max\left[V_\tau - \mathcal{L}V - \mathcal{J}V, V - \inf_{\eta\in\mathcal{Z}}(\mathcal{M}(\eta)\ V)\right] = 0 \quad ; \quad \text{if } (s,b)\in\mathcal{S} \tag{2.26}$$

$$\max\left[V_\tau = \mathcal{R}(b)\ bV_b, V - \inf_{\eta\in\mathcal{Z}}(\mathcal{M}(\eta)\ V)\right] = 0 \quad ; \quad \text{if } s = 0 \tag{2.27}$$

$$V(s,b,\tau) = V(0, W_{liq}(s,b),\tau) \quad ; \quad \text{if } (s,b)\in\mathcal{B}\ , \tag{2.28}$$

$$V(s,b,0) = (W_{liq}(s,b) - \gamma/2)^2 \quad ; \quad \text{if } \tau = 0\ , \tag{2.29}$$

defined on the domain $(s,b,\tau)\in\Omega^\infty \equiv [0,\infty)\times(-\infty,+\infty)\times[0,T]$. Equation (2.26) follows from standard arguments. Equation (2.28) is a result of the enforced liquidation if the investor is insolvent. Equation (2.28) can be replaced by a Dirichlet condition

$$V(s,b,\tau) \quad = \quad V(0, W_{liq}(s,b)e^{\mathcal{R}(s+b)\tau}, 0) \quad ; \quad \text{if } (s,b)\in\mathcal{B}\ . \tag{2.30}$$

Equation (2.29) follows from equation (2.15). We can also write equation (2.26) as

$$\max_{\phi\in\{0,1\}}\left[(1-\phi)(V_\tau - \mathcal{L}V - \mathcal{J}V) + \phi(V - \inf_{\eta\in\mathcal{Z}}(\mathcal{M}(\eta)\ V)) = 0\right]\ . \tag{2.31}$$

Thus, the optimal impulse control $\mathcal{C}^*$ for the value function can be represented by the pair $(\phi^*(s,b,\tau), \eta^*(s,b,\tau))$. For consistency of notation, in the insolvent region, i.e. $(s,b)\in\mathcal{B}$, we have that $\phi^*(s,b,\tau)\equiv 1$ and $\eta^*(s,b,\tau)\equiv W_{liq}(s,b)$.

### 2.6.2 Localization

The domain for the value function (2.26-2.29) is $\Omega^\infty$. For computational purposes, we localize this domain to the set of points

$$(s,b,\tau)\in\Omega \quad = \quad [0, s_{\max})\times[-b_{\max}, b_{\max}]\times[0,T]\ , \tag{2.32}$$

where $s_{\max}$ and $b_{\max}$ are positive and sufficiently large. Let $s^* < s_{\max}$. Define the following domains

$$\begin{aligned}
\Omega_{\tau_0} &= [0, s_{\max}]\times[-b_{\max}, b_{\max}]\times\{0\} \\
\Omega_{s^*} &= (s^*, s_{\max}]\times[-b_{\max}, b_{\max}]\times(0,T] \\
\Omega_{s_0} &= \{0\}\times[-b_{\max}, b_{\max}]\times(0,T] \\
\Omega_{\mathcal{B}} &= \{(s,b,\tau)\in\Omega\backslash\Omega_{\tau_0}\backslash\Omega_{s^*}\backslash\Omega_{s_0}\ :\ W_{liq}(s,b)\le 0\} \\
\Omega_{in} &= \Omega\backslash\Omega_{\tau_0}\backslash\Omega_{s^*}\backslash\Omega_{s_0}\backslash\Omega_{\mathcal{B}}\ .
\end{aligned} \tag{2.33}$$

We also define the region

$$\begin{aligned}
\Omega_{b_{\max}} \quad = \quad & (0,s^*]\times[-b_{\max}e^{r_{\max}T}, -b_{\max}) \cup (b_{\max}, b_{\max}e^{r_{\max}T}]\times(0,T]\ , \\
& r_{\max} = \max(r_b, r_\ell)\ .
\end{aligned} \tag{2.34}$$

An illustration of the spatial computational domain is given in Figure 2.1. We emphasize that we do not actually solve the HJB equation in $\Omega_{b_{\max}}$. However, we may use an approximate value to the

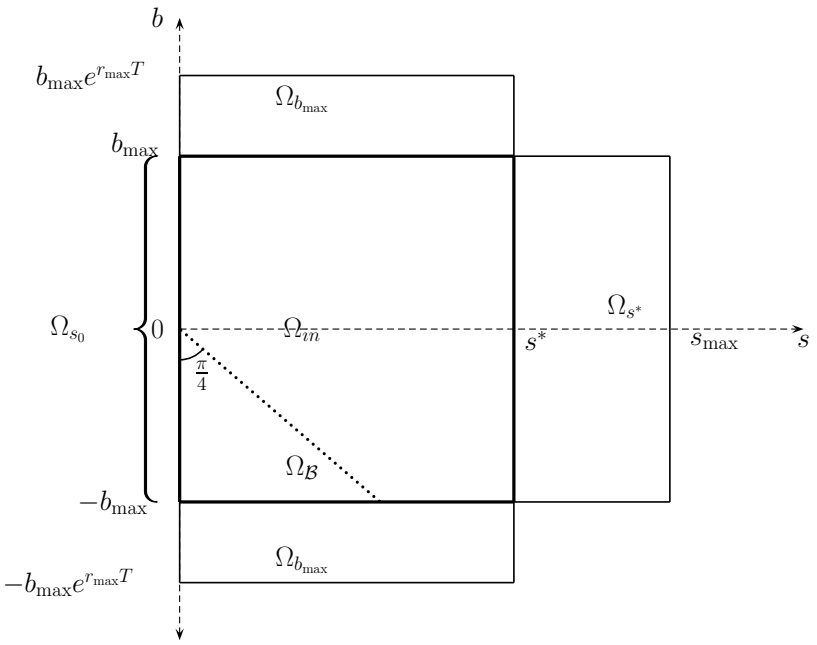


FIGURE 2.1: *Spatial computational domain at each timestep. For ease of exposition, we have illustrated the domain for the case where the transaction cost parameters in equation (2.18) are* $c_1 = c_2 = 0$.

solution in $\Omega_{b_{\max}}$, obtained by means of extrapolation of the computed solution in $\Omega_{in}$, to provide any information required by the HJB PDE in $\Omega$.

We now describe the equation for the localized domains defined in (2.33) and (2.34). From equation (2.29), we have that, for fixed $b$, $V(s \to \infty, b, 0) \simeq (1 - c_2)^2 s^2$. Now, given the PIDE

$$\hat{V}_\tau - \mathcal{L}\hat{V} - \mathcal{J}\hat{V} = 0 \ , \tag{2.35}$$

with $\mathcal{L}, \mathcal{J}$ defined in equation (2.25), making the assumption that $\hat{V}(s, b, \tau) \simeq A(\tau)s^2$, when $s \to \infty$, for some unknown function $A(\tau)$, and substituting this asymptotic form into the PIDE (2.35) gives

$$\begin{aligned} \hat{V}_\tau &= (\sigma^2 + 2\mu + \lambda\kappa_2)\hat{V} \ ; \ s \to \infty \\ &\kappa_2 = E[(J-1)^2] \ . \end{aligned} \tag{2.36}$$

We assume that $s^*$ is selected sufficiently large so that $V(s, b, \tau) \simeq A(\tau)s^2$ in $\Omega_{s^*}$. Using equation (2.36), we can approximate the solution in the domain $\Omega_{s^*}$ by

$$\max\left[ V_\tau - (\sigma^2 + 2\mu + \lambda\kappa_2)V, V - \left\{ \inf_{\substack{\eta \in \mathcal{Z} \\ (S^+, B^+) \in \Omega}} \left( \mathcal{M}(\eta)V(s, b, \tau) \right) \right\} \right] = 0 \ . \tag{2.37}$$

In view of the finite range of $s$, we replace $\mathcal{J}$ in equation (2.26) by the localized operator $\mathcal{J}_\ell$

$$\mathcal{J}_\ell V = \int_0^{s_{\max}/s} p(\xi) V(\xi s, b, \tau) \ d\xi \ . \tag{2.38}$$

Some guidelines for choosing $s^*, s_{\max}$ which minimize the effect of the localization error for the jump terms can be found in [24].

Given the initial condition (2.29), we assume that, for fixed $s$, $V(s,b,\tau) \simeq C(s,\tau)b^2$, when $|b| \to \infty$. Alternatively, we can write this assumption as

$$V(s,|b|>|b_{\max}|,\tau) = \left(\frac{b}{b_{\max}}\right)^2 V(s,\mathrm{sgn}(b)b_{\max},\tau) \;;\; (s,b,\tau) \in \Omega_{b_{\max}} \;. \tag{2.39}$$

With this assumption, we could replace the term $bV_b$ in $\mathcal{L}V$ by $2V$ at $b = \pm b_{\max}$. However, we find it conceptually clearer to define the solution as in equation (2.39) for $(s,b,\tau) \in \Omega_{b_{\max}}$. Putting this all together gives us the following complete localized problem :

$$\begin{aligned}
&\max\left[V_\tau - \mathcal{L}V - \mathcal{J}_\ell V, V - \inf_{\substack{\eta\in\mathcal{Z}\\(S^+,B^+)\in\Omega}} (\mathcal{M}(\eta)\ V)\right] = 0 && ;\ (s,b,\tau)\in\Omega_{in}\\
&\max\left[V_\tau - (\sigma^2+2\mu+\lambda\kappa_2)V, V - \inf_{\substack{\eta\in\mathcal{Z}\\(S^+,B^+)\in\Omega}} \big(\mathcal{M}(\eta)V\big)\right] = 0 && ;\ (s,b,\tau)\in\Omega_{s^*}\\
&\max\left[V_\tau - \mathcal{R}(b)\ bV_b, V - \inf_{\substack{\eta\in\mathcal{Z}\\(S^+,B^+)\in\Omega}} (\mathcal{M}(\eta)\ V)\right] = 0 && ;\ (s,b,\tau)\in\Omega_{s_0}\\
&V(s,b,\tau) - V(0,W_{liq}(s,b),\tau) = 0 && ;\ (s,b,\tau)\in\Omega_{\mathcal{B}}\\
&V(s,b,0) - (W_{liq}(s,b)-\gamma/2)^2 = 0 && ;\ (s,b,\tau)\in\Omega_{\tau_0}\\
&V - \left(\tfrac{b}{b_{\max}}\right)^2 V(s,\mathrm{sgn}(b)b_{\max},\tau) = 0 && ;\ (s,b,\tau)\in\Omega_{b_{\max}}
\end{aligned} \tag{2.40}$$

The localized equations in the domains $\Omega_{b_{\max}}, \Omega_{s^*}$ are clearly approximations. However, the errors in regions of interest are expected be small, if $s_{\max}, (s_{\max}-s^*)$, and $b_{\max}$ are sufficiently large. We verify this in some numerical experiments in Section 7.

## 2.7 Expected Value Problem

### 2.7.1 PDE Formulation

Given the solution for the value function (2.15), with the optimal control $\mathcal{C}^* = (\phi^*(s,b,\tau), \eta^*(s,b,\tau))$, it is also desirable to determine the quantity $\bar{U}(x,t)$ defined as

$$\bar{U}(x,t) = E^{x,t}_{\mathcal{C}^*}[W_{liq}(T)] \;, \tag{2.41}$$

since this information is required in order to determine the corresponding point on the efficient frontier.

Let $\tau = T-t$, $U(s,b,\tau) = \bar{U}(s,b,T-\tau)$. Using standard arguments in [15, 23], the linear PDE satisfied by $U(s,b,\tau)$ in the domain $(s,b,\tau) \in [0,\infty)\times(-\infty,+\infty)\times[0,T]$ can be described

by

$$
\begin{aligned}
&(1-\phi^*)(U_\tau - \mathcal{L}U - \mathcal{J}U) + \phi^*(U - (\mathcal{M}(\eta^*)\ U)) = 0 \quad ; \quad \text{if } (s,b) \in \mathcal{S}\ , && (2.42)\\
&(1-\phi^*)(U_\tau - \mathcal{R}(b)\ bU_b) + \phi^*(U - (\mathcal{M}(\eta^*)\ U)) = 0 \quad ; \quad \text{if } s = 0\ , && (2.43)\\
&U(s,b,\tau) = U(0, W_{liq}(s,b), \tau) \quad ; \quad \text{if } (s,b) \in \mathcal{B}\ , && (2.44)\\
&U(s,b,0) = W_{liq}(s,b) \quad ; \quad \text{if } \tau = 0\ , && (2.45)
\end{aligned}
$$

where $(\phi^*, \eta^*)$ given from the solution to equation (2.31).

### 2.7.2 Localization

From the initial condition (2.45), we make the assumptions that $U(s,b,\tau) \simeq A'(\tau)s$ in $\Omega_{s^*}$ , and that

$$
U(s, |b| > |b_{\max}|, \tau) = \left(\frac{b}{b_{\max}}\right) U(s, \mathrm{sgn}(b) b_{\max}, \tau)\ ; (s,b,\tau) \in \Omega_{b_{\max}}\ . \tag{2.46}
$$

Following similar reasoning used to derive equation (2.40), we obtain

$$
\begin{aligned}
&(1-\phi^*)(U_\tau - \mathcal{L}U - \mathcal{J}_\ell U) + \phi^*(V - \mathcal{M}(\eta^*)\ V) = 0 \quad ; \quad (s,b,\tau) \in \Omega_{in}\ ,\\
&(1-\phi^*)(U_\tau - \mu U) + \phi^*(U - \mathcal{M}(\eta) U) \quad ; \quad (s,b,\tau) \in \Omega_{s^*}\ ,\\
&(1-\phi^*)(U_\tau - \mathcal{R}(b)\ bU_b) + \phi^*(V - \mathcal{M}(\eta^*)\ V) = 0 \quad ; \quad (s,b,\tau) \in \Omega_{s_0}\ ,\\
&U(s,b,\tau) - U(0, W_{liq}(s,b), \tau) = 0 \quad ; \quad (s,b,\tau) \in \Omega_{\mathcal{B}}\ ,\\
&U - W_{liq}(s,b) = 0 \quad ; \quad (s,b,\tau) \in \Omega_{\tau_0}\ ,\\
&U - \left(\tfrac{b}{b_{\max}}\right) U(s, \mathrm{sgn}(b) b_{\max}, \tau) \quad ; \quad (s,b,\tau) \in \Omega_{b_{\max}}\ .
\end{aligned} \tag{2.47}
$$

Again, we remind the reader that we only solve the PDE in $\Omega$. The values in $\Omega_{b_{\max}}$, obtained by means of extrapolation of the computed solution in $\Omega_{in}$, are only used if required by the PDE in $\Omega$.

# 3 Value Function: Compact Representation and Viscosity Solution

## 3.1 Compact Presentation

In general, we cannot expect solutions to HJB equations of the form (2.40) to be sufficiently smooth. Hence, we seek the viscosity solution of equations (2.40). To make the statement of the problem more precise in the context of viscosity solutions, we now write the localized problem for the value function, i.e. equations (2.40), in a compact form, which includes the terminal and boundary equations in a single equation. To this end, define $\mathbf{x} = (s,b,\tau)$, and let $DV(\mathbf{x}) = (V_s, V_b, V_\tau)$ and $D^2V(\mathbf{x}) = V_{ss}$. In addition, let $\mathbf{x}^+ = (S^+(s,b), B^+(s,b), \tau)$, and

$$
\mathcal{M}V(\mathbf{x}) \quad = \quad \inf_{\substack{\eta \in \mathcal{Z} \\ \mathbf{x}^+ \in \Omega}} (\mathcal{M}(\eta)\ V(\mathbf{x}))\ . \tag{3.1}
$$

We then write equations (2.40) as

$$
FV \equiv F(\mathbf{x}, V(\mathbf{x}), DV(\mathbf{x}), D^2V(\mathbf{x}), \mathcal{M}V(\mathbf{x}), \mathcal{J}_\ell V(\mathbf{x})) = 0\ , \tag{3.2}
$$

where the operator $FV$ is defined by

$$FV = \begin{cases} F_{\text{in}}V \equiv F_{\text{in}}(\mathbf{x}, V(\mathbf{x}), DV(\mathbf{x}), D^2V(\mathbf{x}), \mathcal{M}V(\mathbf{x}), \mathcal{J}_\ell V(\mathbf{x})), & \mathbf{x} \in \Omega_{in}, \\ F_{s^*}V \equiv F_{s^*}(\mathbf{x}, V(\mathbf{x}), DV(\mathbf{x}), \mathcal{M}V(\mathbf{x})), & \mathbf{x} \in \Omega_{s^*}, \\ F_{s_0}V \equiv F_{s_0}(\mathbf{x}, V(\mathbf{x}), DV(\mathbf{x}), \mathcal{M}V(\mathbf{x})), & \mathbf{x} \in \Omega_{s_0}, \\ F_{\mathcal{B}} \equiv F_{\mathcal{B}}(\mathbf{x}, V(\mathbf{x})), & \mathbf{x} \in \Omega_{\mathcal{B}}, \\ F_{\tau_0}V \equiv F_{\tau_0}(\mathbf{x}, V(\mathbf{x})), & \mathbf{x} \in \Omega_{\tau_0}, \\ F_{b_{\max}}V \equiv F_{b_{\max}}(\mathbf{x}, V(\mathbf{x}), DV(\mathbf{x}), D^2V(\mathbf{x}), \mathcal{M}V(\mathbf{x}), \mathcal{J}_\ell V(\mathbf{x})), & \mathbf{x} \in \Omega_{b_{\max}} \ . \end{cases} \tag{3.3}$$

Here,

$$F_{in}V = \max\left[V_\tau - \mathcal{L}V - \mathcal{J}_\ell V, V - \mathcal{M}V\right], \tag{3.4}$$

$$F_{s^*}V = \max\left[V_\tau - (\sigma^2 + 2\mu + \lambda\kappa_2)V, V - \mathcal{M}V\right], \tag{3.5}$$

$$F_{s_0}V = \max\left[V_\tau - \mathcal{R}(b)\ bV_b, V - \mathcal{M}V\right], \tag{3.6}$$

$$F_{\mathcal{B}}V = V - V(0, W_{liq}(s,b), \tau), \tag{3.7}$$

$$F_{\tau_0}V = V - \left(W_{liq}(s,b) - \gamma/2\right)^2 \ , \tag{3.8}$$

$$F_{b_{\max}}V = V - \left(\frac{b}{b_{\max}}\right)^2 V(s, \text{sgn}(b)b_{\max}, \tau) \ . \tag{3.9}$$

**Definition 3.1** (Value Function Problem)**.** *The HJB equation for the value function (2.15) on the localized domain $\Omega \cup \Omega_{b_{\max}}$ is given by*

$$F(\mathbf{x}, V(\mathbf{x}), DV(\mathbf{x}), D^2V(\mathbf{x}), \mathcal{M}V(\mathbf{x}), \mathcal{J}_\ell V(\mathbf{x})) = 0 \ . \tag{3.10}$$

*Equation (3.10) includes the HJB PDE in the interior and all the boundary conditions in equation (2.40).*

## 3.2 Viscosity Solution

Before defining the viscosity solution of equation (3.10), we first recall the definitions of upper and lower semi-continuous envelopes. Given a function $f : \bar{\Omega} \to \mathbb{R}$, $\bar{\Omega} \subseteq \mathbb{R}^n$, the upper semi-continuous envelope of $f$, denoted by $f^*$, is defined as

$$f^*(\bar{x}) = \lim_{\bar{r} \to 0^+} \sup\left\{ f(y) \ \middle| \ y \in \bar{B}(\bar{x}, \bar{r}) \cap \bar{\Omega} \right\} \tag{3.11}$$

where $\bar{B}(\bar{x}, \bar{r}) = \{y \in \mathbb{R}^n \ \middle| \ |\bar{x} - y| < \bar{r}\}$. We also have the obvious definition for a lower semi-continuous envelope $f_*(\bar{x})$.

We also define

$$\limsup_{y \to \bar{x}} f(\bar{x}) = \lim_{\bar{r} \to 0^+} \sup\left\{ f(y) \ \middle| \ y \in \bar{B}(\bar{x}, \bar{r}) \cap \bar{\Omega} - \{\bar{x}\} \right\} \ , \tag{3.12}$$

with the corresponding definition of $\liminf$.

**Definition 3.2** (Viscosity solution of equation (3.10))**.** *A locally bounded function $V : \Omega \cup \Omega_{b_{\max}} \to \mathbb{R}$ is a viscosity sub-solution (resp. super-solution) of PDE (3.10) if, for all test functions $\phi(\mathbf{x}) \in C^\infty(\Omega \cup \Omega_{b_{\max}})$, and all $\mathbf{x}$, such that $V - \phi$ has a strict global maximum (resp. minimum) with $\phi(\mathbf{x}) = V^*(\mathbf{x})$ (resp. $V_*(\mathbf{x})$), we have*

$$F_*(\mathbf{x}, \phi(\mathbf{x}), D\phi(\mathbf{x}), D^2\phi(\mathbf{x}), \mathcal{M}\phi(\mathbf{x}), \mathcal{J}_\ell\phi(\mathbf{x})) \leq 0 \ , \tag{3.13}$$

*(resp.*

$$F^*(\mathbf{x}, \phi(\mathbf{x}), D\phi(\mathbf{x}), D^2\phi(\mathbf{x}), \mathcal{M}\phi(\mathbf{x}), \mathcal{J}_\ell\phi(\mathbf{x})) \geq 0 \quad ) \ . \tag{3.14}$$

*$V$ is a viscosity solution if it is both a viscosity sub-solution and a viscosity super-solution.*

**Remark 3.1** (Equivalent definitions: viscosity solutions)**.** *There are many equivalent definitions of viscosity solutions. For example, one can replace $\phi(\mathbf{x}) \in C^\infty(\Omega \cup \Omega_{b_{\max}})$ by $\phi(\mathbf{x}) \in C^2(\Omega \cup \Omega_{b_{\max}})$ [25]. It is also possible to replace $\phi(\mathbf{x})$ by $V^*(\mathbf{x})$ (resp. $V_*(\mathbf{x})$) in the non-local terms $\mathcal{J}_\ell\phi(\mathbf{x})$ and $\mathcal{M}\phi(\mathbf{x})$ [26]. This is possible, since these terms contain no derivatives. However, for the purposes of verifying consistency of a numerical scheme, it is convenient to use Definition 3.2. Note that $F(\cdot)$ is proper and degenerate elliptic [27].*

We make the following assumption.

**Assumption 3.1** (Strong Comparison)**.** *The value function as given in Definition 3.1 satisfies a strong comparison result in $\Omega_{in} \cup \Gamma$, where $\Gamma \subseteq \partial\Omega_{in}$. Hence, a unique continuous viscosity solution exists in $\Omega_{in} \cup \Gamma$.*

**Remark 3.2.** *Strong comparison has been proven for similar impulse control problems in [17]. However, some of the assumptions in [17] do not appear to hold for our particular problem. In general, the viscosity solution can be discontinuous on parts of the boundary $\Gamma$. Note that the precise specification of $\Gamma$ has virtually no impact on a computational algorithm. The boundary data is either used or is irrelevant. In all cases, we consider the computed solution as the limiting value approaching $\partial\Omega_{in}$ from the interior.*

# 4 Discretization

## 4.1 Computational grid

We discretize our problem on the localized domain $\Omega$. Define a set of nodes in the $s$-direction by $\{s_1, s_2, \ldots, s_{i_{\max}}\}$, and in the $b$-direction $\{b_1, \ldots, b_{j_{\max}}\}$. Denote the $n^{th}$ discrete timestep by $\tau^n$. For ease of notation, in the following, we assume constant timestep sizes, i.e. $\Delta\tau = \tau^{n+1} - \tau^n$ is constant. However, the actual implementation could make use of variable timestep sizes. The nodes in the $s$- and $b$-directions are not necessarily equally spaced.

Let $\Delta s_{\max} = \max_i(s_{i+1} - s_i)$, $\Delta b_{\max} = \max_j(b_{j+1} - b_j)$, $\Delta\tau_{\max} = \max_n(\tau^{n+1} - \tau^n)$. In addition, we suppose that the control $\eta$ in equation (2.18) is discretized so that $\eta_j = b_j$, with $\Delta\eta_{\max} = \max_j(\eta_{j+1} - \eta_j) = \Delta b_{\max}$. We assume that there is a positive discretization parameter $h$ such that

$$\Delta s_{\max} = C_1 h \ ; \ \Delta b_{\max} = C_2 h \ ; \ \Delta\tau_{\max} = C_3 h \ ; \ \Delta\eta_{\max} = C_4 h \ , \tag{4.1}$$

where $C_p, p = 1, \ldots, 4$, are positive and independent of $h$.

We denote by $V(s_i, b_j, \tau^n)$ the exact solution to the non-linear value equation (3.10) evaluated at the reference node $(s_i, b_j, \tau^n)$, and by $V_h(s, b, \tau)$ the approximate solution at the point $(s, b, \tau)$ obtained using the discretization parameter $h$. We refer to the approximate solution at the reference node $(s_i, b_j, \tau^n)$ as $V^n_{i,j} \equiv V_h(s_i, b_j, \tau^n)$. In the event that we need to evaluate $V_h$ at a point other than nodal values, linear interpolation is used. Similarly, $U_h(s_i, b_j, \tau^n) = U^n_{i,j}$ is the approximation to $U(s_i, b_j, \tau^n)$, which is the solution to the linear expected value equation (2.47). Let $N = i_{\max} \times j_{\max}$ be the number of nodes in the computational grid, and let $V^n$ be the $N$ length vector at time $\tau^n$, i.e.

$$V^n = [V^n_{1,1}, \ldots, V^n_{i_{\max},1}, \ldots, V^n_{1,j_{\max}}, \ldots, V^n_{i_{\max},j_{\max}}]' \ , \tag{4.2}$$

with a similar definition of $U^n$.

We denote by $\mathcal{Z}_h$ the discrete set of admissible controls

$$\mathcal{Z}_h \quad = \quad \{b_1, \ldots, b_{j_{\max}}\} \cap \mathcal{Z} \ . \tag{4.3}$$

We determine the infimum of the intervention operator by a linear search over the discrete set of controls (4.3). Using this approach, we can guarantee convergence to the viscosity solution as $h \to 0$. The obvious alternative is the use of a 1-D optimization algorithm. However, this alternative cannot guarantee convergence to the global minimum. More specifically, in numerical experiments with this alternative approach, we have seen convergence to local minima, and hence, non-convergence to the viscosity solution.

## 4.2 Discretization

As example, we give the details of the discretization for $F_{in}$ in $\Omega_{in}$ as given in equation (3.4). The derivation of the discretizations of $F$ in the remaining sub-domains of $\Omega$, as well as the discretization of the expected value equation for $U$ defined in (2.47), is similar, and hence, is omitted.

Let

$$\mathcal{P}V = \frac{\sigma^2 s^2}{2} V_{ss} + (\mu - \lambda\kappa) s V_s - \lambda V \ , \tag{4.4}$$

and recall that

$$\mathcal{L}V = \mathcal{P}V + \mathcal{R}(b) b V_b \ . \tag{4.5}$$

We denote by $\mathcal{P}_h$ the discrete approximation to $\mathcal{P}$. For $\mathcal{P}_h$, we use the standard three point approximations to the derivatives in equation (2.25), with central, forward and backward differencing. Central differencing is used as much as possible, but we require that the scheme be a positive coefficient method [28]:

$$\begin{aligned} \mathcal{P}_h V^n_{i,j} \quad &= \quad \alpha_{i,j} V_{i-1,j} + \beta_{i,j} V^n_{i+1,j} - (\alpha_{i,j} + \beta_{i,j} + \lambda) V^n_{i,j} \\ &\qquad \alpha_{i,j} \geq 0 \ ; \ \beta_{i,j} \geq 0 \ . \end{aligned} \tag{4.6}$$

Define $(\mathcal{J}_\ell)_h$ to be the discrete form of the localized jump operator (2.38). We use a midpoint rule to approximate this integral, followed by a linear interpolation onto an equally spaced grid. This

facilitates use of an FFT to evaluate the integral [24]. This results in a discretization of the form

$$(\mathcal{J}_\ell)_h V^n_{i,j} = \sum_k q^{i,j}_k V^n_{k,j} \tag{4.7}$$

$$0 \le q^{i,j}_k \le 1 \; ; \; \sum_k q^{i,j}_k \le 1 \; . \tag{4.8}$$

Equations (4.8) hold since $p(\xi)$ defined in (2.1) is a probability density. For details regarding the discretization of the jump term, we refer the reader to [24].

The term $\mathcal{R}(b)V_b$ in equation (4.5) is handled by a semi-Lagrangian timestepping scheme, details of which can be found in Appendix B. Using the implicit timestepping method, our discretization for equation (3.4) at reference point $\mathbf{x}^{n+1}_{i,j} = (s_i, b_j, \tau^{n+1}) \in \Omega_{in}$, is given by

$$\begin{aligned}
&\frac{V^{n+1}_{i,j}}{\Delta\tau} - P_h V^{n+1}_{i,j} - (\mathcal{J}_\ell)_h V^{n+1}_{i,j} = \frac{\widetilde{V}^n_{i,j}}{\Delta\tau} \\
&\widetilde{V}^n_{i,j} = \left( \min \left[ V_h(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, \tau^n), \min_{\substack{B^+ \in \mathcal{Z}_h \\ (S^+, B^+) \in \Omega}} V_h(S^+(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, B^+), B^+, \tau^n) \right] \right) \; .
\end{aligned} \tag{4.9}$$

It is important to note that the semi-Lagrangian timestepping decouples the PIDE $V_\tau = \mathcal{L}V + \mathcal{J}_\ell V$ for each $b_j$ value, $j = 1, \ldots, j_{max}$. More specifically, for a fixed $j$, once the quantity $\widetilde{V}^n_{i,j}$, $i = 1, \ldots, i_{max}$, is computed, the discretized equations for the PIDE corresponding to $b_j$ can then be obtained from (4.9), and can be solved independently from those of other non-controlled PIDEs. Note that, equation (4.9) is the form actually used in the computation, and has a simple intuitive interpretation. An intuitive derivation of (4.9) is presented in Appendix B. In Section 5.1, we prove that discretization (4.9) is a consistent approximation to equation (3.4).

**Remark 4.1** (Solution of the discretized equations)**.** *Note that the semi-Lagrangian discretization* (4.9) *requires only the solution of local non-linear optimization problems and solution of linear equations at each step. In order to avoid a dense matrix solve (due to the presence of the jump term) we use a fixed-point iteration to solve the discrete equations, details of which can be found in [24]. Regarding the the convergence of the fixed-point iteration, since*
*1.* $\alpha_{i,j} \ge 0$ *and* $\beta_{i,j} \ge 0$ *(see 4.6),*
*2.* $0 \le q^{i,j}_k \le 1$*, and* $\sum_k q^{i,j}_k \le 1$*, (see 4.8)*
*3. the weights for linear interpolation are in* $[0, 1]$*,*
*4.* $\mu > 0$ *and* $\lambda > 0$*,*
*the fixed point iteration is guaranteed to converge. For proof details, see [24].*

The dense matrix-vector product (arising from the jump term) is computed in $O(1/h^2 |\log h|)$ operations using an FFT [24].

# 5 Value Function: Convergence to the Viscosity Solution

## 5.1 Consistency

While equation (4.9) is convenient for computation, it is not in a form amenable for analysis. For purposes of proving consistency, it is more convenient to rewrite equation (4.9) in an equivalent

form. Let $\mathcal{G}(\cdot)$ be the discrete approximation to $F_{in}$ for $\mathbf{x} \in \Omega_{in}$. Let $\mathbf{x}_{i,j}^{n+1} = (s_i, b_j, \tau^{n+1})$. We rearrange equation (4.9) so that our formal discretization of $F_{in}$ is

$$\begin{aligned}
&\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, V_{i,j}^{n+1}, \left\{V_{a,b}^{n+1}\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{V_{k,\ell}^{n}\right\}\right) \\
&= \max\Bigg[\frac{V_{i,j}^{n+1} - V_h(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, \tau^n)}{\Delta\tau} - P_h V_{i,j}^{n+1} - (\mathcal{J}_\ell)_h V_{i,j}^{n+1}, \\
&V_{i,j}^{n+1} - \min_{\substack{B^+\in\mathcal{Z}_h \\ (S^+,B^+)\in\Omega}} V_h(S^+(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, B^+), B^+, \tau^n) - \Delta\tau P_h V_{i,j}^{n+1} - \Delta\tau(\mathcal{J}_\ell)_h V_{i,j}^{n+1}\Bigg] \\
&= 0 \quad ; \quad \mathbf{x}_{i,j}^{n+1} \in \Omega_{in} \quad . \qquad (5.1)
\end{aligned}$$

It is easily seen that a solution of equation (5.1) is a solution of equation (4.9). For $\mathbf{x} \in \Omega_{s^*}$, our formal discretization of $F_{s^*}$ is given by

$$\begin{aligned}
&\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, V_{i,j}^{n+1}, \left\{V_{a,b}^{n+1}\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{V_{k,\ell}^{n}\right\}\right) \\
= \;& \max\Bigg[\frac{V_{i,j}^{n+1} - V_h(s_i, b_j), \tau^n)}{\Delta\tau} - (\sigma^2 + 2\mu + \lambda\kappa_2) V_{i,j}^{n+1}, \\
& V_{i,j}^{n+1} - \min_{\substack{B^+\in\mathcal{Z}_h \\ (S^+,B^+)\in\Omega}} V_h(S^+(s_i, b_j, B^+), B^+, \tau^n) - \Delta\tau(\sigma^2 + 2\mu + \lambda\kappa_2) V_{i,j}^{n+1}\Bigg] \\
& = 0 \quad ; \quad \mathbf{x}_{i,j}^{n+1} \in \Omega_{s^*} \quad . \qquad (5.2)
\end{aligned}$$

For $\mathbf{x} \in \Omega_{s_0}$, we approximate $F_{s_0}$ by

$$\begin{aligned}
&\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, V_{i,j}^{n+1}, \left\{V_{a,b}^{n+1}\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{V_{k,\ell}^{n}\right\}\right) \\
= \;& \max\left[\frac{V_{i,j}^{n+1} - V_h(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, \tau^n)}{\Delta\tau}, V_{i,j}^{n+1} - \min_{\substack{B^+\in\mathcal{Z}_h \\ (S^+,B^+)\in\Omega}} V_h(S^+(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, B^+), B^+, \tau^n)\right] \\
& = 0 \quad ; \quad \mathbf{x}_{i,j}^{n+1} \in \Omega_{s_0} \quad . \qquad (5.3)
\end{aligned}$$

Finally, we have

$$\mathcal{G}(\cdot) = 0 = \begin{cases} V(s_i, b_j, \tau^{n+1}) - V(0, W_{liq}(s_i, b_j), \tau^{n+1}) \;, & \mathbf{x}_{i,j}^{n+1} \in \Omega_{\mathcal{B}} \;, \\ V(s_i, b_j, 0) - (W_{liq}(s_i, b_j) - \gamma/w)^2 \;, & \mathbf{x}_{i,j}^{n+1} \in \Omega_{\tau_0} \;, \\ V(s_i, b, \tau^{n+1}) - \left(\frac{b}{b_{\max}}\right)^2 V(s_i, \mathrm{sgn}(b) b_{\max}, \tau^{n+1}) \;, & \mathbf{x}_{i,j}^{n+1} \in \Omega_{b_{\max}} \;. \end{cases} \qquad (5.4)$$

**Remark 5.1** (Size of $\Omega_{b_{\max}}$)**.** *From the above discretization, we can see that $\Omega_{b_{\max}}$ needs only be the region*

$$\Omega_{b_{\max}} \quad = \quad (0, s^*] \times [-b_{\max} e^{r_b \Delta\tau}, -b_{\max}) \cup (b_{\max}, b_{\max} e^{r_\ell \Delta\tau}] \times (0, T] \;, \qquad (5.5)$$

*in order to provide all the information necessary. If we define $\Omega_{b_{\max}}$ as in equation (5.5), then the measure of $\Omega_{b_{\max}}$ tends to zero as $h \to 0$.*

**Lemma 5.1** (Local consistency)**.** *Suppose the mesh, timestep parameter, and control discretization satisfy equations (4.1-4.3), then for any $C^\infty$ function $\phi(s,b,\tau)$ in $\Omega \cup \Omega_{b_{\max}}$, with $\phi_{i,j}^{n+1} = \phi(s_i, b_j, \tau^{n+1}) = \phi(\mathbf{x}_{i,j}^{n+1})$, and for $h, \psi$ sufficiently small, $\psi$ a constant, we have that*

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1}+\psi, \left\{\phi_{a,b}^{n+1}+\psi\right\}_{\substack{a\neq i \\ or\ b\neq j}}, \left\{\phi_{k,\ell}^{n}+\psi\right\}\right)$$
$$= \begin{cases} F_{in}\phi_{i,j}^{n+1} + O(h) + O(\psi), & (s_i, b_j, \tau^{n+1}) \in \Omega_{in} \ , \\ F_{s^*}\phi_{i,j}^{n+1} + O(h) + O(\psi), & (s_i, b_j, \tau^{n+1}) \in \Omega_{s^*} \ , \\ F_{s_0}\phi_{i,j}^{n+1} + O(h) + O(\psi), & (s_i, b_j, \tau^{n+1}) \in \Omega_{s_0} \ , \\ F_{\mathcal{B}}\phi_{i,j}^{n+1} + O(\psi), & (s_i, b_j, \tau^{n+1}) \in \Omega_{B} \ , \\ F_{\tau_0}\phi_{i,j}^{n+1} + O(\psi), & (s_i, b_j, \tau^{n+1}) \in \Omega_{\tau_0} \ , \\ F_{b_{\max}}\phi_{i,j}^{n+1} + O(\psi), & (s_i, b_j, \tau^{n+1}) \in \Omega_{b_{\max}} \ . \end{cases} \tag{5.6}$$

*Proof.* To be precise, define the following notation.

$$\begin{aligned} \mathcal{P}\phi_{i,j}^{n+1} \equiv \mathcal{P}\phi(s_i, b_j, \tau^{n+1}) \ : \ ; \ (\phi_b)_{i,j}^{n+1} \equiv \phi_b(s_i, b_j, \tau^{n+1}) \\ (\phi_\tau)_{i,j}^{n+1} \equiv \phi_\tau(s_i, b_j, \tau^{n+1}) \ ; \ \mathcal{J}_\ell\phi_{i,j}^{n+1} \equiv \mathcal{J}_\ell\phi(s_i, b_j, \tau^{n+1}) \ . \end{aligned} \tag{5.7}$$

We will also use the notation $\phi_h(s,b,\tau^n)$ and $(\phi(\cdot)+\psi)_h$ to denote the linearly interpolated value of $\phi$ and $\phi(\cdot)+\psi$ on the grid with parameter $h$, at timestep $\tau^n$. Note that

$$\begin{aligned} (\phi(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, \tau^n) + \psi)_h &= \phi_h(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, \tau^n) + \psi, \\ \mathcal{P}_h(\phi_{i,j}^{n+1} + \psi) &= \mathcal{P}_h\phi_{i,j}^{n+1} - \lambda\psi \ , \\ (\mathcal{J}_\ell)_h(\phi_{i,j}^{n+1} + \psi) &= (\mathcal{J}_\ell)_h\phi_{i,j}^{n+1} + O(\psi) \ , \end{aligned} \tag{5.8}$$

and consider the case where $\mathbf{x}_{i,j}^{n+1} = (s_i, b_j, \tau^{n+1}) \in \Omega_{in}$, so that, from equation (5.1)

$$\begin{aligned} &\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}\phi_{i,j}^{n+1}+\psi, \left\{\phi_{a,b}^{n+1}+\psi\right\}_{\substack{a\neq i \\ \text{or } b\neq j}}, \left\{\phi_{k,\ell}^{n}+\psi\right\}\right) \\ &= \max\Bigg[\frac{\phi_{i,j}^{n+1} - \phi_h(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, \tau^n)}{\Delta\tau} - \mathcal{P}_h\phi_{i,j}^{n+1} - (\mathcal{J}_\ell)_h\phi_{i,j}^{n+1} + O(\psi), \\ &\phi_{i,j}^{n+1} - \min_{\substack{B^+\in\mathcal{Z}_h \\ (S^+,B^+)\in\Omega}} \phi_h(S^+(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, B^+), B^+, \tau^n) - \Delta\tau\mathcal{P}_h\phi_{i,j}^{n+1} - \Delta\tau(\mathcal{J}_\ell)_h\phi_{i,j}^{n+1} + O(\psi)\Bigg] \\ &= 0 \ . \end{aligned} \tag{5.9}$$

Noting that

$$\begin{aligned} \phi_h(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, \tau^n) &= \phi_{i,j}^n + \mathcal{R}(b_j)\ b_j(\phi_b)_{i,j}^n\Delta\tau + O(h^2) \ ; \\ \frac{\phi_{i,j}^{n+1} - \phi_{i,j}^n}{\Delta\tau} = (\phi_\tau)_{i,j}^{n+1} + O(h) \ ; \ \mathcal{P}_h\phi_{i,j}^{n+1} &= \mathcal{P}\phi_{i,j}^{n+1} + O(h) \ ; \\ (\mathcal{J}_\ell)_h\phi_{i,j}^{n+1} = \mathcal{J}_\ell\phi_{i,j}^{n+1} + O(h) \ ; \ S^+(s_i, b_j e^{\mathcal{R}\Delta\tau}, B^+) &= S^+(s_i, b_j, B^+) + O(h) \ , \end{aligned} \tag{5.10}$$

then equation (5.9) becomes

$$
\begin{aligned}
&\mathcal{G}\Big(h, \mathbf{x}_{i,j}^{n+1}\phi_{i,j}^{n+1}+\psi, \Big\{\phi_{a,b}^{n+1}+\psi\Big\}_{\substack{a\neq i\\ \text{or } b\neq j}}, \{\phi_{k,\ell}^{n}+\psi\}\Big)\\
&= \max\Bigg[(\phi_\tau)_{i,j}^{n+1} - \mathcal{R}(b_j)\ b_j(\phi_b)_{i,j}^{n} - \mathcal{P}\phi_{i,j}^{n+1} - \mathcal{J}_\ell\phi_{i,j}^{n+1} + O(\psi) + O(h),\\
&\phi_{i,j}^{n+1} - \min_{\substack{B^+\in\mathcal{Z}_h\\ (S^+,B^+)\in\Omega}} \phi_h(S^+(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, B^+), B^+, \tau^n) - \Delta\tau\mathcal{P}\phi_{i,j}^{n+1} - \Delta\tau\mathcal{J}_\ell\phi_{i,j}^{n+1} + O(\psi) + O(h)\Bigg]\\
&= \max\Bigg[(\phi_\tau)_{i,j}^{n+1} - \mathcal{R}(b_j)\ b_j(\phi_b)_{i,j}^{n+1} - \mathcal{P}\phi_{i,j}^{n+1} - \mathcal{J}_\ell\phi_{i,j}^{n+1} + O(\psi) + O(h),\\
&\phi_{i,j}^{n+1} - \min_{\substack{B^+\in\mathcal{Z}_h\\ (S^+,B^+)\in\Omega}} \phi_h(S^+(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, B^+), B^+, \tau^n) + O(\psi) + O(h)\Bigg]\\
&= \max\Bigg[(\phi_\tau)_{i,j}^{n+1} - \mathcal{L}\phi_{i,j}^{n+1} - \mathcal{J}_\ell\phi_{i,j}^{n+1},\\
&\phi_{i,j}^{n+1} - \min_{\substack{B^+\in\mathcal{Z}_h\\ (S^+,B^+)\in\Omega}} (\phi_h(S^+(s_i, b_j, B^+), B^+, \tau^n) + O(h))\Bigg] + O(h) + O(\psi)\\
&= 0\ .
\end{aligned}
\tag{5.11}
$$

Noting that since $\eta = B^+$, $\phi$ is smooth, and $\mathcal{Z}$ is compact, we have that

$$
\begin{aligned}
\min_{\substack{B^+\in\mathcal{Z}_h\\ (S^+,B^+)\in\Omega}} \phi_h(S^+(s_i, b_j, B^+), B^+, \tau^n) &= \Big(\inf_{\substack{B^+\in\mathcal{Z}\\ (S^+,B^+)\in\Omega}} \phi(S^+(s_i, b_j, B^+), B^+, \tau^{n+1})\Big) + O(h)\\
&= \mathcal{M}\phi_{i,j}^{n+1} + O(h)\ .
\end{aligned}
\tag{5.12}
$$

Using equation (5.12) in equation (5.11) gives us the final result

$$
\begin{aligned}
&\mathcal{G}\Big(h, \mathbf{x}_{i,j}^{n+1}\phi_{i,j}^{n+1}+\psi, \Big\{\phi_{a,b}^{n+1}+\psi\Big\}_{\substack{a\neq i\\ \text{or } b\neq j}}, \{\phi_{k,\ell}^{n}+\psi\}\Big)\\
&= \max\Big[(\phi_\tau)_{i,j}^{n+1} - \mathcal{L}\phi_{i,j}^{n+1} - \mathcal{J}_\ell\phi_{i,j}^{n+1}, \phi_{i,j}^{n+1} - \mathcal{M}\phi_{i,j}^{n+1}\Big] + O(h) + O(\psi)\\
&= F_{in}\phi_{i,j}^{n+1} + O(h) + O(\psi)\ .
\end{aligned}
\tag{5.13}
$$

Following similar steps, we can easily prove the remaining results in equation (5.6). □

**Definition 5.1** (Consistency: viscosity sense)**.** *Suppose the mesh, timestep parameter, and control discretization satisfy equations (4.1-4.3). For any $C^\infty$ function $\phi(s, b, \tau)$ in $\Omega \cup \Omega_{b_{\max}}$, with $\phi_{i,j}^{n+1} = \phi(s_i, b_j, \tau^{n+1}) = \phi(\mathbf{x}_{i,j}^{n+1})$, the numerical scheme $\mathcal{G}(\cdot)$ (equations (5.1-5.4)) is consistent in the viscosity sense, if, $\forall\hat{\mathbf{x}} = (\hat{s}, \hat{b}, \hat{\tau})$ with $\mathbf{x}_{i,j}^{n+1} = (s_i, b_j, \tau^{n+1})$, the following holds*

$$
\begin{aligned}
&\limsup_{\substack{h\to 0\\ \psi\to 0\\ \mathbf{x}_{i,j}^{n+1}\to\hat{\mathbf{x}}}} \mathcal{G}\Big(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1}+\psi, \Big\{\phi_{a,b}^{n+1}+\psi\Big\}_{\substack{a\neq i\\ or\ b\neq j}}, \{\phi_{k,\ell}^{n}+\psi\}\Big)\\
&\qquad \le F^*(\hat{\mathbf{x}}, \phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), D^2\phi(\hat{\mathbf{x}}), \mathcal{M}\phi(\hat{\mathbf{x}}), \mathcal{J}_\ell\phi(\hat{\mathbf{x}})),
\end{aligned}
\tag{5.14}
$$

*and*

$$\liminf_{\substack{h\to 0\\ \psi\to 0\\ \mathbf{x}_{i,j}^{n+1}\to\hat{\mathbf{x}}}} \mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, \phi_{i,j}^{n+1}+\psi, \left\{\phi_{a,b}^{n+1}+\psi\right\}_{\substack{a\neq i\\ or\ b\neq j}}, \left\{\phi_{k,\ell}^{n}+\psi\right\}\right) \geq F_*(\hat{\mathbf{x}}, \phi(\hat{\mathbf{x}}), D\phi(\hat{\mathbf{x}}), D^2\phi(\hat{\mathbf{x}}), \mathcal{M}\phi(\hat{\mathbf{x}}), \mathcal{J}_\ell\phi(\hat{\mathbf{x}}))\ . \tag{5.15}$$

**Lemma 5.2** (Consistency of Scheme (5.1-5.4))**.** *Provided all the conditions for Lemma 5.1 are satisfied then scheme (5.1-5.4) is consistent according to definition 5.1.*

*Proof.* This follows in straightforward fashion from Lemma 5.1, using the same steps as in, for example [29]. □

## 5.2 Monotonicity and Stability

Monotonicity is defined as follows

**Definition 5.2** (Monotonicity)**.** *The numerical scheme (5.1-5.4) is monotone if for all* $Y_{i,j}^n \geq X_{i,j}^n, \forall i,j,n$

$$\mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, V_{i,j}^{n+1}, \left\{Y_{a,b}^{n+1}\right\}_{\substack{a\neq i\\ or\ b\neq j}}, \left\{Y_{k,\ell}^{n}\right\}\right) \leq \mathcal{G}\left(h, \mathbf{x}_{i,j}^{n+1}, V_{i,j}^{n+1}, \left\{X_{a,b}^{n+1}\right\}_{\substack{a\neq i\\ or\ b\neq j}}, \left\{X_{k,\ell}^{n}\right\}\right) \tag{5.16}$$

**Lemma 5.3** (Monotonicity)**.** *If the scheme (5.1-5.4) has the properties*

- *The positive coefficient condition is satisfied (equation (4.6)).*
- *The discretization of* $\mathcal{J}_\ell$ *has quadrature weights satisfying (4.8).*
- *Linear interpolation is used, if necessary, to compute* $V_h(\cdot)$*,*

*then the discretization is monotone, according to Definition 5.2.*

*Proof.* This is easily done using the same steps as in [30]. □

Finally, the discretization (5.1-5.4) is $\ell_\infty$-stable, which is a consequence of the following Lemma.

**Lemma 5.4** (Stability)**.** *If the conditions for Lemma 5.3 are satisfied, then the discretization (5.1-5.4) satisfies*

$$\begin{aligned} 0 \leq V_{i,j}^n &\leq \|V^0\|_\infty e^{r_{\max}T} \\ r_{\max} &= \max(r_\ell, r_b) \\ (s_i, b_j, \tau^n) &\in \Omega\ , \end{aligned} \tag{5.17}$$

*for* $0 \leq n \leq N$*,* $T = N\Delta\tau$*, as* $\Delta\tau \to 0$*,* $h \to 0$*.*

*Proof.* This follows from a straightforward maximum analysis (e.g. the same steps as in [30]), since $\Omega$ is a bounded domain. The term $e^{r_{\max}T}$ in equation (5.17) is a result of the evaluation of $V_h(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, \tau^n)$ using equation (5.4) at points near $\pm b_{\max}$. □

### 5.3 Convergence

**Theorem 5.1** (Convergence)**.** *Assume that discretization (5.1-5.4) satisfies all the conditions required for Lemmas 5.2, 5.3 and 5.4 and that Assumption 3.1 holds, then scheme (5.1-5.4) converges to the unique continuous viscosity solution of Problem 3.1 in* $\Omega_{in} \cup \Gamma$.

*Proof.* Since the scheme is monotone, consistent and $\ell_\infty$-stable, this follows from the results in [16]. □

**Remark 5.2.** *Since we have assumed strong comparison holds only in* $\Omega_{in} \cup \Gamma$*, then we can guarantee uniqueness and continuity only in* $\Omega_{in} \cup \Gamma$.

## 6 Implementation Details

### 6.1 Complexity

Examination of equation (4.9) reveals that each timestep requires

- Solution of a local optimization problem at each node (evaluation of $\widetilde{V}^n_{i,j}$).
- A linear time advance step. At each time step, each fixed point iteration for the solution of the discretized equations requires $2j_{\max}$ FFT evaluations and solution of $j_{\max}$ tridiagonal systems. This is a result of implicit treatment of the jump term [24].

In order to solve the local optimization problems, we use simple linear search to find the minimum for $B^+ \in \mathcal{Z}_h$. We have found that using a continuous 1-D optimization method is unreliable, and often converges to a local, not global, minimum. The complexity of the time advance is thus dominated by the solution of the local optimization problems. Each optimization problem is resolved by evaluating the objective function $O(1/h)$ times. There are $O(1/h^2)$ nodes, and $O(1/h)$ timesteps giving a total complexity of $O(1/h^4)$.

### 6.2 Construction of the Efficient Frontier

At each timestep, we solve a discrete approximation to equation (2.31). The optimal controls $(\phi^*, \eta^*)$ at each node are then used to solve a discrete approximation to the expected value equation (2.47) for this same timestep. We continue to alternate solution of equations (2.31) and equation (2.47) at each timestep until the stopping time is reached.

For fixed $\gamma$, let

$$\begin{aligned} V_0(W_{init}) &= V_h(s=0, b=W_{init}, \tau=T) \\ U_0(W_{init}) &= U_h(s=0, b=W_{init}, \tau=T) \\ & W_{init} = \text{ initial wealth } . \end{aligned} \tag{6.1}$$

From

$$\begin{aligned} V_0(W_{init}) &= \left( E^{t=0}_{\mathcal{C}^*}[(W_{liq}(T) - \gamma/2)^2] \right)_h \\ U_0(W_{init}) &= \left( E^{t=0}_{\mathcal{C}^*}[W_{liq}(T)] \right)_h , \end{aligned} \tag{6.2}$$

where $(\cdot)_h$ refers to a discrete approximation to the expression in the brackets, we have that

$$\begin{aligned}\left(Var_{\mathcal{C}^*}^{t=0}[W_{liq}(T)]\right)_h &= V_0(W_{init}) + \gamma U_0(W_{init}) - \frac{\gamma^2}{4} - U_0(W_{init})^2 \\ \left(E_{\mathcal{C}^*}^{t=0}[W_{liq}(T)]\right)_h &= U_0(W_{init}) \ , \end{aligned} \tag{6.3}$$

which gives us a single point $\mathcal{Y}_{Q(\gamma)}$. Repeating this for many values of $\gamma$ gives us an approximation to $\mathcal{Y}_Q$ (see Theorem 2.1). Finally, the efficient frontier is constructed from the upper left convex hull of $\mathcal{Y}_Q$ [21] to remove spurious points. In our case, it turns out that all the points in $\mathcal{Y}_Q$ are Pareto points (i.e. there are no spurious points).

Note that the smallest possible value of $\gamma$ is

$$\gamma_{\min} = 2W_{init}e^{r_\ell T} \ , \tag{6.4}$$

which corresponds to an infinitely risk averse investor ($\rho^* \to \infty$, see equation (2.16)), who invests only in the risk-free asset. In practice, the interesting part of the efficient frontier is in the range $\gamma \in [\gamma_{\min}, 10\gamma_{\min}]$.

**Remark 6.1** (Computational grid)**.** *For ease of exposition, we have outlined the discretization method for a rectangular $(s, b)$ grid. However, since the semi-Lagrangian timestepping decouples each PIDE for each $b_j$ value, there is no need to use the same $s$ grid for every $b_j$ value. Our actual implementation makes use of this to concentrate nodes near the liquidation boundary. Consequently, it is a simple matter to handle cases where the liquidation boundary $W_{liq}(s, b) = 0$ is an arbitrary curve.*

## 6.3 An Improved Linear Interpolation Scheme

Recall that, when solving the value function problem (2.15) or the expected value problem (2.41) on a computational grid, it is usually required to evaluate $V_h(\cdot)$ or $U_h(\cdot)$, respectively, at points other than a node of the computational grid. Hence, interpolation must be employed. As mentioned earlier, to preserve the monotonicity of the numerical schemes, linear interpolation on nodal values is used in our implementation. In this subsection, we discuss a special linear interpolation scheme applied along the $b$-direction at $s = 0$ which, as illustrated by numerical results in Subsection 7.1, can significantly improve the accuracy of the interpolation.

### 6.3.1 Value Function

Recall that the value function is the viscosity solution of the HJB PDE equation (2.40). Recall the initial condition

$$V(s, b, 0) = (W_{liq}(s, b) - \gamma/2)^2 ; \quad \text{if } \tau = 0 \ .$$

Following from this equation, we have

$$V_h(0, \gamma/2, \tau^0) = 0 \ , \tag{6.5}$$

even in the presence of transaction costs. Note that the optimal rebalancing at time $\tau^0$ does not reallocate the point $(s, b) = (0, \gamma/2)$, since the minimum value of the objective function is attained at this point.

Next, we draw the reader's attention to the fact that, at $s = 0$, the PIDE degenerates to the first-order hyperbolic equation $V_\tau = \mathcal{R}(b)\ bV_b$. Note that this hyperbolic equation is part of Equation (2.27). The exact solution of this first-order hyperbolic equation must satisfy

$$V(0, be^{-\mathcal{R}(b)\tau}, \tau) = V(0, b, 0)\ . \tag{6.6}$$

Combining (6.5) and (6.6), we obtain

$$V_h(0, (\gamma/2)\ e^{-\mathcal{R}(\gamma/2)\tau^n}, \tau^n) = V_h(0, \gamma/2, \tau^0) = 0\ . \tag{6.7}$$

That is, for any timestep $\tau^n$, the exact solution for the value function problem at the special point $(s, b) = (0, (\gamma/2)\ e^{-\mathcal{R}(\gamma/2)\tau^n})$ is zero. Since the value function can never be less than zero, no reallocation takes place at this node. Figure 6.1 (a) illustrates how this special point moves along the $b$-direction from the time $\tau^0$ to $\tau^n$. Below, we discuss how the result (6.7) could be incorporated into the (linear) interpolation scheme.

Assume that we want to proceed from timestep $\tau^n$ to $\tau^{n+1}$, and that we want to compute $V_h(0, \bar{b}, \tau^n)$, where $\bar{b}$ is neither a grid point in the $b$-direction nor the special value $(\gamma/2)\ e^{-\mathcal{R}(\gamma/2)\tau^n}$. This situation could happen when solving Equation (2.27) or the local optimization problem on the right-side of (4.9). Furthermore, assume that $b_j < \bar{b} < b_{j+1}$ for some grid points $b_j$ and $b_{j+1}$ in the $b$-direction. For presentation purposes, let $b_{special} = (\gamma/2)\ e^{-\mathcal{R}(\gamma/2)\tau^n}$ and $V_{special} = 0$. A linear interpolation scheme for computing $V_h(0, \bar{b}, \tau^n)$ is presented in Algorithm 6.1. Figure 6.1 (b) provides a pictorial illustration of this interpolation scheme.

**Algorithm 6.1** Improved linear interpolation scheme along the $b$-direction at $s = 0$ for the function value problem (2.15).

**if** $b_{special} < b_j$ or $b_{special} > b_{j+1}$ **then**
  set $b_{left} = b_j$, $V_{left} = V_h(0, b_j, \tau^n)$, $b_{right} = b_{j+1}$, and $V_{right} = V_h(0, b_{j+1}, \tau^n)$;
**else**
  **if** $b_{special} \leq \bar{b}$ **then**
    set $b_{left} = b_{special}$, $V_{left} = V_{special}$, $b_{right} = b_{j+1}$, and $V_{right} = V_h(0, b_{j+1}, \tau^n)$;
  **else**
    set $b_{left} = b_j$, $V_{left} = V_h(0, b_{j+1}, \tau^n)$, $b_{right} = b_{special}$, and $V_{right} = V_{special}$;
  **end if**
**end if**
apply linear interpolation to $(b_{left}, V_{left})$ and $(b_{right}, V_{right})$ to compute $V_h(0, \bar{b}, \tau^n)$;

### 6.3.2 Expected Value Problem (2.41)

Following the same lines of reasoning used for the function value problem (2.15), we have that

$$U_h(0, \gamma/2, \tau^0) = \gamma/2\ , \tag{6.8}$$

and

$$U_h(0, (\gamma/2)\ e^{-\mathcal{R}(\gamma/2)\tau^n}, \tau^n) = U_h(0, \gamma/2, \tau^0) = \gamma/2\ . \tag{6.9}$$

even in the presence of transaction costs. (Note Equation (2.43) and the first first-order hyperbolic equation $U_\tau = \mathcal{R}(b)\ bU_b$.) Algorithm 6.1 with $b_{special} = (\gamma/2)\ e^{-\mathcal{R}(\gamma/2)\tau^n}$ and $V_{special} = \gamma/2$ can be used to compute $U_h(0, \bar{b}, \tau^n)$, where $\bar{b}$ is neither a grid point in the $b$-direction nor the special value $(\gamma/2)\ e^{-\mathcal{R}(\gamma/2)\tau^n}$.

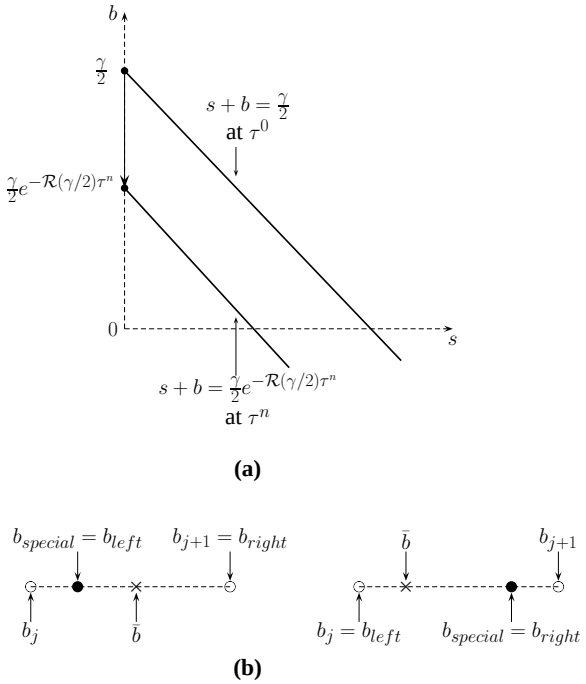


FIGURE 6.1: *(a) Special interpolation point movements in backward time; (b) Pictorial illustration of the (linear) interpolation scheme described in Algorithm 6.1*

# 7 Numerical Examples

In this section, we present selected numerical results of our PDE approach applied to the continuous time mean-variance portfolio allocation problem. For all the experiments, unless otherwise noted, the details of grid and timestep refinement levels used are given in Table 7.1.

| Refinement | Timesteps | $S$ Nodes | $B$ Nodes |
|---|---|---|---|
| 0 | 50 | 58 | 115 |
| 1 | 100 | 115 | 229 |
| 2 | 200 | 229 | 457 |
| 3 | 400 | 457 | 913 |

TABLE 7.1: *Grid and timestep refinement levels used during numerical tests. On each refinement, a new grid point is placed halfway between all old grid points and the number of timesteps is doubled. A constant timestep size is used.* $s_{\max} = 20000$, $b_{\max} = 10000$, $s^* = 10000$.

## 7.1 Effects of the improved interpolation scheme

In this subsection, we discuss the effects on the numerical results of the linear interpolation scheme described in Subsection 6.3, where, at each time $\tau^n$, exact boundary conditions are used for the special point $(s, b) = (0, (\gamma/2)\ e^{-\mathcal{R}(\gamma/2)\tau^n})$. As an illustrating example, we consider the no-jump case with data in Table 7.2, with the exception that trading continues even if insolvent. This allows

us to compare with the exact closed form solution [1, 2].

**Remark 7.1** (Unbounded expected value)**.** *For this example, the expected value becomes unbounded, violating one of the conditions needed to apply Theorem 2.1. However, in this case, we have a closed form solution available, which confirms that there are no spurious points on the computed efficient frontier.*

Figure 7.1 (a) presents the numerical efficient frontiers obtained using standard linear interpolation. The exact efficient frontier is constructed using the expression in [1, 2]. It is clear that, while the numerical efficient frontiers agree well with the exact efficient frontier for relatively large standard deviations, they are very inaccurate for small standard deviations. More specifically, it appears that, in this case, the numerical methods were not able to construct, to within the accuracy of methods, the special point on the exact efficient frontier

$$(Var^{x,t}_{\mathcal{C}^*}[W_{liq}(T)], E^{x,t}_{\mathcal{C}^*}[W_{liq}(T)]) = (0, 100e^{10\times 0.0445}) \approx (0, 156.049). \tag{7.1}$$

This trivial point corresponds to the case where the investor invests only in the bond, and not in the risky asset (hence, there is no risk).

Figure 7.1 (b) presents the numerical efficient frontiers obtained with the improved linear interpolation scheme. More specifically, at each timestep $\tau^n$, for interpolation along the $b$-direction at $s = 0$, Algorithm 6.1 is utilized, and, otherwise, standard linear interpolation is used. It is obvious that the numerical efficient frontiers obtained with the improved linear interpolation scheme agree very well with the exact efficient frontier, even for small standard deviations. In particular, the special point on the exact efficient frontier (7.1) is now approximated accurately. This result highlights the importance of using exact boundary conditions (where available) for linear interpolation in constructing accurate numerical efficient frontiers. In all our numerical experiments in this section, unless otherwise stated, the improved linear interpolation scheme is used.

## 7.2 Validation Examples

In this subsection, we provide select examples to validate our proposed numerical approach. For comparison purposes, we only consider several special cases of the continuous time mean-variance portfolio allocation problem where exact efficient frontiers can be constructed.

### 7.2.1 No jumps, insolvency not allowed, no maximum leverage, no transaction costs

We consider the example where (i) the underlying asset follows a GBM without jumps, (ii) insolvency is not allowed, (iii) $q_{\max} = \infty$, and (iv) no transaction costs. Input parameters and data for this test is given in Table 7.3. In this case, exact efficient frontiers can be constructed using algorithms in [4]. That is, given a value for the mean, the exact standard deviation of the point on the efficient frontier having that mean can be found. Alternatively, one could fix the standard deviation and compute the exact mean.

Table 7.4 presents computed means and standard deviations for different refinement levels when $\gamma = 800$. To provide an estimate of the convergence rate of the algorithm, we compute the "change" as the difference in values from the coarser grid and the "ratio" as the ratio of changes between successive grids. The numerical results indicate first-order convergence is achieved for the algorithm.

| | Jumps | No Jumps |
|---|---|---|
| Investment Horizon $T$ | 10 | 10 |
| Lending rate $r_\ell$ | .0445 | .0445 |
| Borrowing rate $r_b$ | .0445 | .0445 |
| Trading ceases if insolvent (2.23) | yes | yes |
| Volatility $\sigma$ | 0.1765 | .281751 |
| Drift $\mu$ | .0795487 | .0795487 |
| Initial Wealth | 100 | 100 |
| Maximum (Risky Asset)/Wealth Ratio $q_{\max}$ | $\infty$ | $\infty$ |
| $\nu$ | -.788325 | N/A |
| $\lambda$ | .0585046 | N/A |
| $\zeta$ | .450500 | N/A |
| Fixed Transaction Cost $c_1$ | 0.0 | 0.0 |
| Proportional Transaction Cost $c_2$ | 0.0 | 0.0 |

TABLE 7.2: *Input parameters and data for comparison of the jump and no jump cases. Effective volatility for the no jump case based on jump parameters and computed as in [31]. For the no-jumps case, insolvency is not allowed. For the jump case, immediate liquidation is enforced in the case of insolvency.*

| | |
|---|---|
| Investment Horizon | 10 |
| Lending rate $r_l$ | .04 |
| Borrowing rate $r_b$ | .04 |
| Trading ceases if insolvent (2.23) | yes |
| Volatility $\sigma$ | 0.15 |
| Drift $\mu$ | 0.15 |
| Initial Wealth | 100 |
| Maximum (Risky Asset)/Wealth Ratio $q_{\max}$ | $\infty$ |
| $\nu$ | N/A |
| $\lambda$ | 0.0 |
| $\zeta$ | 0.0 |
| Fixed Transaction Cost $c_1$ | 0.0 |
| Proportional Transaction Cost $c_2$ | 0.0 |

TABLE 7.3: *No-jump test case. Input parameters and data for the validation test in Subsubsection 7.2.1.*

Of course, we would like to verify that the numerical solution converges to the known exact solution in this case. However, Table 7.4 shows convergence for a fixed $\gamma$. The exact solution in [4] gives points on the efficient frontier, i.e. given a value of the mean, then the standard deviation is determined. One way around this problem would be to use a very fine grid to construct a numerical solution for fixed $\gamma$, and then verify that the standard deviation is consistent with the exact standard deviation. However, we were not able to compute benchmark means (or standard deviations) on a grid finer than the finest grid in Table 7.1, due to the high computational cost.

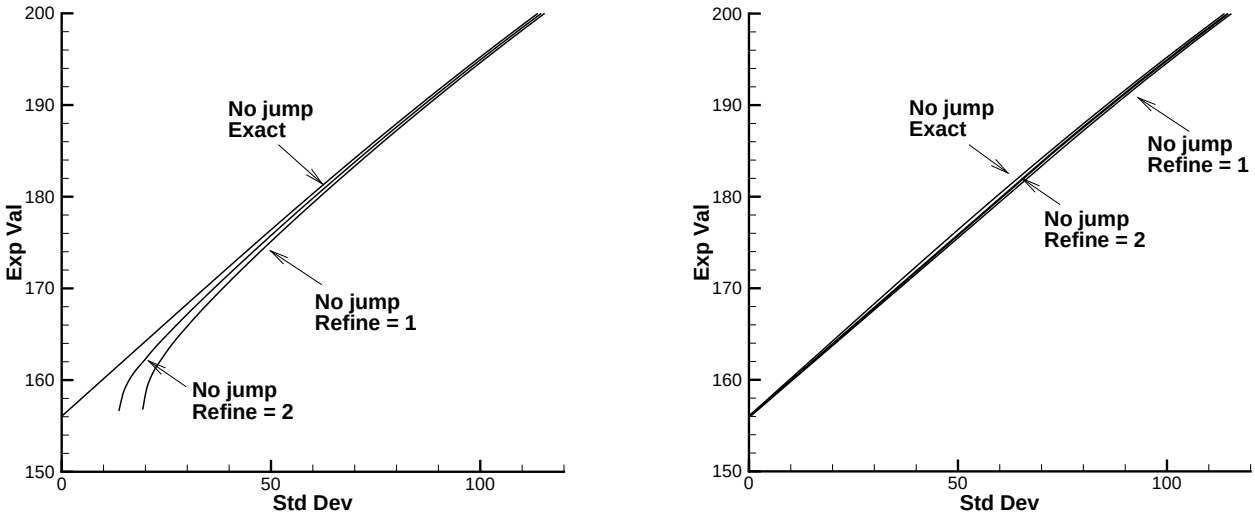


FIGURE 7.1: *Efficient frontier, no-jump cases, data in Table 7.2, with the exception that trading continues if insolvent. The exact from [1, 2] is used. Close-up of efficient frontier for small standard deviations. (a) standard linear interpolation. (b) improved interpolation method, using the exact boundary condition at one point along the $s = 0$ boundary (see Subsection 6.3), otherwise standard linear interpolation.*

| Refine | Mean | Change | Ratio | Standard Deviation | Change | Ratio |
|---|---|---|---|---|---|---|
| 0 | 377.714323 | | | 62.069472 | | |
| 1 | 381.379938 | 3.665615 | | 56.292507 | -5.776965 | |
| 2 | 383.104304 | 1.724366 | 2.1 | 53.503351 | -2.789156 | 2.1 |
| 3 | 383.966487 | 0.862182 | 2.0 | 52.108774 | -1.394578 | 2.0 |

TABLE 7.4: *Validation test (data in Table 7.3), $\gamma = 800$. No jumps, insolvency not allowed, no maximum leverage, no transaction costs*

In order to reconcile these two different forms of the solution, we proceed as follows.

1. Step 1: apply extrapolation to the numerical means in Table 7.4, assuming first-order convergence (which is what we observe), to obtain a benchmark mean.

2. Step 2: compute the corresponding benchmark standard deviation using algorithms in [4] and the mean obtained in Step 1.

3. Step 3: check whether the numerical standard deviations in Table 7.4 exhibit first-order convergence to the benchmark standard deviation obtained in Step 2.

From Table 7.4, the benchmark mean is 384.828663. Using algorithms in [4], the corresponding

benchmark standard deviation is 50.686326. Table 7.5 presents the results of the above-described convergence check. Here, we compute the "error" as the difference in values between computed standard deviations and the benchmark value. Clearly, the numerical standard deviations exhibit first-order convergence to the benchmark value.

| Refine | Standard Deviation | Error | Ratio |
|---|---|---|---|
| 0 | 62.069472 | 11.383145 | |
| 1 | 56.292507 | 5.606181 | 2.0 |
| 2 | 53.503351 | 2.817025 | 2.0 |
| 3 | 52.108774 | 1.422447 | 2.0 |

TABLE 7.5: *Validation test (data in Table 7.3). No jumps, insolvency not allowed, no maximum leverage condition, no transaction costs. Exact standard deviation* 50.686326 *computed using [4].*

### 7.2.2 Jumps, insolvency allowed, no maximum leverage, no transaction costs

As the second validation example, we consider the example where (i) the underlying asset follows a GBM with jumps described in (2.2), (ii) insolvency is allowed, (iii) $q_{\max} = \infty$, and (iv) no transaction costs. For this special case, the exact efficient frontier can be constructed using results in Appendix A. Input parameters and data for this test are given in Table 7.2, except that, in this test, insolvency is allowed.

Table 7.6 presents computed means and standard deviations for different refinement levels. The results indicate that first-order convergence for computed means is attained, while, for computed standard deviations, we observe slightly less than first-order convergence.

Next, we illustrate the convergence of computed means and standard deviations to the exact mean and standard deviation of a point on the efficient frontier. We proceed in the same fashion as in Section 7.2.1. In this case, the benchmark mean is 163.156862. Using Appendix A, the benchmark standard deviation is 13.304860. Table 7.7 presents the results of this convergence check. We observe that the numerical standard deviations exhibit convergence to the benchmark at roughly the same convergence rate as observed in Table 7.6, i.e. slightly less than first-order.

| Refine | Expected Value | Change | Ratio | Standard Deviation | Change | Ratio |
|---|---|---|---|---|---|---|
| 0 | 166.538872 | | | 30.402662 | | |
| 1 | 164.889882 | -1.648989 | | 23.403936 | -6.998726 | |
| 2 | 164.023370 | -0.866512 | 1.9 | 19.288471 | -4.115465 | 1.7 |
| 3 | 163.590115 | -0.433256 | 2.0 | 16.867609 | -2.4208617 | 1.7 |

TABLE 7.6: *Validation example, jump case,* $\gamma = 380$. *Insolvency allowed, i.e.* $\mathcal{B} = \emptyset$, *otherwise data in Table 7.2.*

| Refine | Standard Deviation | Error | Ratio |
|---|---|---|---|
| 0 | 30.402662 | -17.097802 | |
| 1 | 23.403936 | -10.099076 | 1.7 |
| 2 | 19.288471 | -5.9836114 | 1.7 |
| 3 | 16.867609 | -3.5627494 | 1.7 |

TABLE 7.7: *Validation test (data in Table 7.2). Jumps, insolvency allowed, no maximum leverage, no transaction costs. Exact standard deviation is 13.304860 computed using Appendix A.*

## 7.3 Comparison between jump and no-jump cases

In this section, we compare the jump and no-jump cases in terms of mean-variance efficiency for the continuous time portfolio allocation problem. As an illustrative example, we consider the example where (i) lending and borrowing rates are the same, (ii) insolvency not allowed, (iii) no maximum leverage, and (iv) no transactions costs. Input parameters and data for this test are given in Table 7.2. For the no-jump case, the exact efficient frontier can be constructed using algorithms in [4].

Figure 7.2 presents efficient frontiers for the jump and no-jump cases for various refinement levels. Observe that, for each case, the difference between various refinement levels (i.e. the discretization errors) is small. In addition, for the no-jump case, it is clear that the numerical efficient frontiers for all refinement levels agrees well with the exact solution. Recall that the no-jump parameters are computed by determining an *effective* volatility which approximates the behaviour of a jump diffusion by a diffusion process [31]. Figure 7.2 illustrates the fact that the efficient frontiers computed using a jump diffusion are considerably different from the efficient frontier computed using a diffusion approximation to a jump diffusion, even for relatively long investment horizons (e.g. 10 years).

## 7.4 Sensitivity of Efficient Frontiers

In this section, we illustrate the effects on the efficient frontiers when realistic financial modeling, as well as realistic constraints on the portfolio, are included. In particular, we consider the presence of (i) different borrowing and lending interest rates, (ii) transaction costs, and (iii) a maximum leverage condition.

We consider four examples with details listed in Table 7.8. Note that the example with $c_2 = 0.005$ can be viewed as a relatively extreme case, since this value of $c_2$ is equivalent to a proportional cost of about 50 bps per transaction.

| | rates | | leverage cond. | trans. costs. | | Other |
|---|---|---|---|---|---|---|
| Experiment | $r_\ell$ | $r_b$ | $q_{\max}$ | $c_1$ | $c_2$ | data |
| Example (a) | $r$ | $r$ | 1.5 | 0.0 | 0.0 | Table 7.2 |
| Example (b) | $r-1\%$ | $r+1\%$ | 1.5 | 0.0 | 0.0 | —"— |
| Example (c) | $r$ | $r$ | 1.5 | 0.001 | 0.001 (0.005) | —"— |
| Example (d) | $r-1\%$ | $r+1\%$ | 1.5 | 0.001 | 0.001 (0.005) | —"— |

TABLE 7.8: *Details of experiments with realistic modeling and constraints. Here,* $r = 0.0445$.

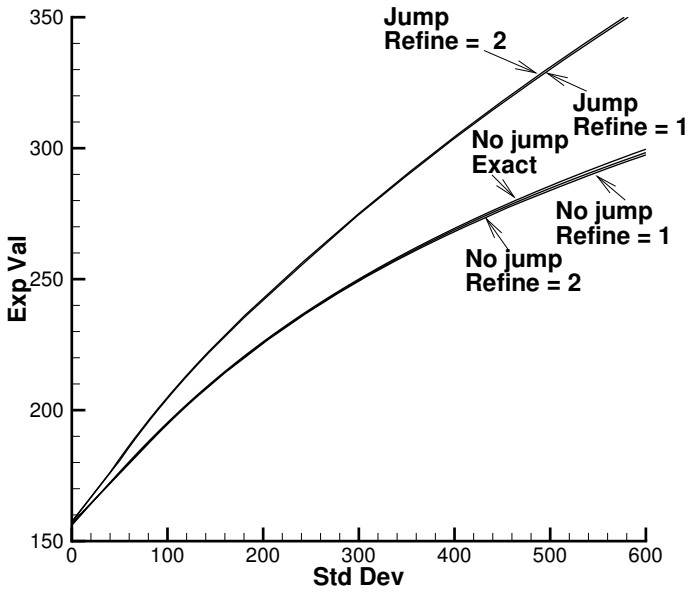


FIGURE 7.2: *Efficient frontier, jump and no-jump cases, data in Table 7.2. For the no-jump case, the exact efficient frontier from [4] is used. Large scale plot, improved interpolation using the exact boundary condition at one point along the $S = 0$ boundary (see Section 6.3) used.*

Figure 7.3 presents numerical efficient frontiers obtained with refinement level 2 for Examples (a) and (b). For comparison purposes, in this figure, we also include the exact efficient frontier for the corresponding jump case where trading is allowed even if insolvent. In this case, the exact efficient frontier is a straight line. (See Appendix A for exact solution.)

Figure 7.4 (a) and (b) present numerical efficient frontiers obtained with refinement level 2 for Examples (c) and (d). Note that, for comparison purposes, the numerical efficient frontier with $c_1 = c_2 = 0$ from Figure 7.3 are repeated in these figures.

A common observation is that, when more (realistic) constraints or modeling features are included, the expected means (for fixed standard deviation) become significantly smaller. The curves also flatten out quickly. From the perspective of the investor, this observation is important, since, in these cases, accepting substantially more risk does not necessarily result in considerably higher rate of return of the portfolio.

We conclude this section by emphasizing that, all PDE results presented in this section have also been verified by Monte-Carlo (MC) simulation. More specifically, we carried out MC simulations using the optimal strategies obtained from the PDE methods. In all case, we observed convergence of the MC simulations means and standard deviations to the respective PDE values, as the numbers of simulations and timesteps increase.

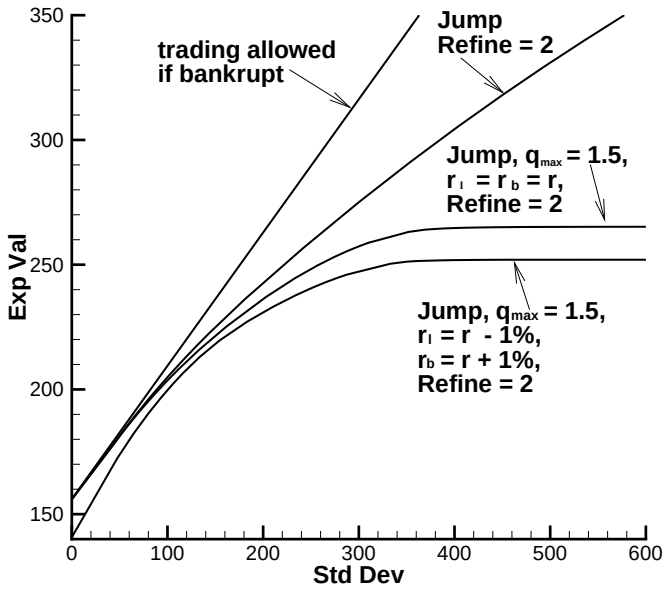


FIGURE 7.3: *Efficient frontiers for Examples (a) and (b) in Table 7.8. For the case trading is allowed even if bankruptcy in (a), the efficient frontier in Appendix A is used..*

### 7.5 Localization Error

All the previous computations used $s_{\max} = 20000$, $s^* = 10000$, $b_{\max} = 10000$. Increasing these values by an order of magnitude resulted in no change to the points on the efficient frontier to eight digits.

## 8 Conclusion

In this paper, we develop an efficient fully numerical PDE approach for the pre-commitment continuous time mean-variance asset allocation problem when the underlying asset follows a jump diffusion process. A standard formulation of this problem gives rise to a 1-D non-linear HJB PIDE with the control present in the integrand of the jump term, which is very challenging to solve efficiently numerically. Using the impulse control framework, we formulate the asset allocation problem as the solution to a 2-D impulse control problem in the form of a non-linear HJB PIDE, with one dimension for each asset in the portfolio. We then develop a numerical scheme based on a semi-Lagrangian type method, which decouples each PIDE for each discrete value of the riskless asset, i.e. the bond. More specifically, our numerical approach involves solving, at each timestep, a sequence of 1-D non-controlled PIDEs. The optimal controls are then obtained from solving the optimization problem originated from an optimal rebalancing of the portfolio. We show that our numerical scheme is monotone, consistent, and $\ell_\infty$-stable. Hence, the numerical solution is guaranteed to converge to the unique viscosity solutions of the corresponding HJB PIDE, assuming that the HJB PIDE satisfies a strong comparison property.

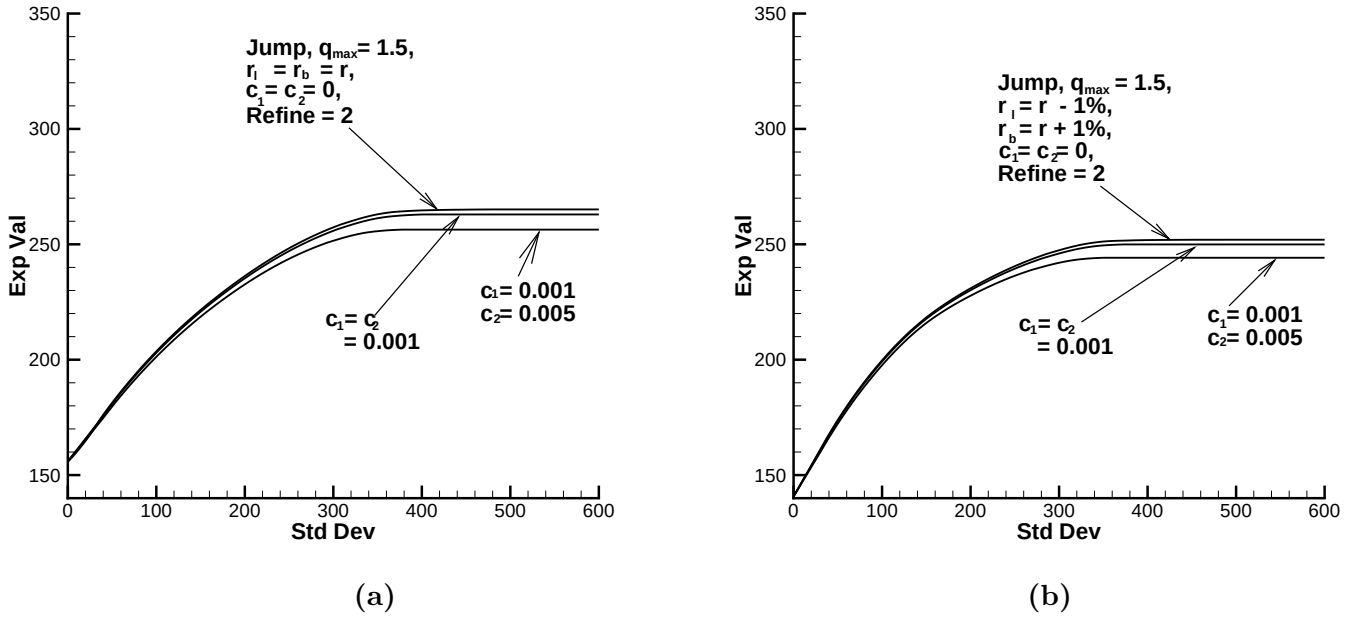


FIGURE 7.4: *Efficient frontiers for Examples (c) and (d) in Table 7.8.*

Use of monotone (linear) interpolation at each semi-Lagrangian timestep results in an inaccurate efficient frontier at small values of the standard deviation. This problem is eliminated by using the exact value function solution at a single point along the $s = 0$ boundary.

Another focus of this paper is the inclusion of realistic financial modeling, such as different borrowing and lending interest rates and transaction costs, as well as realistic constraints on the portfolio, such as maximum limits on borrowing and solvency of the portfolio. Our numerical results indicate that these realistic features have considerable effect on the efficient frontier.

Due to the long time horizon of the investment, the assumption that the interest rate is constant/deterministic is questionable. One approach for taking into account stochastic interest rates, in an economically appropriate way for long term contracts, is by means of a regime switching process [32]. This approach will be computationally efficient compared to a multi-factor stochastic interest rate model. We intend to investigate this approach in the future.

# Appendix

## A Exact Efficient Frontier: Jump Diffusion, Trading Continues if Insolvent

Based on [33] and [34], the exact efficient frontier for the investment problem can be determined. It is assumed that trading continues even if the investor is insolvent, infinite borrowing is permitted, and there are no transaction costs.

Assuming processes (2.2) and (2.3) then in this case, the efficient frontier is a straight line, given

by

$$\text{Var}[W_T] \;=\; \frac{\Big(E[W_T] - W_0 e^{rT}\Big)^2}{(e^{AT}-1)}\;, \tag{A.1}$$

where

$$A \;=\; \frac{(\mu - r)^2}{\sigma^2 + \lambda E[(\nu-1)^2]}\;, \tag{A.2}$$

and

$$\begin{aligned} \sigma = \text{ Volatility} \quad &; \quad \mu = \text{ real world drift} \;\; ; \\ r = \text{ risk-free rate} \quad &; \quad \lambda = \text{ intensity of Poisson jump process} \;\; ; \\ \nu = \text{ jump size} \quad &; \quad S \to \nu S \text{ when a jump occurs} \;; \\ T = \text{ Investment Horizon} \quad &; \quad W_T = \text{ Wealth at time } T \;; \\ W_0 = \text{ Wealth at initial time} \quad &; \quad S_t = \text{ amount invested in risky asset} \;. \end{aligned} \tag{A.3}$$

Note that $E[\xi] = e^{\xi+\zeta^2/2}$ and $E[\xi^2] = e^{2\xi+2\zeta^2}$. It is also interesting to note that the optimal strategy in this case is [33]

$$\begin{aligned} S_t \;&=\; \frac{\mu - r}{\sigma^2 + \lambda E[(\nu-1)^2]}\left(\frac{\gamma e^{-r(T-t)}}{2} - W_t\right) \\ &\quad W_t = \text{ Wealth at time } t \\ &\quad W_0 < \frac{\gamma e^{-r(T)}}{2} \end{aligned} \tag{A.4}$$

from which we can see that $S_t$ cannot exceed the discounted target unless a jump occurs.

# B Intuitive Derivation of the Discretization (4.9)

In this Appendix, we provide an intuitive derivation of the discretization (4.9). Below, we first discuss the evolution in forward time of the value function (2.15), then provide an intuitive derivation of (4.9).

## B.1 Evolution of the value function in forward time

Consider a set of discrete rebalancing times $\{t_1, t_2, \ldots\}$ where $t_{i+1} - t_i = \Delta t$. We also define

$$t_i^+ = t_i + \epsilon \quad ; \quad t_i^- = t_i + \epsilon \tag{B.1}$$

where $\epsilon \ll 1$ but finite. Assume that the portfolio consists of $s = S(t)$ and $b = B(t)$ amount of the stock and bond at time $t = t_i^+$, respectively, i.e. after the rebalancing at time $t_i$. Let $t_{i+1} = t_i + \Delta t$. Over the time period $[t_i^+, t_{i+1}^+]$, the evolution of the portfolio can be viewed as consisting of the following three steps.

1. Over the time period $[t_i^+, t_{i+1}^-]$, the stock amount evolves randomly from $S$ to $S + dS$, where $dS$ follows (2.6); however, the bond amount remains unchanged, since no interest is paid.

2. Over the time period $[t_{i+1}^-, t_{i+1}]$, the stock amount remains unchanged, but the bond amount changes from $B$ to $Be^{\mathcal{R}(B)\Delta t}$, due to the interest payment.

3. Over the time period $[t_{i+1}, t_{i+1}^+]$, the rebalancing of the portfolio occurs.

We now investigate how the value function $\bar{V}(s = S(t), b = B(t), t)$ changes over the above-mentioned three time periods.

1. Over the time period $[t_i^+, t_{i+1}^-]$, the value function $\bar{V}(s, b, t)$ evolves according to the PIDE

$$\bar{V}_t + \mathcal{P}\bar{V} + \mathcal{J}\bar{V} = 0, \tag{B.2}$$

where the differential operator $\mathcal{P}$ is defined in (4.5). Note that the term $\mathcal{R}(b)b\bar{V}_b$ does not appear in $\mathcal{P}\bar{V}$, since the bond amount remains constant over this time period. Denote by $\bar{V}(s, b, t_{i+1}^-)$ the resulting value function at time $t_{i+1}^-$.

2. Over the time period $[t_{i+1}^-, t_{i+1}]$, where the interest payment occurs, by no-arbitrage arguments, we have

$$\bar{V}(s, b, t_{i+1}^-) = \bar{V}(s, be^{\mathcal{R}(b)\Delta t}, t_{i+1}) \tag{B.3}$$

3. Over the time period $[t_{i+1}, t_{i+1}^+]$, an optimal rebalancing of the portfolio stipulates that

$$\bar{V}(s, b, t_{i+1}) = \min\left[\bar{V}(s, b, t_{i+1}^+), \min_{B+} \bar{V}(S^+(s, b, B^+), B^+, t_{i+1}^+)\right] \tag{B.4}$$

**Remark B.1.** *Combining* (B.3) *and* (B.4) *gives*

$$\bar{V}(s, b, t_{i+1}^-) = \min\left[\bar{V}(s, be^{\mathcal{R}(b)\Delta t}, t_{i+1}^+), \min_{B+} \bar{V}(S^+(s, be^{\mathcal{R}(b)\Delta t}, B^+), B^+, t_{i+1}^+)\right] \tag{B.5}$$

*Equation* (B.5) *can essentially be viewed as the optimization problem originating from an optimal rebalancing which occurs at time* $t_{i+1}$.

## B.2 A derivation of (4.9)

Let

$$\tau_-^n = T - t_{i+1}^+ \ ; \ \tau_+^n = T - t_{i+1}^- \ ; \ \tau_-^{n+1} = T - t_i^+ \ ; \ \tau_+^{n+1} = T - t_i^- \ . \tag{B.6}$$

Assume that we want to proceed from the discrete time $\tau_-^n$ to $\tau_-^{n+1}$. This can essentially be split into two steps. In the first step, we proceed from $\tau_-^n$ to $\tau_+^n$, and this step involves solving an optimization problem originated from an optimal rebalancing of the portfolio occurring at time $\tau^n$. The second step involves solving the model PDE from $\tau_+^n$ to $\tau_-^{n+1}$ with the initial condition obtained from the first step.

For the first step, we make use of Remark B.1. More specifically, with the discretization notation described in Section 4, the optimization problem in (B.5) essentially becomes

$$V_h(s_i, b_j, \tau_+^n) = \min \left[ V_h(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, \tau_-^n), \min_{B^+ \in \mathcal{Z}_h} V_h(S^+(s_i, b_j e^{\mathcal{R}(b_j)\Delta\tau}, B^+), B^+, \tau_-^n) \right] .$$

In the second step, taking into account (B.2), we need to solve the PIDE

$$V_\tau - \mathcal{P}V - \mathcal{J}V = 0,$$

from $\tau_+^n$ to $\tau_-^{n+1}$ with $V_h(s_i, b_j, \tau_+^n)$ as the initial condition. Using fully implicit timestepping, we have

$$V_h(s_i, b_j, \tau_-^{n+1}) - \Delta\tau P_h V_h(s_i, b_j, \tau_-^{n+1}) - \Delta\tau(\mathcal{J}_\ell)_h V_h(s_i, b_j, \tau_-^{n+1}) = V_h(s_i, b_j, \tau_+^n) ,$$

which is equivalent to (4.9).

# References


[1] X.Y. Zhou and D. Li. Continuous time mean variance portfolio selection: A stochastic LQ framework. *Applied Mathematics and Optimization*, 42:19–33, 2000.

[2] D. Li and W.-L. Ng. Optimal dynamic portfolio selection: Multiperiod mean variance formulation. *Mathematical Finance*, 10:387–406, 2000.

[3] M. Leippold, F. Trojani, and P. Vanini. A geometric approach to mulitperiod mean variance optimization of assets and liabilities. *Journal of Economic Dynamics and Control*, 28:1079–1113, 2004.

[4] T. Bielecki, S. Pliska, and X.Y. Zhou. Continuous time mean-variance portfolio selection with bankruptcy prohibition. *Mathematical Finance*, 15:213244, 2005.

[5] J. Wang and P.A. Forsyth. Numerical solution of the Hamilton-Jacobi-Bellman formulation for continuous time mean variance asset allocation. *Journal of Economic Dynamics and Control*, 34:207–230, 2010.

[6] E. Vigna. On the efficiency of mean-variance based portfolio selection in DC pension schemes. Working Paper, University of Torino, 2011.

[7] T. Bjork and A. Murgoci. A general theory of Markovian time inconsistent stochastic control problems. Available at SSRN: http://ssrn.com/abstract=1694759, 2010.

[8] S. Basak and G. Chabakauri. Dynamic mean-variance asset allocation. *The Review of Financial Studies*, 23:2970–3016, 2010.

[9] J. Wang and P.A. Forsyth. Comparison of mean variance like strategies for optimal asset allocation problems. *International Journal of Theoretical and Applied Finance*, 15:2, 2012. DOI: 10.1142/S0219024912500148.

[10] X. Cui, D. Li, S. Wang, and S. Zhu. Better than dynamic mean-variance: time-inconsistency and free cash flow stream. *Mathematical Finance*, 22:346–378, 2012.

[11] N. Bauerle. Benchmark and mean-variance problems for insurers. *Mathematical Methods of Operations Research*, 62:159–162, 2005.

[12] L. Delong and R. Gerrard. Mean-variance portfolio selection for a non-life insurance company. *Mathematical Methods of Operations Research*, 66:339–367, 2007.

[13] L. Delong, R. Gerrard, and S. Haberman. Mean-variance optimization problems for an accumulation phase in a defined benefit plan. *Insurance: Mathematics and Economics*, 42:107–118, 2008.

[14] R. Jose-Fombellida and J. Rincon-Zapatero. Mean variance portfolio and contribution selection in stochastic pension funding. *European Journal of Operational Research*, 187:120–137, 2008.

[15] B. Oksendal and A. Sulem. *Applied Control of Jump Diffusions*. Springer, 2009.

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

[17] R. Seydel. Existence and uniqueness of viscosity solutions for qvi associated with impulse control of jump-diffusions. *Stochastic Processes and Their Applications*, 119:3719–3748, 2009.

[18] R.C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.

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

[20] C. Fu, A. Lari-Lavassani, and X. Li. Dynamic mean variance portfolio selection with a borrowing constraint. *European Journal of Operational Research*, 200:312–319, 2010.

[21] S.T. Tse, P.A. Forsyth, and Y. Li. Preservation of scalarization optimal points in the embedding technique for the continuous time mean variance optimization. Working Paper, Cheriton School of Computer Science, University of Waterloo, 2012.

[22] V. I. Zakamouline. A unified approach to portfolio optimization with linear transaction costs. *Mathematical Methods of Operations Research*, 62:319–343, 2005.

[23] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer, 2009.

[24] Y. d'Halluin, P.A. Forsyth, and K.R. Vetzal. Robust numerical methods for contingent claims under jump diffusion processes. *IMA Journal of Numerical Analysis*, 25:87–112, 2005.

[25] G. Barles. Solutions de viscosité et équations elliptiques du deuxième order. Lecture Notes Université de Tours, 1997.

[26] M. H. A. Davis, X. Guo, and G. Wu. Impulse control of jump diffusions. *SIAM Journal on Control and Optimization*, 48:5276–5293, 2010.

[27] E. Jakobsen. Monotone schemes. In R. Cont, editor, *Encyclopedia of Quantitative Finance*, pages 1253–1263. Wiley, New York, 2010.

[28] J. Wang and P.A. Forsyth. Maximal use of central differencing for Hamilton-Jacobi-Bellman PDEs in finance. *SIAM Journal on Numerical Analysis*, 46:1580–1601, 2008.

[29] Y. Huang and P.A. Forsyth. Analysis of a penalty method for pricing a guaranteed minimum withdrawal benefit (GMWB). *IMA Journal of Numerical Analysis*, 32:320–351, 2012.

[30] P. A. Forsyth and G. Labahn. Numerical methods for controlled Hamilton-Jacobi-Bellman PDEs in finance. *Journal of Computational Finance*, 11 (Winter):1–44, 2008.

[31] J.F. Navas. On jump diffusion processes for asset returns. Working paper, Instituto de Empresa, 2000.

[32] R.J. Elliot, T.K. Siu, and A. Badescu. Bond valuation under a discrete-time regime-switching term-structure model and its continuous-time extension. Working paper, University of Calgary, 2011.

[33] Y. Zweng and Z. Li. Asset liability management under benchmark and mean-variance criteria in a jump diffusion market. *Journal of Systems Science and Complexity*, 24:317–327, 2011.

[34] A. E. B. Lim. Mean variance hedging when there are jumps. *SIAM Journal on Control and Optimization*, 44:1893–1922, 2005.