Pérez López, Iker and Hodge, David and Le, Huiling (2016) MDP algorithms for wealth allocation problems with defaultable bonds. Advances in Applied Probability, 48 (2). ISSN 1475-6064 (In Press)

This is a pre-peer-reviewed version of an article to appear in Advances in Applied Probability 48.2 (June 2016)

MDP ALGORITHMS FOR WEALTH ALLOCATION PROBLEMS WITH DEFAULTABLE BONDS

Iker Perez, David Hodge and Huiling Le

University of Nottingham

Abstract

This paper is concerned with analysing optimal wealth allocation techniques within a defaultable financial market similar to Bielecki and Jang (2007). It studies a portfolio optimization problem combining a continuous-time jump market and a defaultable security; and presents numerical solutions through the conversion into a Markov decision process and characterization of its value function as a unique fixed point to a contracting operator. This work analyses allocation strategies under several families of utilities functions, and highlights significant portfolio selection differences with previously reported results.

Keywords: portfolio optimization; defaultable bonds; markov decision processes;

2010 Mathematics Subject Classification: Primary 91G10

Secondary 90C40

1. Introduction

Let T be a finite time horizon and denote by $X = (X_t)_{t\geq 0}$ a continuous-time stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. Assume that X describes the evolution of a wealth process dependent on an allocation *strategy* or *policy*, taking values on a set Π. This paper is concerned with the study of a variation of a *portfolio optimization* problem of the form

$$V(t,x) = \sup_{\pi\in\Pi} \mathbb{E}[U(X_T^\pi)|X_t^\pi = x] \ , \tag{1}$$

for all $(t,x) \in [0,T] \times \mathbb{R}_+$. Here, the supremum is taken over all admissible policies in Π and function U is the *utility* determining a certain performance criterion.

Research within the field of portfolio optimization was triggered during the late 60s with the work of Merton [16], who made use of stochastic control techniques to maximize expected discounted utilities

* Postal address: Iker Perez, School of Mathematical Sciences, University of Nottingham. Email: ikertxo1986@gmail.com, iker.perez@nottingham.ac.uk.

of consumption. Later, his work was extended to different default-free frameworks where market uncertainty was mainly modelled by continuous processes with Brownian components, such work includes those of Fleming and Pang [11], Karatzas and Shreve [13] and Pham [17], among others. In the last decade, however, it is the optimal investment linked to defaultable claims that has attracted major attention. High yield corporate bonds offer attractive risk-return profiles and have become popular in comparison to stocks or default-free bonds; recent work in this area includes those of Bielecki and Jang [6], Bo *et. al.* [8], Lakner and Liang [15] and Capponi and Figueroa-Lopez [9].

Bielecki and Jang [6] first considered a market including a defaultable bond, a risk-free account and a stock driven by Brownian dynamics, and analysed optimal asset allocations for a variation of problem (1) with a risk averse CRRA utility, given by

$$V(t,x,h) = \sup_{\pi\in\Pi} \mathbb{E}\Big[\frac{(X_T^\pi)^\gamma}{\gamma}\Big| X_t^\pi = x, H_t = h\Big] \ , \quad \text{with } 0<\gamma<1,$$

for all $(t,x,h)\in[0,T]\times\mathbb{R}_+\times\{0,1\}$; here h denotes the current value of a *default process* $H=(H_t)_{t\geq 0}$ that models the state of the defaultable bond under the intensity based approach to credit risk (see Bielecki and Rutkowski [7]). For this matter, the authors assumed constant parameters governing the system and default intensity, and derived closed form solutions for the optimal allocations, pointing out that investments on the defaultable security are only justified under the presence of reasonable interest premiums. In addition, the results allocated a constant fraction of wealth in the Brownian asset, in a similar fashion to Merton [16].

Bo *et. al.* [8] approached a perpetual allocation problem for an investor with logarithmic utility, considering a defaultable perpetual bond along with a traditional stock and a risk-free account in a similar manner to Bielecki and Jang [6]. Their work modelled stochastically the intensities and premium process including a common Brownian factor, and made use of heuristic arguments in order to postulate the price dynamics of the defaultable bond, instead of arbitrage-free arguments. Their results established, in the same fashion to that of Bielecki and Jang [6], monotonicity conditions on the optimal investment on defaultable bonds with respect to the risk premium and recovery of wealth at default.

Lakner and Liang [15] employed duality theory to obtain similar optimal allocation strategies in a 2-way market, including a continuous-time money market account and a defaultable bond whose prices can jump; and Capponi and Figueroa-Lopez [9] extended the work in Bielecki and Jang [6] to a defaultable market with different economical regimes, where a defaultable bond, a money market and a stock are all dependent on a finite state continuous-time Markov process $Y=(Y_t)_{t\geq 0}$; in their work they obtained the explicit solution to the optimization problem

$$V(t,x,h;y) = \sup_{\pi\in\Pi} \mathbb{E}[U(X_T^\pi)|X_t^\pi = x, H_t = h, Y_t = y]$$

with logarithmic and risk averse CRRA utilities, for all $(t,x,h) \in [0,T]\times\mathbb{R}_+\times\{0,1\}$ and market regimes $y \in \{y_1, ...y_N\}$, with $N > 0$. Their numerical economic analysis highlighted the preference of investors to buy defaultable bonds when the macroeconomic regimes yields high expected returns and the planning horizon is large.

Results in the literature do however primarily relate to markets incorporating Brownian-driven assets and are limited with regards to the choices of utility functions that they provide solutions for. This work incorporates the presence of a defaultable bond in a finite horizon market with a bank account and a continuous-time jump asset driven by a *piecewise deterministic Markov process* as introduced in Almudevar [1]. In this circumstance, it is possible to build a bridge between a problem formulated in continuous-time and the theory of discrete-time Markov decision processes (MDPs), reducing the optimization problem to a discrete-time model by considering an embedded state process. Similar financial markets, in absence of the defaultable claim, have previously been explored by Kirch and Runggaldier [14] and Bäuerle and Rieder [2]. Kirch and Runggaldier [14] presented an algorithm for the evaluation of hedging strategies for European claims, addressing the optimization problem

$$V(t,x,s) = \min_{\pi\in\Pi} \mathbb{E}[l(F(S_T) - x - \int_t^T \pi_s \mathrm{d}S_s)|X_t^\pi = x, S_t = s] \ ,$$

which aims to minimize the expected value of a convex loss function l of the hedging error of a claim with payoff F, for all $(t,x,s) \in [0,T]\times\mathbb{R}_+^2$. Here, S is an asset whose dynamics are driven by a geometric Poisson process and X^π is the available capital under π. Strategies in Π are given by units held in the risky asset at different times.

On the other hand, Bäuerle and Rieder [2] considered the general portfolio utility maximization problem (1). In their case, the wealth process X reflects the evolution of wealth in a portfolio mixing a bank account and a generalized family of pure jump models; in addition, utility U is any increasing concave function. The authors make use of the embedding procedure previously explored by Almudevar [1] in order to convert the problem into a discrete-time MDP, and offer a proof for the validity of value iteration and policy improvement algorithms to approximate optimal allocation policies.

This paper makes use of the results on credit risk presented in Bielecki and Rutkowski [7] along with the theory for MDPs reviewed in Putterman [18] and Bäuerle and Rieder [4]; and extends the work of Bäuerle and Rieder [2, 3] to the context of defaultable markets explored by Bielecki and Jang, Bo *et. al.*, Lakner and Liang, Capponi and Figueroa-Lopez [6, 8, 15, 9] and references therein. By means of a conversion of the optimization problem into a MDP, its value function is characterized as the unique fixed point to a dynamic programming operator and optimal wealth allocations are numerically approximated through value iteration. This enables us to present a computational methodology which overcomes the need to assume any particular form for the utility function on the above mentioned classic optimization problems. Furthermore, this provides means of analyzing portfolio strategies incorporating illiquid markets. Thus, it allows for us to undertake a numerical analysis exploring the dependence of portfolio selections on the *risk premium* and different parameters describing the system. In doing so, we are able to examine extensions of the work in [6, 8, 15, 9] to more general families of logarithmic and exponential utility functions. The results highlight the nature of the significantly different allocation procedures under an exponential family of utilities, and the existence of a dependency on optimal stock allocation to default event, in a model with short selling restrictions. In order for the presented procedure to hold, default intensities and interest rates are assumed constant in a similar manner to that in Bielecki and Jang [6]. However, an extension to Markov modulated regimes similar to Capponi and Figueroa-Lopez [9] is discussed in the closing section.

The paper is organised as follows. Section 2 provides an introduction to the market analysed and presents the problem of interest. Sections 3 derives the infinitesimal dynamics of the financial products in the market and characterizes the evolution of the joint wealth process within the optimization problem. Sections 4 and 5 are concerned with validating a procedure in order to introduce an equivalent MDP to our optimization problem, and present the main technical results in the paper. Sections 6 and 7 will provide the proof for our results and justify the use of value iteration techniques in order to approximate optimal solutions. Finally, Sections 8 and 9 present a numerical analysis and make comments on optimal portfolio strategies, drawing comparisons with previous results that lead to the key contributions of this work. In addition, possible extensions of the model and drawbacks of this approach are discussed.

2. Introduction to the Market and Formulation of the Problem

Let $(\Omega, \mathcal{G}, \mathbb{P})$ denote a complete probability space equipped with a filtration $\{\mathcal{G}_t\}_{t\geq 0}$. Here $\mathbb{P}$ refers to the real world (also called historical) probability measure and $\{\mathcal{G}_t\}_{t\geq 0}$ is the *enlargement* of a *reference filtration* $\{\mathcal{F}_t\}_{t\geq 0}$ denoted $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$ and satisfying the usual assumptions of completeness and right continuity; $\mathcal{H}_t$ will be introduced later. We consider a frictionless financial market consisting of a risk-free bank account $B = (B_t)_{0\leq t\leq T}$, a pure-jump asset $S = (S_t)_{0\leq t\leq T}$ and a defaultable bond

$P = (P_t)_{0 \leq t \leq T}$. The dynamics of each of the components of the market are given as follows.

Risk free bank account. Let $B_0 = 1$ and $r > 0$ denote the market fixed-interest rate. The deterministic dynamics of B are given by

$$\mathrm{d}B_t = rB_t\mathrm{d}t \ .$$

Pure jump asset. Let $C = (C_t)_{0 \leq t \leq T}$ be a compound Poisson process defined on $(\Omega, \mathcal{G}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$, given by

$$C_t = \sum_{n=1}^{N_t} Y_n \ , \tag{2}$$

where $N = (N_t)_{0 \leq t \leq T}$ denotes a Poisson process with intensity $\nu > 0$ and $(Y_n)_{n \in \mathbb{N}}$ is a sequence of independent and identically distributed random variables, with $\mathbb{E}[Y_n] < \infty$, $Y_n \geq -1$ and distribution $\gamma(\mathrm{d}y)$. Here $\{\mathcal{F}_t\}_{t \geq 0}$ is a suitable complete and right-continuous filtration.

Asset S is a piecewise deterministic Markov process (cf. Almudevar [1]) adapted to $\mathcal{F}_t$ and is given by

$$\mathrm{d}S_t = S_{t^-}(\mu\mathrm{d}t + \mathrm{d}C_t) \ ,$$

where μ is the constant *appreciation rate* of the asset and $S_0 > 1$. Figure 1 illustrates some sample realisations of the process S.

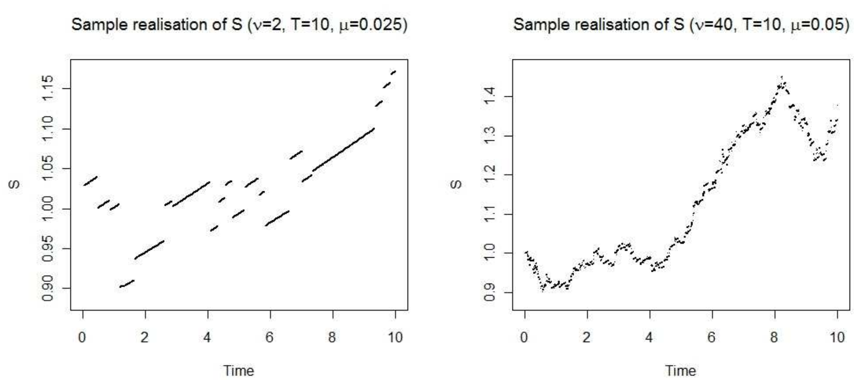

FIGURE 1: Sample realisations of the piecewise deterministic Markov process S, with varying parameters. On the left hand side $\nu = 2$, $T = 10$, $\mu = 0.025$; on the right hand side $\nu = 40$, $T = 10$, $\mu = 0.05$. Jumps Y follow truncated normal distributions.

Defaultable bond. We consider a tradeable zero coupon bond with face value of one unit and recovery at default. Let $\tau > 0$ be an exponentially distributed random variable defined on $(\Omega, \mathcal{G}, \{\mathcal{H}_t\}_{t \geq 0}, \mathbb{P})$

with intensity $\lambda_{\mathbb{P}}$; we make use of the intensity-based approach for modelling credit risk as introduced in Bielecki and Rutkowski [7] and let the τ model the default time of the bond P. Here $\mathcal{H}_t = \sigma(H_s : s \leq t)$ is the filtration generated by the one-jump process $H_t = 1_{\{\tau \leq t\}}$, after completion and regularization on the right; C_t and H_t, as well as $\mathcal{F}_t$ and $\mathcal{H}_t$, are assumed to be independent and $\lambda_{\mathbb{P}}$ denotes the *hazard rate* of τ, so that the compensated process

$$\mathrm{d}M_t = \mathrm{d}H_t - \lambda_{\mathbb{P}}\mathrm{d}(t \wedge \tau) \tag{3}$$

with $M_0 = 0$ is a $(\mathcal{G}_t, \mathbb{P})$-martingale, with $\mathcal{G}_t = \mathcal{F}_t \vee \mathcal{H}_t$. Lastly, we denote by $Z = (Z_t)_{0 \leq t \leq T}$ the $\mathcal{F}_t$-adapted *recovery process* of P, i.e. the process determining the wealth recovery upon default.

Then, the time-t price of this defaultable bond P with maturity at T is given by

$$P_t = B_t \mathbb{E}_{\mathbb{Q}}\Big[B_T^{-1}(1 - H_T) + \int_t^T B_u^{-1} Z_u \mathrm{d}H_u \Big| \mathcal{G}_t\Big] \ , \tag{4}$$

where $\mathbb{Q}$ is a martingale measure equivalent to $\mathbb{P}$. Intuitively, P_t models the discounted $\mathbb{Q}$-expected value of the pay-off $(1 - H_T) + H_T Z_\tau$. The existence of such an equivalent measure on $(\Omega, \mathcal{G})$ follows from the results on change of measures presented in Bielecki and Rutkowski [7] (Chapter 4).

Consider now an investor wishing to invest in this market. Denote by π_t^B the percentage of total wealth at time t invested on the risk-less bond; analogously π_t^S and π_t^P denote the time-t proportions on the asset and defaultable bond. The *portfolio process* $\pi = (\pi_t^B, \pi_t^S, \pi_t^P)_{0 \leq t \leq T}$ is a $\mathcal{G}_t$-predictable process taking values in

$$\mathcal{U} = \{(u_1, u_2, u_3) \in \mathbb{R}_+^3 : \sum_{i=1}^{3} u_i = 1\} \ , \tag{5}$$

so that short selling is not allowed and wealth is fully invested at all times and remains positive; in addition, note that $\pi_t^P = 0$ for $t > \tau$ is a must.

Denote by $X^\pi = (X_t^\pi)_{0 \leq t \leq T}$ the wealth process associated to a strategy $\pi \in \mathcal{U}$; its infinitesimal dynamics and explicit form are derived later. Also, let Π denote the family of all measurable portfolio processes π taking values in $\mathcal{U}$. In view of (1), for a given increasing and concave utility function $U : (0, \infty) \to \mathbb{R}^+$, let

$$V_\pi(t, x, h) = \mathbb{E}_{t,x,h}[U(X_T^\pi)] \tag{6}$$

denote the expected terminal reward associated to a portfolio strategy $\pi \in \Pi$, at time t and with values $X_t^\pi = x$ and $H_t = h$. Here, $\mathbb{E}_{t,x,h}$ denotes the expectation under the conditional probability measure $\mathbb{P}|_{(X_t^\pi = x, H_t = h)}$. This paper is concerned with identifying the optimal policy $\pi^* \in \Pi$ maximizing rewards

(6), so that

$$V_{\pi^*}(t,x,h) = \sup_{\pi\in\Pi} V_\pi(t,x,h)\ , \tag{7}$$

for all $(t,x,h) \in [0,T]\times\mathbb{R}^+\times\{0,1\}$. Note that $V_{\pi^*}(T,x,h) = U(x)$ for all $(x,h)\in\mathbb{R}^+\times\{0,1\}$ and problem V_{π^*} is tractable since $\mathbb{E}[Y_n] < \infty$.

3. The $\mathbb{P}$-Dynamics of the Defaultable Bond and Wealth Evolution

Following the results in Bielecki and Rutkowski [7] (Section 4.4) and Jeanblanc *et. al.* [12] (Section 8.6), let $\eta = \eta(\tau) = \phi e^{-\lambda_{\mathbb{P}}(\phi-1)\tau}$ be a random variable satisfying $\eta > 0$ and $\mathbb{E}_{\mathbb{P}}[\eta] = 1$, where ϕ is a strictly positive constant. Then, the change of measure with *Radon-Nikodým density process*

$$\eta_t = \frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\Big|_{\mathcal{G}_t} = \mathbb{E}_{\mathbb{P}}[\eta(\tau)|\mathcal{G}_t] = \mathbb{E}_{\mathbb{P}}[\eta(\tau)|\mathcal{H}_t]\ , \tag{8}$$

is such that τ is an exponentially distributed random variable under $\mathbb{Q}$, with intensity $\lambda_{\mathbb{Q}} = \phi\lambda_{\mathbb{P}}$. In practice, default intensities are independently estimated, using credit ratings and company data for the real world intensity $\lambda_{\mathbb{P}}$ and derivatives prices (including CDS and Options) for $\lambda_{\mathbb{Q}}$; their underlying ratio ϕ is named the 'Risk Premium' and represents the reward investors claim for bearing the risk of default in P.

Proposition 1. *The stochastic process η_t defined by* (8) *is a $(\mathcal{G}_t,\mathbb{P})$-martingale with $\eta_0 = 1$ and*

$$\mathrm{d}\eta_t = \eta_{t^-}(\phi-1)\mathrm{d}M_t\ ,$$

where M_t is defined by (3).

Proof. Expanding the conditional expectation in (8) we get

$$\begin{aligned}\eta_t &= \mathbb{E}_{\mathbb{P}}[\eta(\tau)|\mathcal{H}_t] = H_t\phi e^{-\lambda_{\mathbb{P}}(\phi-1)\tau} + (1-H_t)\int_t^\infty \phi e^{-\lambda_{\mathbb{P}}(\phi-1)x}\lambda_{\mathbb{P}} e^{-\lambda_{\mathbb{P}}(x-t)}\mathrm{d}x \\ &= H_t\phi e^{-\lambda_{\mathbb{P}}(\phi-1)\tau} + (1-H_t)e^{-\lambda_{\mathbb{P}}(\phi-1)t} \;=\; \phi^{H_t} e^{-\lambda_{\mathbb{P}}(\phi-1)(\tau\wedge t)}\ .\end{aligned}$$

Then, direct application of Itô's formula for non-continuous semi-martingales to η_t yields

$$\begin{aligned}\mathrm{d}\eta_t &= \eta_{t^-}(\phi-1)[\mathrm{d}H_t - (1-H_t)\lambda_{\mathbb{P}}\mathrm{d}t] \\ &= \eta_{t^-}(\phi-1)[\mathrm{d}H_t - \lambda_{\mathbb{P}}\mathrm{d}(t\wedge\tau)]\end{aligned}$$

proving the result.

In order to obtain the $\mathbb{P}$-dynamics of P defined by (4) we make use of the models for valuation of contingent claims subject to default risk in Duffie and Singleton [10]. We first define the concept of a *gain process*; we denote by $G = (G_t)_{0\leq t\leq T}$ the wealth gain process resulting from holding one defaultable bond P, given by

$$\mathrm{d}G_t = \mathrm{d}P_t + Z_t \mathrm{d}H_t \ , \tag{9}$$

with $G_0 = P_0$. Note that P and G differ in the sense that G incorporates the wealth recovered in case of default in P, so that $G_t = Z_\tau$ for $t \geq \tau$. In addition, we make the following assumption.

Assumption 1. (Recovery of Market.) *The wealth recovery upon default in P is given by a fraction of its current market value, i.e. $Z_t = (1-L)P_{t^-}$ for all $t < T$, with $0 \leq L \leq 1$ constant.*

Lemma 1. *The price dynamics of the defaultable bond P in* (4)*, under Assumption 1 and real world probability measure $\mathbb{P}$, are given by*

$$\mathrm{d}P_t = \begin{cases} P_{t^-}[(r+\phi\lambda_{\mathbb{P}}L)\mathrm{d}t - \mathrm{d}H_t] & \text{if} \quad t \leq T\wedge\tau, \\ 0 & \text{if} \quad \tau < t \leq T, \end{cases} \tag{10}$$

with $P_0 = e^{-(r+\phi\lambda_{\mathbb{P}}L)T}$.

Proof. The derivation of these equations follows from the application of Theorem 1 in Duffie and Singleton [10]. We use arbitrage-free arguments to obtain a pricing expression for P_t; the key observation is that its future expected gain G in (9), up to time $\tau\wedge T$, must match the attainable risk-less reward under measure $\mathbb{Q}$. That is, the discounted gain $e^{-rt}G_t$, given by

$$e^{-rt}G_t = e^{-rt}P_t + (1-L)\int_0^t e^{-rs}P_{s^-}\mathrm{d}H_s \tag{11}$$

for $t \in [0, \tau\wedge T]$, must be a $\mathbb{Q}$-martingale. Noting that $\mathbb{P}(\tau = T) = 0$ a.s., we may assume that default does not occur at maturity time. Recall from (4) that P is discontinuous only at the default time and that $P_t = 0$ for $t \geq \tau$, we may denote $P_t = (1-H_t)U_t$, where U_t is a continuous process. Plugging this expression for P into (11) above and applying Itô's formula we obtain

$$\mathrm{d}(e^{-rt}G_t) = e^{-rt}\Big[(1-H_{t^-})\mathrm{d}U_t - r(1-H_{t^-})U_{t^-}\mathrm{d}t - LU_{t^-}\mathrm{d}H_t\Big]$$

for $t \in [0, \tau\wedge T]$. It is possible to rewrite the above equation in terms of a compensated jump process through the inclusion and subsequent subtraction of a compensator in the jump differential term $\mathrm{d}H_t$, so that

$$\mathrm{d}(e^{-rt}G_t) = e^{-rt}\Big[(1-H_{t^-})(\mathrm{d}U_t - (r+\lambda_{\mathbb{Q}}L)U_{t^-}\mathrm{d}t) - LU_{t^-}\mathrm{d}M_t^{\mathbb{Q}}\Big] \ ,$$

where

$$\mathrm{d}M_t^{\mathbb{Q}} = \mathrm{d}H_t - \lambda_{\mathbb{Q}}\mathrm{d}(t \wedge \tau)$$

with $M_0^{\mathbb{Q}} = 0$ is a $(\mathcal{G}_t, \mathbb{Q})$-martingale. Therefore, for $e^{-rt}G_t$ to be a $\mathbb{Q}$-martingale the following must hold

$$\mathrm{d}U_t = (r + \phi\lambda_{\mathbb{P}}L)U_t\mathrm{d}t \ ,$$

since we recall that $\lambda_{\mathbb{Q}} = \phi\lambda_{\mathbb{P}}$. Finally, note that $\mathrm{d}P_t = \mathrm{d}U_t - U_{t^-}\mathrm{d}H_t$ and $P_t = U_t$ for $t < \tau$, the result follows.

We note that the price of P in (10) drops to zero at default; however, for portfolio optimization purposes we must account for the gain derived from its recovery value, so that we consider the $\mathbb{P}$-dynamics of the gain process $G = (G_t)_{0\leq t\leq T}$ in (9) for this purpose. We observe in (10) that the dynamics of G are determined by

$$\begin{aligned} \mathrm{d}G_t = G_{t^-}[(r + \phi\lambda_{\mathbb{P}}L)\mathrm{d}t - L\mathrm{d}H_t] \qquad & \text{for} \quad t \leq T \wedge \tau, \quad \text{and} \\ \mathrm{d}G_t = 0 \qquad & \text{for} \quad \tau < t \leq T \ , \end{aligned}$$

with $G_0 = P_0$. Thus, the time-t infinitesimal gain of a wealth process associated to a strategy $\pi \in \mathcal{U}$, denoted by $X^\pi = (X_t^\pi)_{0\leq t\leq T}$ in (6), is given by

$$\mathrm{d}X_t^\pi = X_{t^-}^\pi \cdot \left[(1 - \pi_t^P - \pi_t^S)\frac{\mathrm{d}B_t}{B_t} + \pi_t^S\frac{\mathrm{d}S_t}{S_{t^-}} + \pi_t^P\frac{\mathrm{d}G_t}{G_{t^-}}\right] .$$

The explicit form of X is derived using Itô calculus and is given by

$$X_t^\pi = X_0 e^{\int_0^t (r+\pi_s^S(\mu-r)+\pi_s^P\phi\lambda_{\mathbb{P}}L)\mathrm{d}s}(1 - \pi_\tau^P L)^{H_t}\prod_{n=1}^{N_t}(1 + \pi_{T_n}^S Y_n) \ , \tag{12}$$

where X_0 stands for the initial wealth.

4. A Discrete-Time Markov Decision Process

We follow a similar approach to that in Bäuerle and Rieder [2, 3] in order to first reduce problem (7) to a discrete-time MDP. This will allow for V_{π^*} to be computationally identified as the unique fixed point to a maximal reward operator.

Let $\Psi = (\Psi_n)_{n\geq 0}$ denote the increasing sequence of joint jump times in N and H, given by

$$\Psi_n = T_n 1_{\{T_n<\tau\}} + \tau 1_{\{T_{n-1}<\tau<T_n\}} + T_{n-1}1_{\{\tau<T_{n-1}\}} \ , \tag{13}$$

with $\Psi_0 = 0$. Intuitively, Ψ represents an ordered discrete counting process incorporating default time τ to jump times $(T_n)_{n\geq 0}$ in asset S. In addition, we refer to the counting steps $n \geq 0$ of Ψ as *decision epochs*. We define the MDP composed by the following 4-tuple $(E, \mathcal{A}, Q, R)$, an explanatory diagram is presented in Figure 2.

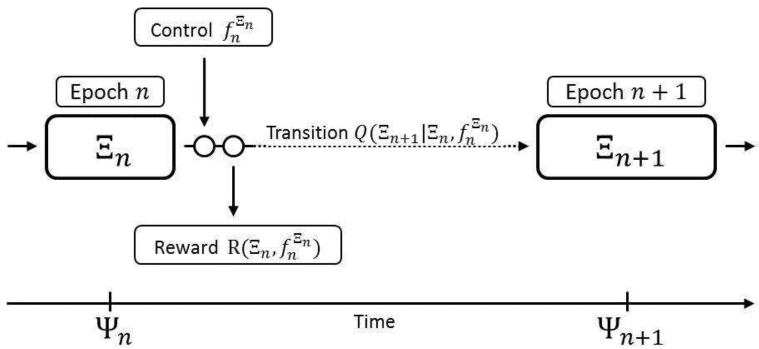

FIGURE 2: Explanatory diagram of the structure of the MDP $(E, \mathcal{A}, Q, R)$; variables Ξ_n and Ξ_{n+1} refer to the states of the system at epochs n and $n+1$ subsequently. We observe that each decision epoch n takes place at time Ψ_n.

The state space $\boldsymbol{E}$ is given by $E = [0, T] \times \mathbb{R}^+ \times \{0, 1\}$ and supports times Ψ_n, with associated wealth X_{Ψ_n} and states of default process H_{Ψ_n}, immediately after each jump. We use the notation Ξ_n to denote the n-th *state of the system*, given by

$$\Xi_n = \begin{cases} (\Psi_n, X_{\Psi_n}, H_{\Psi_n}) \in E & \text{if } \Psi_n \leq T \ , \\ \Delta & \text{otherwise} \ , \end{cases}$$

for $n \geq 0$. $\Delta \notin E$ is an external *absorption state* and allows for us to set up an *infinite horizon* optimization problem as described in Putterman [18] and Bäuerle and Rieder [4].

The action space $\boldsymbol{\mathcal{A}}$ stands for the set of deterministic *control* actions

$$\mathcal{A} = \{\alpha : \mathbb{R}^+ \to \mathcal{U} \text{ measurable}\} \ , \tag{14}$$

where $\mathcal{U}$ is given by (5). A control $\alpha \in \mathcal{A}$ is a function of time and $\alpha(t) \in \mathcal{U}$ determines the allocation of wealth at time t after a jump in Ψ. We note that for a given state $\Xi_n \in E \cup \{\Delta\}$ only a subclass of actions $D_n(\Xi_n) \subseteq \mathcal{A}$ may be admissible (for example, if bond P defaulted).

In addition to $\mathcal{A}$, we denote by F the set of all deterministic *policies* or *decision rules* given by

$$F = \{f : E \cup \{\Delta\} \to \mathcal{A} \text{ measurable}\} \ . \tag{15}$$

At any decision epoch n, a policy $f_n \in F$ maps a state Ξ_n to an admissible control action in $D_n(\Xi_n)$; we denote the resulting control by $f_n^{\Xi_n}$. The policy determines, as a function of the system state, the control chosen at epoch n. This, therefore results in a function $f_n^{\Xi_n} : \mathbb{R}^+ \to \mathcal{U}$ that models the time evolving allocation of wealth in our portfolio π, so that

$$\pi_t = f_n^{\Xi_n}(t - \Psi_n) \quad \text{for} \quad t \in [\Psi_n, \Psi_{n+1}) \ . \tag{16}$$

A portfolio process $\pi \in \Pi$ is called a *Markov portfolio strategy* if it is defined by a *Markov policy*, i.e. a sequence of functions $(f_n)_{n\geq 0}$ with $f_n \in F$. If policies $f_n \equiv f$ for all $n \geq 0$, the Markov policy is called *stationary*, implying that decisions are independent of the epoch number and only dependent on the system state. Figure 3 illustrates the characterization of such a Markovian portfolio strategy in a diagram. It is key to note that for a specified Markov policy, the controls to take at each epoch are random, since they depend on the system states to be observed.

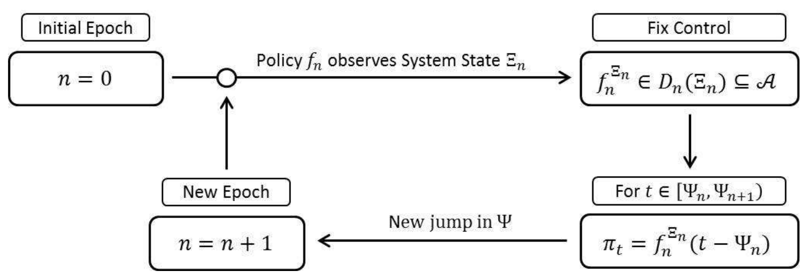

Figure 3: Characterization of a Markovian portfolio strategy $\pi \in \Pi$ defined by a Markov policy $(f_n)_{n\geq 0}$, with $f_n \in F$.

The transition probability Q. For current state $\Xi_n \in E$ and control $f_n^{\Xi_n} \in D_n(\Xi_n)$, the transition probability describes the probability for the system to adopt a specific state in epoch $n+1$ (or time Ψ_{n+1}). Let $f_n^{\Xi_n}(t) = (\alpha_t^B, \alpha_t^S, \alpha_t^P) \in \mathcal{U}$ denote the proportions of wealth allocated to each financial instrument at t time units after jump time Ψ_n, according to control $f_n^{\Xi_n}$; we note from (16) that this is equivalent to the global portfolio wealth allocation $\pi_{t+\Psi_n}$ at time $t + \Psi_n$. Analogously, let $\Gamma_t^{f_n^{\Xi_n}}$ denote the associated wealth at t time units after Ψ_n; this is equivalent to the global wealth $X_{t+\Psi_n}^{\pi}$ at time $t + \Psi_n$. Note from (12) that $\Gamma_t^{f_n^{\Xi_n}}$ is a deterministic function of the last system state, given by

$$\Gamma_t^{f_n^{\Xi_n}}(X_{\Psi_n}, H_{\Psi_n}) = X_{\Psi_n} e^{\int_0^t (r + \alpha_s^S(\mu - r))\mathrm{d}s} [H_{\Psi_n} + (1 - H_{\Psi_n}) e^{\int_0^t \alpha_s^P \lambda_{\mathbb{P}} L\phi \mathrm{d}s}] \ . \tag{17}$$

For an arbitrary $\Xi_n = (t', x, h)$, Lemma 7 in the Appendix shows that the transition probability Q is given by

$$\begin{aligned} Q(B|\Xi_n, f_n^{\Xi_n}) &= \mathbb{P}(\Xi_{n+1} \in B|\mathcal{G}_{\Psi_n}, f_n^{\Xi_n}) \\ &= \nu \int_0^{T-t'} e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s} \int_{-1}^{\infty} 1_B(t'+s, \Gamma_s^{f_n^{\Xi_n}}(x,h)(1+\alpha_s^S y), h)\gamma(\mathrm{d}y)\mathrm{d}s \\ &+ (1-h)\lambda_{\mathbb{P}} \int_0^{T-t'} e^{-(\nu+\lambda_{\mathbb{P}})s} 1_B(t'+s, \Gamma_s^{f_n^{\Xi_n}}(x,0)(1-\alpha_s^P L), 1)\mathrm{d}s \ , \end{aligned} \tag{18}$$

for $B \subseteq E$; in addition

$$Q(\{\Delta\}|\Xi_n, f_n^{\Xi_n}) = 1 - Q(E|\Xi_n, f_n^{\Xi_n}) \ .$$

Since Δ is an absorbing state we define $Q(\{\Delta\}|\Delta, \alpha) = 1$ for all controls $\alpha \in \mathcal{A}$. Intuitively, formula (18) gives the probability for the system state at epoch $n+1$ to fall within a subset B of the state space, given all information in $\mathcal{G}_{\Psi_n}$.

The reward function $\boldsymbol{R}$ is a function $R : E \times \mathcal{A} \to \mathbb{R}$ given by

$$R(t, x, h, \alpha) = e^{-(\nu+(1-h)\lambda_{\mathbb{P}})(T-t)} U(\Gamma^{\alpha}_{T-t}(x, h)) \ . \tag{19}$$

The adoption of such a non-negative reward function ensures the reducibility of optimization problem (7) to an infinite horizon discrete-time Markov decision process, as it will be shown in Lemma 2 below. We note that the term $e^{-(\nu+(1-h)\lambda_{\mathbb{P}})(T-t)}$ defines the likelihood of no jumps in a Poisson process with rate $\nu + (1-h)\lambda_{\mathbb{P}}$ over a period of time $T-t$, this will be a key observation in the proof of Lemma 2. In addition, we define $R(\Delta, \alpha) = 0$ for all $\alpha \in \mathcal{A}$.

For an arbitrary state $(t, x, h) \in E$, we let $v(t, x, h)$ denote the *optimal total expected reward* over all Markov policies $(f_n)_{n\geq 0}$ with $f_n \in F$, given by

$$v(t, x, h) = \sup_{(f_n)} \mathbb{E}_{t,x,h}\Big[\sum_{k=0}^{\infty} R(\Xi_k, f_k^{\Xi_k})\Big] \ . \tag{20}$$

5. Main Results

We now present the result on the equivalence between the portfolio optimization problem (7) and the MDP$(E, \mathcal{A}, Q, R)$.

Lemma 2. *For any $(t, x, h) \in E$, we have $V_{\pi^*}(t, x, h) = v(t, x, h)$, where $v(t, x, h)$ is defined by* (20).

Proof. We treat the case $t = 0$. The result at arbitrary time points can be proved similarly upon redefinition of terminal time $T' = T - t$ and adjustment of notation as pointed out in Bäuerle and Rieder

[4] (Chapter 8). Denote by Π_M the set of all Markovian portfolio strategies and note that $\Pi_M \subseteq \Pi$. Due to the Markovian structure of the state process the optimal strategy in (7) must be Markovian (cf. Bertsekas and Shreve [5]), so that

$$V_{\pi^*}(0,x,h) = \sup_{\pi\in\Pi} V_\pi(0,x,h) = \sup_{\pi\in\Pi_{\mathrm{M}}} \mathbb{E}_{x,h}[U(X_T^\pi)] \ ,$$

i.e. the supremum is attained in the set Π_M. Any $\pi \in \Pi_M$ is defined by a sequence of decision rules $f_n \in F$ forming a Markov policy $(f_n)_{n\geq 0}$ as described in (16). Therefore, for such a policy we need to show that

$$\mathbb{E}_{x,h}[U(X_T^\pi)] = \mathbb{E}_{x,h}\Big[\sum_{k=0}^{\infty} R(\Xi_k, f_k^{\Xi_k})\Big] \ .$$

For this, we note that

$$\begin{aligned}
\mathbb{E}_{x,h}[U(X_T^\pi)] &= \mathbb{E}_{x,h}\Big[\sum_{k=0}^{\infty} U(X_T^\pi)1_{\{\Psi_k\leq T<\Psi_{k+1}\}}\Big] \\
&= \sum_{k=0}^{\infty} \mathbb{E}_{x,h}\Big[\mathbb{E}_{x,h}\Big[U(X_T^\pi)1_{\{\Psi_k\leq T<\Psi_{k+1}\}}\Big|\mathcal{G}_{\Psi_k}\Big]\Big] \ ,
\end{aligned}$$

where Ψ is the non-decreasing counting process in (13) incorporating default time in H_t to jump times in N_t; we recall that these are $\mathcal{G}_t$-adapted processes with exponentially distributed jumps of intensities $\lambda_{\mathbb{P}}$ and ν. In view of (16) and (17) we note that wealth X^π can be expressed as a deterministic function of the previous system state, i.e.

$$X_t^\pi = \Gamma_{t-\Psi_k}^{f_k^{\Xi_k}}(X_{\Psi_k}, H_{\Psi_k}) \ ,$$

for $t \in [\Psi_k, \Psi_{k+1})$, with $X_0^\pi = x$. Therefore

$$\begin{aligned}
\mathbb{E}_{x,h}[U(X_T^\pi)] &= \sum_{k=0}^{\infty} \mathbb{E}_{x,h}\Big[\mathbb{E}_{x,h}\Big[U(\Gamma_{T-\Psi_k}^{f_k^{\Xi_k}}(X_{\Psi_k}, H_{\Psi_k}))1_{\{\Psi_k\leq T<\Psi_{k+1}\}}\Big|\mathcal{G}_{\Psi_k}\Big]\Big] \\
&= \sum_{k=0}^{\infty} \mathbb{E}_{x,h}\Big[U(\Gamma_{T-\Psi_k}^{f_k^{\Xi_k}}(X_{\Psi_k}, H_{\Psi_k}))\mathbb{P}(\Psi_{k+1} > T \geq \Psi_k|\mathcal{G}_{\Psi_k})\Big] \ .
\end{aligned}$$

In addition, we note that

$$\begin{aligned}
\mathbb{P}(\Psi_{k+1} > T \geq \Psi_k|\mathcal{G}_{\Psi_k}) &= 1_{\{T\geq\Psi_k\}}\mathbb{P}(\Psi_{k+1} > T|\mathcal{G}_{\Psi_k}) \\
&= 1_{\{T\geq\Psi_k\}}e^{-(\nu+(1-H_{\Psi_k})\lambda_{\mathbb{P}})(T-\Psi_k)} \ .
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{E}_{x,h}[U(X_T^\pi)] &= \sum_{k=0}^{\infty} \mathbb{E}_{x,h}\Big[1_{\{T\geq\Psi_k\}}e^{-(\nu+(1-H_{\Psi_k})\lambda_{\mathbb{P}})(T-\Psi_k)}U(\Gamma_{T-\Psi_k}^{f_k^{\Xi_k}}(X_{\Psi_k}, H_{\Psi_k}))\Big] \\
&= \sum_{k=0}^{\infty} \mathbb{E}_{x,h}\Big[R(\Xi_k, f_k^{\Xi_k})\Big] \ ,
\end{aligned}$$

completing the proof.

It has been shown that value function V_{π^*} in (7) can be derived as the sum of expected rewards v in (20). In what follows, we make use of the theory of MDPs exposed in Putterman [18] and Bäuerle and Rieder [4] and present results confirming the usefulness of iterative methods in order to approximate optimal portfolio strategies for our problem. The efforts are directed towards the construction of a complete metric space with a reward operator in a similar manner to Bäuerle and Rieder [2, 3], where V_{π^*} is identified as its fixed point.

Let $\mathbb{M}(E)$ define the set of measurable functions mapping the state space E into the positive subset of the real line, i.e.

$$\mathbb{M}(E) = \{g : E \to \mathbb{R}^+ : g \text{ measurable}\} \ .$$

We note that the *maximal reward operator* $\mathcal{T}$ for the MDP$(E, \mathcal{A}, Q, R)$ is a *dynamic programming* operator acting on $\mathbb{M}(E)$, such that

$$(\mathcal{T}g)(t,x,h) = \sup_{\alpha \in \mathcal{A}} \left\{ R(t,x,h,\alpha) + \sum_k \int g(s,y,k) Q(\mathrm{d}s, \mathrm{d}y, k | t,x,h,\alpha) \right\} \ ,$$

for all $g \in \mathbb{M}(E)$ and $(t,x,h) \in E$. Additionally, we denote by $(\mathcal{L}g)(t,x,h|\alpha)$ the term within brackets, i.e.

$$(\mathcal{L}g)(t,x,h|\alpha) = R(t,x,h,\alpha) + \sum_k \int g(s,y,k) Q(\mathrm{d}s, \mathrm{d}y, k | t,x,h,\alpha) \ , \tag{21}$$

and refer to it as the *reward operator*, so that

$$(\mathcal{T}g)(t,x,h) = \sup_{\alpha \in \mathcal{A}} (\mathcal{L}g)(t,x,h|\alpha) \ . \tag{22}$$

Now, let $\mathbb{C}_\vartheta(E)$ be the function space defined by

$$\mathbb{C}_\vartheta(E) = \{g \in \mathbb{M}(E) : g \text{ continuous and concave in } x \text{ and } \|g\|_\vartheta < \infty\} \ , \tag{23}$$

where

$$\|g\|_\vartheta = \sup_{(t,x,h) \in E} \frac{g(t,x,h)}{(1+x)e^{\vartheta(T-t)}} \ , \tag{24}$$

for fixed $\vartheta \geq 0$ satisfying conditions in Lemma 3 given in the next Section.

Theorem 1. *Operator $\mathcal{T}$ is a contraction mapping on the metric space $(\mathbb{C}_\vartheta(E), \|\cdot\|_\vartheta)$.*

Theorem 2. *There exists an optimal stationary portfolio strategy $\pi^* \in \Pi$, defined by a Markov policy $(f)_{n \geq 0}$ with $f \in F$ as shown in (16), so that V_{π^*} in (7) is the unique fixed point of $\mathcal{T}$ in $\mathbb{C}_\vartheta(E)$.*

Theorem 2 implies that a single decision rule $f : \Xi_n \to \mathcal{A}$ is optimal for all epochs $n \geq 0$, and the control chosen after each jump in Ψ is only dependent on the state of the system Ξ; we note that this incorporates information on time left to deadline, current wealth and event of default in P. Moreover, since V_{π^*} is characterized as a unique fixed point to a dynamic programming operator the use of computational approaches to approximate its value is justified.

6. Proof of Theorem 1

We begin with the presentation of a contraction result for later use. Let $\mathbb{M}_\vartheta(E)$ define the function space given by

$$\mathbb{M}_\vartheta(E) = \{g \in \mathbb{M}(E) : \ \|g\|_\vartheta < \infty\} \ ,$$

where the norm $\|\cdot\|_\vartheta$ is as in (24). We note from the theory in Putterman [18] (Chapter 7) that $\|\cdot\|_\vartheta$ is a *weighted supremum norm* and thus $\mathbb{M}_w(E)$ is a Banach space, since every Cauchy sequence of elements converges to an element in the set.

Lemma 3. *For sufficiently large $\vartheta \in \mathbb{R}_+$, $\|\mathcal{T}g_1 - \mathcal{T}g_2\|_\vartheta < \|g_1 - g_2\|_\vartheta$, for all $g_1, g_2 \in \mathbb{M}_\vartheta(E)$.*

Proof. For all $g_1, g_2 \in \mathbb{M}_\vartheta(E)$,

$$\begin{aligned}
(\mathcal{T}g_1 - \mathcal{T}g_2)(t,x,h) &\leq \sup_{\alpha\in\mathcal{A}}\{(\mathcal{L}g_1)(t,x,h|\alpha) - (\mathcal{L}g_2)(t,x,h|\alpha)\} \\
&= \sup_{\alpha\in\mathcal{A}}\Big\{\sum_k \int (g_1 - g_2)(s,y,k)Q(\mathrm{d}s,\mathrm{d}y,k|t,x,h,\alpha)\Big\} \\
&\leq \|g_1 - g_2\|_\vartheta \sup_{\alpha\in\mathcal{A}}\Big\{\sum_k \int (1+y)e^{\vartheta(T-s)}Q(\mathrm{d}s,\mathrm{d}y,k|t,x,h,\alpha)\Big\} \ .
\end{aligned}$$

Denote by I the expression within brackets on the right hand side. In view of (18), it reads

$$\begin{aligned}
I = \ & \nu\int_0^{T-t} e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s}\int_{-1}^{\infty}(1+\Gamma_s^\alpha(x,h)(1+\alpha^S y))e^{\vartheta(T-t-s)}\gamma(\mathrm{d}y)\mathrm{d}s \\
&+ (1-h)\lambda_{\mathbb{P}}\int_0^{T-t} e^{-(\nu+\lambda_{\mathbb{P}})s}(1+\Gamma_s^\alpha(x,0)(1-\alpha^P L))e^{\vartheta(T-t-s)}\mathrm{d}s \ .
\end{aligned}$$

Note that for all $(t,x,h,\alpha) \in E \times \mathcal{A}$ we have

$$1 + \Gamma_s^\alpha(x,h) < 1 + xe^{(2r+\mu)t+\lambda_{\mathbb{P}}L\phi} \leq k(1+x) \ ,$$

for some $k \in \mathbb{R}^+$. Therefore there exists a constant $c \in \mathbb{R}^+$ such that

$$1 + \Gamma_s^\alpha(x,h)(1-\alpha^P L) \leq c(1+x) \ ,$$

and

$$\int_{-1}^{\infty}(1+\Gamma_s^{\alpha}(x,0)(1+\alpha^S y))\gamma(\mathrm{d}y) = 1+\Gamma_s^{\alpha}(x,0)(1+\alpha^S\bar{y}) \leq c(1+x) \ ,$$

for all $x\in\mathbb{R}^+$, since $\bar{y}=\mathbb{E}[Y]<\infty$. Thus,

$$\begin{aligned} I &\leq c(1+x)e^{\vartheta(T-t)}\cdot\Big\{\nu\int_0^{T-t}e^{-(\nu+(1-h)\lambda_{\mathbb{P}}+\vartheta)s}\mathrm{d}s+(1-h)\lambda_{\mathbb{P}}\int_0^{T-t}e^{-(\nu+\lambda_{\mathbb{P}}+\vartheta)s}\mathrm{d}s\Big\} \\ &\leq c(1+x)e^{\vartheta(T-t)}(1-e^{-(\vartheta+\nu+\lambda_{\mathbb{P}})(T-t)})\Big(\frac{\nu}{\nu+\vartheta}+\frac{\lambda_{\mathbb{P}}}{\nu+\lambda_{\mathbb{P}}+\vartheta}\Big) \ . \end{aligned}$$

We note that there exists a constant $\vartheta\in\mathbb{R}_+$ sufficiently large such that

$$c_{\vartheta} = c(1-e^{-(\vartheta+\nu+\lambda_{\mathbb{P}})(T-t)})\Big(\frac{\nu}{\nu+\vartheta}+\frac{\lambda_{\mathbb{P}}}{\nu+\lambda_{\mathbb{P}}+\vartheta}\Big) < 1 \ .$$

Thus,

$$\|\mathcal{T}g_1-\mathcal{T}g_2\|_{\vartheta} = \sup_{(t,x,h)\in E}\frac{(\mathcal{T}g_1-\mathcal{T}g_2)(t,x,h)}{(1+x)e^{\vartheta(T-t)}} \leq \|g_1-g_2\|_{\vartheta}c_{\vartheta} < \|g_1-g_2\|_{\vartheta} \ ,$$

completing the proof.

Note that for all $(t,x,h,\alpha)\in E\times\mathcal{A}$

$$R(t,x,h,\alpha)\leq\mu(1+x)e^{\vartheta(T-t)} \quad \text{for some } \mu>0.$$

In addition, it follows from the proof of Lemma 3 that

$$\int_E(1+y)e^{\vartheta(T-t)}Q(\mathrm{d}y|t,x,h,\alpha) < (1+x)e^{\vartheta(T-t)} \ .$$

In view of these properties, $(1+x)e^{\vartheta(T-t)}$ is referred to as a *bounding function* of the MDP$(E,\mathcal{A},Q,R)$ (cf. Bäuerle and Rieder [4]); this ensures that the optimal total expected reward v in (20) is *well-defined*, i.e. $v<\infty$.

Now, since $\mathbb{C}_{\vartheta}(E)$ in (23) is a closed subset of $\mathbb{M}_{\vartheta}(E)$, the contracting property of $\mathcal{T}$ in Theorem 1 follows. However, we must provide proof for the concavity of the mapping $x\mapsto(\mathcal{T}g)(t,x,h)$, along with the continuity of $(t,x,h)\mapsto(\mathcal{T}g)(t,x,h)$, so that $(\mathcal{T}g)\in\mathbb{C}_{\vartheta}(E)$ for all $g\in\mathbb{C}_{\vartheta}(E)$; here, we do so separately.

6.1. The Proof of Concavity

Lemma 4. *For all $g\in\mathbb{C}_{\vartheta}(E)$, the mapping $x\mapsto(\mathcal{T}g)(t,x,h)$ is concave.*

Proof. We begin introducing the concept of *invested amounts*. In view of (17), at t time units after a last decision epoch in E with wealth x and default state h, a control action $\alpha \in \mathcal{A}$ with fractions $\alpha(t) \in \mathcal{U}$ for all $t \geq 0$ yields the wealth amounts $a(t) = \alpha(t)\Gamma_t^{\alpha}(x,h)$. It is therefore possible to define an alternative convex action space of invested amounts, given by

$$\mathbb{A}_{x,h} = \{a : \mathbb{R}_+ \to \mathbb{R}_+^3 \ : \ \sum_{i=1}^{3} a_i(t) = \Gamma_t^{\alpha}(x,h) \text{ for some } \alpha \in \mathcal{A}\} \ .$$

We denote by $\Gamma_t^a(x,h)$ the deterministic wealth evolution in time for a control $a \in \mathbb{A}_{x,h}$; in addition, we refer to controls a and α as being *equivalent* if $\Gamma_t^a(x,h) = \Gamma_t^{\alpha}(x,h)$.

The dynamics of $\Gamma_t^a(x,h)$ are expressed in terms of invested amounts and given by

$$\frac{\mathrm{d}\Gamma_t^a(x,h)}{\mathrm{d}t} = \Gamma_t^a(x,h)r + a_t^S(\mu - r) + (1-h)a_t^P\lambda_{\mathbb{P}}L\phi \ .$$

This is a first order linear differential equation and its general form solution is given by

$$\Gamma_t^a(x,h) = e^{rt}\Big(x + \int_0^t e^{-rs}[a_t^S(\mu - r) + (1-h)a_t^P\lambda_{\mathbb{P}}L\phi]\mathrm{d}t\Big) \ ,$$

which is a linear function on (x,a). For an arbitrary fixed $t' \geq 0$ and $h \in \{0,1\}$, fix wealths $x_1, x_2 \geq 0$ with $x_1 \neq x_2$ and set controls $\alpha_1, \alpha_2 \in \mathcal{A}$ so that for $i \in \{1,2\}$

$$(\mathcal{T}g)(t', x_i, h) = (\mathcal{L}g)(t', x_i, h|\alpha_i) \ ,$$

where operators $\mathcal{L}$ and $\mathcal{T}$ are given by (21) and (22) respectively. Now, choose equivalent controls $a_1 \in \mathbb{A}_{x_1,h}$ and $a_2 \in \mathbb{A}_{x_2,h}$ so that

$$a_1(t) = \alpha_1(t)\Gamma_t^{\alpha_1}(x_1,h) \quad \text{and} \quad a_2(t) = \alpha_2(t)\Gamma_t^{\alpha_2}(x_2,h) \ ,$$

for $t \geq 0$. Fix $\kappa \in (0,1)$ and let

$$x_3 = \kappa x_1 + (1-\kappa)x_2 \quad \text{and} \quad a_3 = \kappa a_1 + (1-\kappa)a_2 \ .$$

Note that $a_3 \in \mathbb{A}_{x_3,h}$ since $\sum_{i=1}^3 a_{3,i}(0) = x_3$. Hence,

$$(\mathcal{T}g)(t', x_3, h) = \sup_{\alpha \in \mathcal{A}}(\mathcal{L}g)(t, x_3, h|\alpha) = \sup_{a \in \mathbb{A}}(\mathcal{L}g)(t, x_3, h|a) \geq (\mathcal{L}g)(t, x_3, h|a_3) \ ,$$

with

$$\begin{aligned}(\mathcal{L}g)(t', x_3, h|a_3) = {} & e^{-(\nu+\lambda_{\mathbb{P}}(1-h))(T-t')}U(\Gamma_{T-t'}^{a_3}(x_3,h)) \\ & + (1-h)\lambda_{\mathbb{P}}\int_0^{T-t'} e^{-(\nu+\lambda_{\mathbb{P}})s}g(t+s, \Gamma_{t'}^{a_3}(x_3,h) - La_{3,s}^P, 1)\mathrm{d}s \\ & + \nu\int_0^{T-t'} e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s}\int_{-1}^{\infty} g(t+s, \Gamma_{t'}^{a_3}(x_3,h) + ya_{3,s}^S, 1)\gamma(\mathrm{d}y)\mathrm{d}s \ ,\end{aligned}$$

where $a^P_{3,s}$ and $a^S_{3,s}$ denote the wealth amounts invested in the defaultable bond P and stock S respectively $s \geq 0$ time units after t', according to control $a_3 \in \mathbb{A}_{x_3,h}$. We recall that $(x, a) \mapsto \Gamma^a_t(x, h)$ is a linear mapping, utility U is a concave function and g is concave on its second argument, so that

$$\begin{aligned}(\mathcal{T}g)(t', x_3, h) &\geq \kappa(\mathcal{L}g)(t', x_1, h|a_1) + (1-\kappa)(\mathcal{L}g)(t', x_2, h|a_2) \\ &= \kappa(\mathcal{T}g)(t', x_1, h) + (1-\kappa)(\mathcal{T}g)(t', x_2, h),\end{aligned}$$

completing the proof.

6.2. Enlargement of the Action Space

In order to show the continuity of the mapping $(t, x, h) \mapsto (\mathcal{T}g)(t, x, h)$, we will make use of the enlargement of the action space $\mathcal{A}$ in (14) to the set of *randomized controls* given by

$$\mathcal{R} = \{\rho : \mathbb{R}^+ \to \mathbb{P}(\mathcal{U}) \text{ measurable}\} ,$$

where $\mathbb{P}(\mathcal{U})$ defines the set of probability measures on the Borel subsets $\mathcal{B}(\mathcal{U})$ of the compact set $\mathcal{U}$ in (5). Such an enlargement of the action space is common in these circumstances as seen in Putterman [18], Bertsekas and Shreve [5] and Bäuerle and Rieder [4], and it will provide us with tools to obtain the desired result. We note that $\mathcal{A} \subseteq \mathcal{R}$, since all deterministic controls are attainable in $\mathcal{R}$ through the adoption of measures with single mass points. Also, we endow $\mathcal{R}$ with the *Young Topology* as proposed in Bäuerle and Rieder [2, 3]. This is the coarsest topology such that for a sequence of controls $(\rho_n)_{n\geq 1} \subset \mathcal{R}$ and fixed control $\rho \in \mathcal{R}$, $\lim_{n\to\infty} \rho_n = \rho$ if and only if

$$\lim_{n\to\infty} \int_0^T \int_{\mathcal{U}} g(t, u)\rho_{n,t}(\mathrm{d}u)\mathrm{d}t = \int_0^T \int_{\mathcal{U}} g(t, u)\rho_t(\mathrm{d}u)\mathrm{d}t \tag{25}$$

for all functions $g : [0, \infty] \times \mathcal{U} \to \mathbb{R}$ which are measurable in the first argument and continuous in the second, and satisfy

$$\int_0^\infty \max_{u\in\mathcal{U}} |g(t, u)|\mathrm{d}t < \infty . \tag{26}$$

Under the Young topology, $\mathcal{R}$ is a separable, metric and compact Borel space.

As a standard procedure, the functions given by (17) and (18) and defined on the set of deterministic Markovian controls $\mathcal{A}$ need to be extended to $\mathcal{R}$. For $\rho \in \mathcal{R}$, we define the infinitesimal wealth dynamics between jump times in (17) by

$$\mathrm{d}\Gamma^\rho_t(x, h) = \int_{\mathcal{U}} \Gamma^\rho_t(x, h)[r + u^S(\mu - r) + (1-h)u^P \lambda_{\mathbb{P}} L\phi]\rho_t(\mathrm{d}u)\mathrm{d}t ,$$

for all $(x,h) \in \mathbb{R}^+ \times \{0,1\}$, so that the randomized allocation of wealth is accounted for. We note that the above wealth evolution can be expressed in terms of a deterministic allocation

$$\Gamma_t^{\rho}(x,h) = \Gamma_t^{\bar{\rho}}(x,h) \ ,$$

with $\bar{\rho} \in \mathcal{A}$ defined by mean-average allocations according to control ρ, i.e $\bar{\rho}_t = \int_{\mathcal{U}} u\rho_t(\mathrm{d}u)$. On the other hand the transition probability Q in (18) extends to

$$\begin{aligned} Q(B|t,x,h,\rho) \ = \\ \nu \int_0^{T-t} e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s} \int_{-1}^{\infty} \int_{\mathcal{U}} 1_B(t+s, \Gamma_s^{\rho}(x,h)(1+u^S y), h)\rho_s(\mathrm{d}u)\gamma(\mathrm{d}y)\mathrm{d}s \\ + \ (1-h)\lambda_{\mathbb{P}} \int_0^{T-t} e^{-(\nu+\lambda_{\mathbb{P}})s} \int_{\mathcal{U}} 1_B(t+s, \Gamma_s^{\rho}(x,0)(1-u^P L), 1)\rho_s(\mathrm{d}u)\mathrm{d}s \ , \end{aligned} \tag{27}$$

where we recall $\gamma(\cdot)$ defines the density distribution of jumps Y in asset S. We note that, by definition, deterministic controls can perform no better than relaxed ones. Here, we introduce a result showing that, in fact, deterministic controls in $\mathcal{A}$ do perform as well as randomized ones in $\mathcal{R}$.

Lemma 5. *For all $g \in \mathbb{C}_{\vartheta}(E)$,*

$$(\mathcal{T}g)(t,x,h) = \sup_{\alpha\in\mathcal{A}} (\mathcal{L}g)(t,x,h|\alpha) = \sup_{\rho\in\mathcal{R}} (\mathcal{L}g)(t,x,h|\rho),$$

for all $(t,x,h) \in E$.

Proof. We recall that $\mathcal{A} \subseteq \mathcal{R}$, so that for all $g \in \mathbb{C}_{\vartheta}(E)$

$$\sup_{\alpha\in\mathcal{A}} (\mathcal{L}g)(t,x,h|\alpha) \leq \sup_{\rho\in\mathcal{R}} (\mathcal{L}g)(t,x,h|\rho) \ ,$$

for all $(t,x,h) \in E$. In addition, recall that for all $\rho \in \mathcal{R}$ we have $\bar{\rho} \in \mathcal{A}$, so that the result will follow from

$$(\mathcal{L}g)(t,x,h|\rho) \leq (\mathcal{L}g)(t,x,h|\bar{\rho}) \ ,$$

for all $\rho \in \mathcal{R}$. Now, note by (19) that $R(t,x,h,\rho) = R(t,x,h,\bar{\rho})$, since $\Gamma_t^{\rho}(x,h) = \Gamma_t^{\bar{\rho}}(x,h)$ by definition. In addition, any function $g \in \mathbb{C}_{\vartheta}$ is concave on its second argument, so that by Jensen's inequality we have

$$\int_{\mathcal{U}} g(t+s, \Gamma_s^{\rho}(x,h)(1+u^S y), h)\rho_s(\mathrm{d}u) \leq g(t+s, \Gamma_s^{\rho}(x,h)(1+\bar{\rho}^S y), h) \ ,$$

and

$$\int_{\mathcal{U}} g(t+s, \Gamma_s^{\rho}(x,0)(1-u^P L), 1)\rho_s(\mathrm{d}u) \leq g(t+s, \Gamma_s^{\rho}(x,0)(1-\bar{\rho}^P L), 1) \ ,$$

for all $(t, x, h) \in E$. Hence,

$$\begin{aligned}(\mathcal{L}g)(t,x,h|\rho) &= R(t,x,h,\rho) + \sum_k \int g(s,y,k) Q(\mathrm{d}s,\mathrm{d}y,\mathrm{d}u,k|t,x,h,\rho) \\ &\leq R(t,x,h,\bar{\rho}) + \sum_k \int g(s,y,k) Q(\mathrm{d}s,\mathrm{d}y,k|t,x,h,\bar{\rho}) \\ &= (\mathcal{L}g)(t,x,h|\bar{\rho}) \ ,\end{aligned}$$

completing the proof.

6.3. The Proof of Continuity

Lemma 6. *The mapping $(t, x, h) \mapsto (\mathcal{T}g)(t, x, h)$ is continuous, for all $g \in \mathbb{C}_\vartheta(E)$.*

Proof. Note that all sets in $\{0, 1\}$ are open and therefore it suffices to prove that $(t, x) \mapsto (\mathcal{T}g)(t, x, h)$ is continuous. In view of Lemma 5, we can make use of relaxed controls within $\mathcal{R}$, since

$$(\mathcal{T}g)(t,x,h) = \sup_{\rho\in\mathcal{R}} (\mathcal{L}g)(t,x,h|\rho).$$

We recall that $\mathcal{R}$ is a compact Borel space with respect to the *Young topology*. In view of the definition of $\mathcal{L}$ in (21) the proof follows from the continuity of the mappings $E \times \mathcal{R} \to \mathbb{R}$ given by

$$(t,x,\rho) \mapsto e^{-(\nu+(1-h)\lambda_{\mathbb{P}})(T-t)} U(\Gamma^{\rho}_{T-t}(x,h)) = R(t,x,h,\rho) \ , \tag{28}$$

and

$$(t,x,\rho) \mapsto \sum_k \int g(s,y,k) Q(\mathrm{d}s,\mathrm{d}y,k|t,x,h,\rho) \ , \tag{29}$$

for fixed $h \in \{0, 1\}$. Since utility U is a continuous function and the exponential term in (28) is continuous in time, continuity of mapping (28) reduces to showing that

$$(t,x,\rho) \mapsto \Gamma^{\rho}_{T-t}(x,h)$$

is continuous. From the definition of Γ in (17), it is sufficient to show that

$$\int_0^t \int_{\mathcal{U}} u^S(\mu - r)\rho_s(\mathrm{d}u)\mathrm{d}s \quad \text{and} \quad \int_0^t \int_{\mathcal{U}} u^P \lambda_{\mathbb{P}} L\phi\rho_s(\mathrm{d}u)\mathrm{d}s \tag{30}$$

are continuous in (t, ρ). Following the approach in Bäuerle and Rieder [2] (Prop. 4.3) we provide proof for the first integral expression in (30), the second is proved in a similar fashion. Let $(t_n, \rho_n)_{n\geq 1} \subset [0, T] \times \mathcal{R}$ be a sequence with $(t_n, \rho_n) \to (t, \rho)$ and, in order to ease notation let $\epsilon_{n,s}$ and ϵ_s denote

$$\epsilon_{n,s} = \int_{\mathcal{U}} u^S(\mu - r)\rho_{n,s}(\mathrm{d}u) \quad \text{and} \quad \epsilon_s = \int_{\mathcal{U}} u^S(\mu - r)\rho_s(\mathrm{d}u) \ .$$

Then,

$$\begin{aligned}\Big|\int_0^{t_n}\epsilon_{n,s}\mathrm{d}s-\int_0^t\epsilon_s\mathrm{d}s\Big| &\le \Big|\int_0^{t_n}\epsilon_{n,s}\mathrm{d}s-\int_0^t\epsilon_{n,s}\mathrm{d}s\Big|+\Big|\int_0^t\epsilon_{n,s}\mathrm{d}s-\int_0^t\epsilon_s\mathrm{d}s\Big|\\ &\le (\mu-r)|t_n-t|+\Big|\int_0^t\epsilon_{n,s}\mathrm{d}s-\int_0^t\epsilon_s\mathrm{d}s\Big|\ .\end{aligned}$$

Noting that function $u\mapsto g(t,u)=g(u)=u^S(\mu-r)$ satisfies (26), it follows from the characterization of convergence in $\mathcal{R}$ of (25) that

$$(\mu-r)|t_n-t|+\Big|\int_0^t\epsilon_{n,s}\mathrm{d}s-\int_0^t\epsilon_s\mathrm{d}s\Big|\xrightarrow{n\to\infty}0\ .$$

We now turn our attention to the mapping (29), we note from the definition of Q in (27) its continuity follows from that of functions

$$W_1(t,x,\rho)=\int_0^{T-t}e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s}\int_{-1}^{\infty}\int_{\mathcal{U}}g(t+s,\Gamma_s^{\rho}(x,h)(1+u^Sy),h)\rho_s(\mathrm{d}u)\gamma(\mathrm{d}y)\mathrm{d}s \tag{31}$$

and

$$W_2(t,x,\rho)=\int_0^{T-t}e^{-(\nu+\lambda_{\mathbb{P}})s}\int_{\mathcal{U}}g(t+s,\Gamma_s^{\rho}(x,0)(1-u^PL),1)\rho_s(\mathrm{d}u)\mathrm{d}s\ , \tag{32}$$

for a fixed $h\in\{0,1\}$. The following procedure proves the continuity of equation (31), that of (32) is proved in a similar fashion. We begin assuming that $g\in\mathbb{C}_\vartheta(E)$ is a bounded function and let $(t_n,x_n,\rho_n)_{n\ge0}\subset[0,T]\times\mathbb{R}^+\times\mathcal{R}$ be a sequence with $(t_n,x_n,\rho_n)\to(t,x,\rho)$. In order to ease notation let g_n' and g' denote functions defined by

$$g_n'(s,u)=g(t_n+s,\Gamma_s^{\rho_n}(x_n,h)(1+u^Sy),h)\quad\text{and}\quad g'(s,u)=g(t+s,\Gamma_s^{\rho}(x,h)(1+u^Sy),h)\ .$$

Then

$$\begin{aligned}&|W_1(t_n,x_n,\rho_n)-W_1(t,x,\rho)|\\ &\qquad\le\Big|\int_{T-t}^{T-t_n}e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s}\int_{-1}^{\infty}\int_{\mathcal{U}}g_n'(s,u)\rho_{n,s}(\mathrm{d}u)\gamma(\mathrm{d}y)\mathrm{d}s\Big|\\ &\qquad\quad+\int_0^{T-t}e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s}\int_{-1}^{\infty}\int_{\mathcal{U}}|g_n'(s,u)-g'(s,u)|\rho_{n,s}(\mathrm{d}u)\gamma(\mathrm{d}y)\mathrm{d}s\\ &\qquad\quad+\Big|\int_0^{T-t}e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s}\int_{-1}^{\infty}\int_{\mathcal{U}}g'(s,u)(\rho_{n,s}(\mathrm{d}u)-\rho_s(\mathrm{d}u))\gamma(\mathrm{d}y)\mathrm{d}s\Big|\ .\end{aligned}$$

Since g is a bounded function, the first term converges to 0 as $n\to\infty$. By the dominated convergence Theorem and the continuity of Γ and g, the second term also converges to 0 as $n\to\infty$. Finally, the third term converges to 0 due to the characterization of convergence in $\mathcal{R}$ in (25).

Now, we recall from (24) that for each $g \in \mathbb{C}_\vartheta$ there exists some constant $c_g \in \mathbb{R}^+$ satisfying $g(t,x,h) \leq c_g(1+x)e^{\vartheta(T-t)}$. Let $w(t,x,h) = g(t,x,h) - c_g(1+x)e^{\vartheta(T-t)}$ define a negative and continuous function. Then, there exists (cf. Bertsekas and Shreve [5], Lemma 7.14) a decreasing sequence of bounded functions $(w_n)_{n\geq 1}$ with $w_n \to w$ pointwisely. Therefore

$$W_n'(t,x,\rho) = \int_0^{T-t} e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s} \int_{-1}^{\infty} \int_{\mathcal{U}} w_n(t+s, \Gamma_s^\rho(x,h)(1+u^S y), h)\rho_s(\mathrm{d}u)\gamma(\mathrm{d}y)\mathrm{d}s$$

defines a bounded and decreasing sequence of continuous functions with

$$\begin{aligned} &W_n'(t,x,\rho) \to \\ &\quad W_1(t,x,\rho) - c_g \int_0^{T-t} e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s} \int_{-1}^{\infty} (1+\Gamma_s^\rho(x,h)(1+\bar{\rho}_s^S y))e^{\vartheta(T-t)}\gamma(\mathrm{d}y)\mathrm{d}s\ . \end{aligned} \tag{33}$$

as $n \to \infty$. Since the pointwise limit of non-increasing sequences of continuous functions is upper semicontinuous, it follows that the function on the right hand side of (33) is upper semicontinuous. In addition, the term

$$c_g \int_0^{T-t} e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s} \int_{-1}^{\infty} (1+\Gamma_s^\rho(x,h)(1+\bar{\rho}_s^S y))e^{\vartheta(T-t)}\gamma(\mathrm{d}y)\mathrm{d}s$$

is continuous. Therefore W_1 is upper semicontinuous. Taking $w(t,x,h) = -g(t,x,h) + c_g(1+x)e^{\vartheta(T-t)}$ lower semicontinuity of W_1 is achieved, proving the result on continuity for W_1 and completing the proof.

7. Proof of Theorem 2

We recall from Lemma 2 that the value of the original portfolio optimization problem (7) can be derived as the sum of expected rewards v in (20). Theorem 2 implies that, in addition, the value function V_{π^*} is characterized as a unique fixed point to a dynamic programming operator, so that the use of computational methods to approximate its value is justified. The main line of the proof is directed towards the use of Theorem 7.3.5 in Bäuerle and Rieder [4].

Proof. Proof of Theorem 2. We recall that the MDP$(E, \mathcal{A}, Q, R)$ is contracting and it holds that $v = V_{\pi^*}$. In addition, the Banach space $\mathbb{C}_\vartheta(E)$ is a closed subset of $\mathbb{M}_\vartheta(E)$ satisfying

i) $0 \in \mathbb{C}_\vartheta(E)$,

ii) $\mathcal{T} : \mathbb{C}_\vartheta(E) \to \mathbb{C}_\vartheta(E)$.

Thus, according to the Theorem 7.3.5 in Bäuerle and Rieder [4], the proof would follow from the existence, for all $g \in \mathbb{C}_\vartheta(E)$, of a deterministic policy $f \in F$ such that $\mathcal{T}g = \mathcal{T}_f g$, where

$$(\mathcal{T}_f g)(t,x,h) = R(t,x,h,f^{(t,x,h)}) + \sum_k \int g(s,y,k) Q(\mathrm{d}s, \mathrm{d}y, k | t,x,h,f^{(t,x,h)}) \ ,$$

for all $(t,x,h) \in E$.

It is known from a result in Bertsekas and Shreve [5] (Chapter 7) that there exists a randomized policy $f : E \to \mathcal{R}$ such that $\mathcal{T}g = \mathcal{T}_f g$ for all functions $g \in \mathbb{C}_\vartheta(E)$. Then, from Lemma 5 we see that the deterministic policy $\bar{f} : E \mapsto \mathcal{A}$, given by

$$\bar{f}_s^{(t,x,h)} = \int_{\mathcal{U}} u f_s^{(t,x,h)}(\mathrm{d}u) \in \mathcal{U}$$

for all $(t,x,h) \in E$, is measurable and satisfies $\mathcal{T}g = \mathcal{T}_{\bar{f}} g$, therefore completing the proof.

8. Numerical Analysis

Making use of the main results obtained in Section 5, we now present and analyse computational results to our discrete-time infinite-horizon optimization problem $(E, \mathcal{A}, Q, R)$ defined in (13)-(20), for different measures of risk aversion. Numerical approximations to optimal allocation strategies $\pi^* \in \Pi$, along with optimal values V_{π^*}, are obtained through the method of value iteration and justified by the result of Theorem 2. For this, we have made use of an homogeneous space discretization as introduced in Bäuerle and Rieder [2] (Section 5.3).

The equivalence result of Lemma 2 warrants the optimality of these strategies in the original portfolio optimization problem (7), where alterations on wealth allocations are only decided at times of jumps in the market (a jump in asset S or a default in P) and span as time-dependent allocation functions until the next market jump; these jumps are referred to as *epochs* within the context of the MDP. Thus, we take advantage of the flexibility of the method regarding the choice of utility function and, in view of the original problem, determine characteristics of optimal wealth allocation strategies under different families of utilities, as well as the impact of generalizing utilities towards risky investments. Additionally, we assess the influence on allocation strategies of the different parameters defining the model and, more importantly, the effect of the short selling restriction imposed on the original definition of the problem.

Numerical calculations in this section are undertaken with a set interest rate of $r = 0.05$. In addition, values such as jump intensities λ and ν, risk premium ϕ, loss at default L and appreciation rate of the stock μ are, unless otherwise stated, fixed to sensible positive values within a financial context. This is done using parameter choices for numerical simulations in Bielecki and Jang [6] and Cappini and Figueroa-Lopez [9] as a reference, therefore allowing for direct comparisons of our results with recent work on portfolio management with defaultable bonds, and establishing general properties on optimal strategies with respect to variations on utility functions and time, wealth and default state values.

The focus is on popular power, logarithmic and exponential utility measures of risk aversion. The *constant relative risk aversion* (CRRA) family of power utility functions is given by

$$U(x) = \frac{x^{1-c}}{1-c} \quad \text{for } 0 < c < 1,$$

so that the level of *relative risk aversion* is constant and given by $R(x) = -\frac{xU''(x)}{U'(x)} = c$. The logarithmic family of utility functions is on the other hand given by

$$U(x) = \log(x+c) \quad \text{for } c \in \mathbb{R}_+,$$

and its level of risk aversion is $R(x) = \frac{x}{x+c}$, so that it is a CRRA utility measure only if $c = 0$; if $c > 0$ this is an *increasing relative risk aversion* (IRRA) measure. Finally, the exponential family of measures is a popular *constant absolute risk aversion* (CARA) family given by

$$U(x) = 1 - \frac{e^{-cx}}{c} \quad \text{for } c \in \mathbb{R}_+,$$

so that the *absolute risk aversion* level is constant and given by $A(x) = -\frac{U''(x)}{U'(x)} = c$.

Figure 4 presents pre-default value functions under different choices of measures. We note that these are increasing in wealth and decreasing in time. In these cases, the optimal allocation strategies correspond to varying fractional distributions of wealth between the defaultable bond and the bank account; and the convergences in the grid have been in all cases achieved under 10 iterations, using an initial candidate V according to the strategy of investing all wealth in bond B.

8.1. Performance Analysis of Utility Functions

Optimal allocations under different utilities vary on time, wealth values and level of aversion towards risky investments. Under an exponential measure of constant absolute risk aversion, the level of optimal risky investments is highly dependent on wealth values; in this case, both π^P and π^S are decreasing functions of wealth for $x > \kappa$, with $\kappa \in \mathbb{R}_+$ small as observed in the case of a defaultable bond in Figure 5. In addition, the optimal allocation in P slightly increases as $t \to T$; this is opposed to previously reported optimal strategies under power and logarithmic utilities, where a mild increase of aversion towards the exposure to risky bonds is observed as time approaches deadline. Here, we also observe

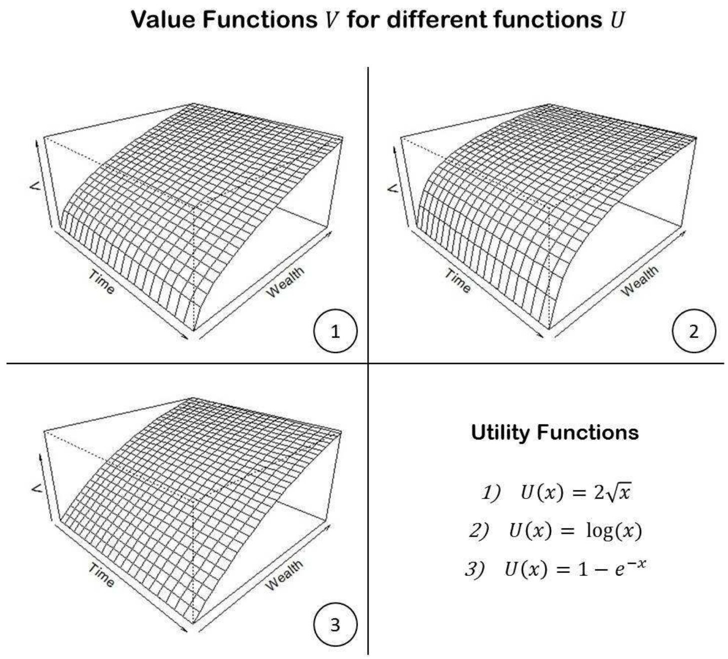

FIGURE 4: Approximation to pre-default V for different utility functions U. Results obtained through the method of value iteration with convergence in 10 iterations. $T = 1$, $r = \mu = 0.05$, $\lambda = 0.25$ $\phi = 1.3$, $L = 0.5$ and $\nu = 10$.

such aversion at times close to the deadline under power and logarithmic utilities, while remaining nearly time-invariant when the planning horizon is large. Additionally, the optimal wealth distribution remains invariant with regards to changes in wealth under these utilities. Certainly, as time approaches the deadline (and maturity in P under definition (4)) there exists an increase on the value of P and a decrease on the likelihood of default, implying that the defaultable bond gets relatively cheap only when the planning horizon is large.

Stock investments remain time-invariant under both power and logarithmic measures, consistent with the previous result. However, the short-selling restriction imposed to the portfolio optimization problem causes allocations π^S to remain invariant to a default event only if pre-default bond allocations π^B are strictly positive; if $\pi^B = 0$ at default time, both bond and stock percentage investments may increase following a default event in P. Figure 6 presents varying levels of the optimal percentage allocation π^S for varying values of the difference between the appreciation rate of the stock μ and the interest rate r under power utility functions $U(x) = \frac{x^{1-c}}{1-c}$, showing that this is a linearly increasing function on $\mu - r$

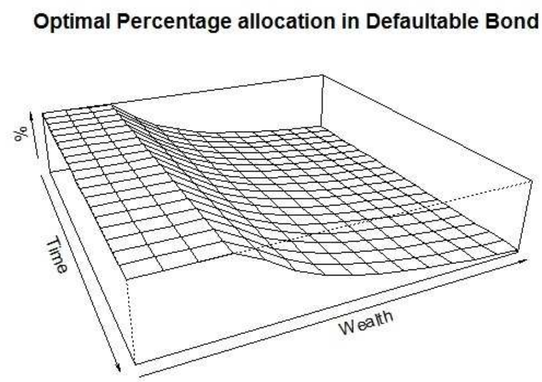

Figure 5: Optimal π^P, for $U(x) = 1 - e^{-x}$ and varying values of $t \in [0, T]$ and $x \geq 0$. Parameters $r = 0.05$, $\nu = 10$ and $\lambda = 0.25$

and a decreasing function on the level of constant relative risk aversion $R(x) = c$.

Moreover, we note in Figure 7 that for fixed $t \in [0, T]$ and wealth $x \in \mathbb{R}_+$, the value function V is such that $V(t, x, 0) \geq V(t, x, 1)$ for all $(t, x) \in [0, t] \times \mathbb{R}_+$. In addition, $V(t, x, 0) - V(t, x, 1)$ is decreasing in time and equal to 0 at $t = T$, a common feature under all utilities. Certainly, a default event decreases the dimensionality of the problem through a reduction in the choices of investment opportunities. Under exponential utilities and for $x > \kappa$, the losses in value are decreasing functions on wealth.

Finally, utilities analysed present common properties with regards to alterations on the values of several parameters defining the model. Optimal allocations π^P are increasing functions of the risk premium ϕ and decreasing functions of the loss value L at default, as illustrated in Figure 8 for a given pre-default state $(t, x, 0) \in E$ and utility $U(x) = 2\sqrt{x}$ in a two-bond market. A higher incentive for bearing risk in P motivates a higher investment; on the contrary, the opposite effect is caused by decreasing the return on recovery, despite the fact that it increases the yield on the bond. It is also never optimal to invest in a defaultable bond provided $\phi \leq 1$. In addition, optimal risky investments present a similar dependency on the level of aversion under different utilities; these are decreasing functions of the level of relative/absolute risk aversion, as observed in Figure 9 for a defaultable bond under power and exponential utilities.

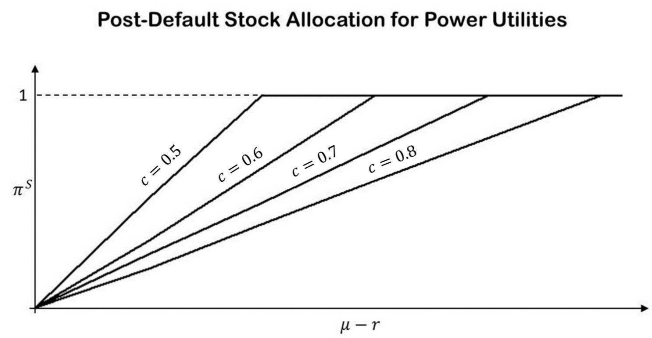

FIGURE 6: Optimal π^S after default, as a function of the distance between the appreciation and interest rate and for different power utility measures $U(x) = \frac{x^{1-c}}{1-c}$. Maximum allocation equals 1, since no short-selling is allowed. Here, $\lambda = 0.25$ $\phi = 1.3$, $L = 0.5$ and $\nu = 10$.

9. Discussion

We have presented an extension of results in Bäuerle and Rieder [2, 3] to the context of a defaultable market, in order to study optimal wealth allocation strategies for risk adverse investors, allowing for the use of broad families of utility functions. The original continuous-time portfolio optimization problem has been transformed into a discrete-time Markov decision process and its value function has been characterized as the unique fixed point to a dynamic programming operator, justifying the use of value iteration algorithms to provide the approximations of results of our interest.

The numerical analysis has been focused on the dependence of optimal portfolio selections on the risk premium, recovery of market value and several other parameters defining the model, and it has extended the scope of the results in [6, 8, 15, 9] to broader families of utility functions, highlighting relevant divergences on optimal strategies with respect to variations and generalizations in choices of utilities. In addition, the work has examined the impact of a short selling restriction within the market, identifying a dependency on optimal stock allocations with respect to default event on a corporate bond.

The analysis in Section 8 suggests that, similarly to [6, 8, 9], investments on defaultable bonds are only justified when the associated risk is correctly priced, measured in terms of risk premium coefficients ϕ. Also, similar monotonicity properties on optimal defaultable bond allocations have been identified in comparison to those presented in Bielecki and Jang [6] and Capponi and Figueroa-Lopez [9], under power and logarithmic utilities, so that these are decreasing on ϕ, increasing on L and there exists a reduction of the risk aversion as time approaches maturity; this work suggests that such properties

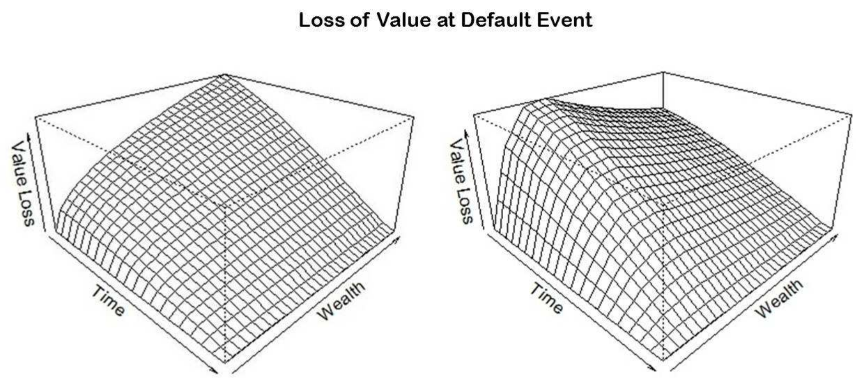

FIGURE 7: Approximation of the loss in V at default. Here $T = 1$, $r = \mu = 0.05$, $\lambda = 0.25$ $\phi = 1.3$, $L = 0.5$ and $\nu = 10$. On the left hand side $U(x) = \frac{\sqrt{x}}{2}$, on the right hand side $U(x) = 1 - e^{-x}$.

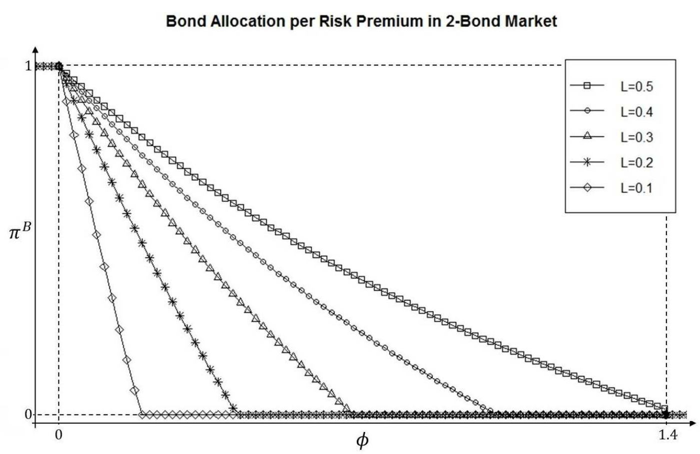

FIGURE 8: Approximation of pre-default π^B in a two-Bond market, for different risk premium ϕ and loss on default L. Parameters $r = 0.05$, $\nu = 10$, $\lambda = 0.25$ and utility $U(x) = 2\sqrt{x}$.

extend to generalizations of logarithmic utility functions. On the contrary, under exponential measures, there exists a slight increase in the risk aversion towards P in time, and optimal defaultable bond allocations are highly dependent on the wealth value and decreasing for $x > \kappa$, for some small $\kappa \in \mathbb{R}_+$. Additionally, we observed that in this case $V(t, x, 0) - V(t, x, 1)$ is decreasing on x for $x \geq \kappa$.

Furthermore, we have shown that the investment in the risky bond and stock is always prioritized as the levels of constant relative or absolute risk aversion are diminished. Also, optimal stock investments

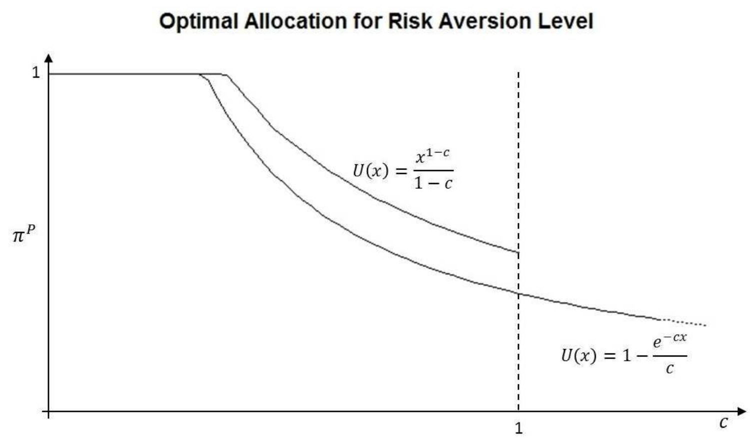

FIGURE 9: Optimal allocation π^P for utilities $U(x) = \frac{x^{1-c}}{1-c}$, $U(x) = 1 - \frac{e^{-cx}}{c}$ and varying values of $c \geq 0$ in a two-Bond market with fixed $(x, t, 0) \in E$. Parameters $r = 0.05$, $\nu = 10$ and $\lambda = 0.25$

have been identified as linear functions of the appreciation rate of the stock and interest rate, similarly to Merton [16]. However, unlike results reported in Bielecki and Jang [6] and Capponi and Figueroa-Lopez [9], a short-selling restriction has been identified to trigger a dependency on the allocation with respect to default event in P.

Finally, we note that the problem of considering a diversified portfolio involving multiple assets and defaultable bonds is a natural extension to this work, but it is not addressed in here to avoid technicalities part of extensive models. Other natural extensions of the model under the reduction to an MDP approach were pointed out in Bäuerle and Rieder [2]. These include the introduction of regime switching markets, where the different economical regimes are modelled by a continuous-time Markov chain $(I_t)_{t\geq 0}$ in a similar manner to Capponi and Figueroa-Lopez [9], so that parameters and coefficients defining the bank account, asset and defaultable bond vary according to the different states of I. In this scenario, the state space within the formulation of the MDP gains a degree of dimensionality, but the embedding procedure remains similar. In addition, models with partial information can be considered upon assuming that I is a *hidden* process and making use of filtering theory. Also, we note that this work has made rather strong assumptions regarding most parameters defining the model. The interest rate, stock appreciation rate, default intensities and loss on default rate have all considered constant. An extension to Brownian models for such parameters would not be tractable under the approach presented. However, the inclusion of different economical regimes as discussed above could present a more realistic case of study.

Appendix A. Transition probability Q

Let $f_n^{\Xi_n}(t) = (\alpha_t^B, \alpha_t^S, \alpha_t^P) \in \mathcal{U}$ denote the proportions of wealth allocated to each financial instrument at t time units after jump time Ψ_n, according to control $f_n^{\Xi_n}$. Analogously, let $\Gamma_t^{f_n^{\Xi_n}}$ in (12) denote the associated wealth t time units after Ψ_n.

Lemma 7. *For an arbitrary $\Xi_n = (t', x, h)$, the transition probability Q for the MDP $(E, \mathcal{A}, Q, R)$ in Section 4 is given by*

$$\begin{aligned} Q(B|\Xi_n, f_n^{\Xi_n}) &= \mathbb{P}(\Xi_{n+1} \in B|\mathcal{G}_{\Psi_n}, f_n^{\Xi_n}) \\ &= \nu \int_0^{T-t'} e^{-(\nu+(1-h)\lambda_{\mathbb{P}})s} \int_{-1}^{\infty} 1_B(t'+s, \Gamma_s^{f_n^{\Xi_n}}(x,h)(1+\alpha_s^S y), h)\gamma(\mathrm{d}y)\mathrm{d}s \\ &+ (1-h)\lambda_{\mathbb{P}} \int_0^{T-t'} e^{-(\nu+\lambda_{\mathbb{P}})s} 1_B(t'+s, \Gamma_s^{f_n^{\Xi_n}}(x,0)(1-\alpha_s^P L), 1)\mathrm{d}s \ , \end{aligned}$$

for $B \subseteq E$. In addition

$$Q(\{\Delta\}|\Xi_n, f_n^{\Xi_n}) = 1 - Q(E|\Xi_n, f_n^{\Xi_n}) \ .$$

Proof. For an arbitrary $\Xi_n = (t', x, h)$ at epoch n, the transition probability to a new state $\Xi_{n+1} \in E \cup \{\Delta\}$ at epoch $n+1$ is given by

$$Q(B|\Xi_n, f_n^{\Xi_n}) = \mathbb{P}(\Xi_{n+1} \in B|\mathcal{G}_{\Psi_n}, f_n^{\Xi_n}) = \mathbb{P}(\Xi_{n+1} \in B|\mathcal{G}_{t'}, f_n^{\Xi_n}) \ , \tag{34}$$

where $\mathcal{G}_{t'}$ intuitively denotes all the information in the system up to time t'. For $B \subseteq E$, the next epoch comes at the time of the first jump in either the asset S or the default process H (and always before the deadline T). We note that cases $h = 0$ and $h = 1$ need to be treated separately since in the latter there are no more jumps in H. Due to the Markovian structure of the problem, we rewrite (34) as

$$\begin{aligned} Q(B|\Xi_n, f_n^{\Xi_n}) &= \mathbb{P}(\Xi_{n+1} \in B|\Xi_n = (t', x, 1), f_n^{\Xi_n}) \cdot h \\ &\quad + \mathbb{P}(\Xi_{n+1} \in B|\Xi_n = (t', x, 0), f_n^{\Xi_n}) \cdot (1-h) \ , \end{aligned} \tag{35}$$

The first term on the right hand side of (35) is derived upon noting that the intensity of the poisson jump process N in (2) is ν, and that the distribution of the jumps $Y \geq -1$ is given by $\gamma(\mathrm{d}y)$. Under control $f_n^{\Xi_n}$, the percentage of wealth invested in asset S at any time t after t' is given by α_t^S. Analogously, the total wealth is given by $\Gamma_t^{f_n^{\Xi_n}}(x, 1)$, so that

$$\begin{aligned} &\mathbb{P}(\Xi_{n+1} \in B|\Xi_n = (t', x, 1), f_n^{\Xi_n}) \\ &= \int_0^{T-t'} \nu e^{-\nu s} \int_{-1}^{\infty} 1_B(t'+s, \Gamma_s^{f_n^{\Xi_n}}(x,1)(1+\alpha_s^S y), 1)\gamma(\mathrm{d}y)\mathrm{d}s \ . \end{aligned} \tag{36}$$

For the second term in (35) we consider the events

- $\mathcal{C}_1 =$ *"Next jump in Asset S arrives before jump in Default process H"* , and
- $\mathcal{C}_2 =$ *"Jump in Default process H arrives before next jump in Asset S"* ,

so that we can extend the above expression according to the laws of conditional probabilities, yielding

$$\begin{aligned}\mathbb{P}(\Xi_{n+1} \in B|\Xi_n = (t', x, 0), f_n^{\Xi_n}) &= \mathbb{P}(\Xi_{n+1} \in B|\Xi_n = (t', x, 0), f_n^{\Xi_n}, \mathcal{C}_1)\mathbb{P}(\mathcal{C}_1) \\ &\quad + \mathbb{P}(\Xi_{n+1} \in B|\Xi_n = (t', x, 0), f_n^{\Xi_n}, \mathcal{C}_2)\mathbb{P}(\mathcal{C}_2) \ .\end{aligned}$$

The jump intensity of H is given by $\lambda_{\mathbb{P}}$. Thus,

$$\mathbb{P}(\mathcal{C}_1) = \int_0^\infty \nu e^{-\nu s} \int_s^\infty \lambda_{\mathbb{P}} e^{-\lambda_{\mathbb{P}} r} \mathrm{d}r \mathrm{d}s = \frac{\nu}{\nu + \lambda_{\mathbb{P}}} \ ,$$

and analogously $\mathbb{P}(\mathcal{C}_2) = \frac{\lambda_{\mathbb{P}}}{\nu+\lambda_{\mathbb{P}}}$. In addition, we denote that ϕ_S and ϕ_H are the next jump times of S and H respectively, so that their conditional probability density functions $f_{\phi_S|\mathcal{C}_1}(\cdot|\mathcal{C}_1)$ and $f_{\phi_H|\mathcal{C}_1}(\cdot|\mathcal{C}_1)$ are given by

$$f_{\phi_S|\mathcal{C}_1}(\cdot|\mathcal{C}_1) = \frac{\mathrm{d}}{\mathrm{d}s} \frac{\mathbb{P}(S \leq s, \mathcal{C}_1)}{\mathbb{P}(\mathcal{C}_1)} = (\lambda_{\mathbb{P}} + \nu) e^{-(\nu+\lambda_{\mathbb{P}})s} = f_{\phi_H|\mathcal{C}_1}(\cdot|\mathcal{C}_1) \ .$$

Then, in a similar manner to (36), we have

$$\begin{aligned}&\mathbb{P}(\Xi_{n+1} \in B|\Xi_n = (t', x, 0), f_n^{\Xi_n}, \mathcal{C}_1)\mathbb{P}(\mathcal{C}_1) \\ &= \int_0^{T-t'} \nu e^{-(\nu+\lambda_{\mathbb{P}})s} \int_{-1}^\infty 1_B(t' + s, \Gamma_s^{f_n^{\Xi_n}}(x, 0)(1 + \alpha_s^S y), 0)\gamma(\mathrm{d}y)\mathrm{d}s \ ,\end{aligned} \tag{37}$$

and

$$\begin{aligned}&\mathbb{P}(\Xi_{n+1} \in B|\Xi_n = (t', x, 0), f_n^{\Xi_n}, \mathcal{C}_2)\mathbb{P}(\mathcal{C}_2) \\ &= \int_0^{T-t'} \lambda_{\mathbb{P}} e^{-(\nu+\lambda_{\mathbb{P}})s} 1_B(t' + s, \Gamma_s^{f_n^{\Xi_n}}(x, 0)(1 - \alpha_s^P L), 0)\mathrm{d}s \ .\end{aligned} \tag{38}$$

Finally, plugging equations (36), (37) and (38) in expression (35) completes the first part of the proof. The additional result

$$Q(\{\Delta\}|\Xi_n, f_n^{\Xi_n}) = 1 - Q(E|\Xi_n, f_n^{\Xi_n})$$

is trivial.

References

[1] Almudevar, A. (2001). A dynamic programming algorithm for the optimal control of piecewise deterministic Markov processes. *SIAM J. Control Optim.* **40,** 525–539.

[2] Bäuerle, N. and Rieder, U. (2009). MDP algorithms for portfolio optimization problems in pure jump markets. *Finance and Stochastics* **13,** 591–611.

[3] Bäuerle, N. and Rieder, U. (2010). Optimal control of piecewise deterministic Markov processes with finite time horizon. *Mod. Trends of Contr. Stoch. Proc.: Theory and Applications* 144–160.

[4] Bäuerle, N. and Rieder, U. (2011). *Markov Decision Processes with Applications to Finance.* Springer.

[5] Bertsekas, D. P. and Shreve, S. (1978). *Stochastic Optimal Control.* Academic Press.

[6] Bielecki, T. R. and Jang, I. (2006). Portfolio optimization with a defaultable security. *Asia Pacific Finan. Markets* **13,** 113–127.

[7] Bielecki, T. R. and Rutkowski, M. (2002). *Credit Risk: Modelling, Valuation and Hedging.* Springer.

[8] Bo, L., Wang, Y. and Yang, X. (2010). An optimal portfolio problem in a defaultable market. *Adv. in Applied Probability* **42,** 689–705.

[9] Capponi, A. and Figueroa-Lopez, J. E. (2011). Dynamic portfolio optimization with a defaultable security and regime switching. *http://arxiv.org/abs/1105.0042*.

[10] Duffie, D. and Singleton, K. J. (1999). Modelling term structures of defaultable bonds. *The Review of Financial Studies* **12,** 687–720.

[11] Fleming, W. H. and Pang, T. (2004). An application of stochastic control theory to financial economics. *SIAM J. Control Optim.* **43,** 502–531.

[12] Jeanblanc, M., Yor, M. and Chesney, M. (2009). *Mathematical Methods for Financial Markets.* Springer.

[13] Karatzas, I. and Shreve, S. (1998). *Methods of Mathematical Finance.* Springer.

[14] Kirch, M. and Runggaldier, W. J. (2004). Efficient hedging when asset prices follow a geometric Poisson process with unknown intensities. *SIAM J. Control Optim.* **43,** 1174–1195.

[15] Lakner, P. and Liang, W. (2008). Optimal investment in a defaultable bond. *Mathematics and Financial Economics* **3,** 283–310.

[16] Merton, R. C. (1969). Lifetime portfolio selection under uncertainty: the continuous case. *Rev. Econom. Statist.* **51,** 247–257.

[17] Pham, H. (2002). Smooth solutions to optimal investment methods with stochastic volatilities and portfolio constraints. *Appl. Math. Optimization* **46,** 55–78.

[18] Puterman, M. L. (2005). *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* Wiley.