
Dynamic Credit Investment in Partially Observed Markets

Agostino Capponi · José E. Figueroa-López · Andrea Pascucci

Abstract We consider the problem of maximizing expected utility for a power investor who can allocate his wealth in a stock, a defaultable security, and a money market account. The dynamics of these security prices are governed by geometric Brownian motions modulated by a hidden continuous time finite state Markov chain. We reduce the partially observed stochastic control problem to a complete observation risk sensitive control problem via the filtered regime switching probabilities. We separate the latter into pre-default and post-default dynamic optimization subproblems, and obtain two coupled Hamilton-Jacobi-Bellman (HJB) partial differential equations. We prove existence and uniqueness of a globally bounded classical solution to each HJB equation, and give the corresponding verification theorem. We provide a numerical analysis showing that the investor increases his holdings in stock as the filter probability of being in high growth regimes increases, and decreases his credit risk exposure when the filter probability of being in high default risk regimes gets larger.

Keywords and phrases: Partial information, Filtering, Risk sensitive control, Default risk, Hidden markov chain.

JEL codes: G11, C61, C11.

MSC codes: Primary 93E20; Secondary 91B28, 49L20, 49L25.

1 Introduction

Regime switching models constitute an appealing framework, stemming from their ability to capture the relevant features of asset price dynamics, which behave differently depending on the specific phase of the business cycle in place. In the context of continuous time utility maximization, some studies have considered observable regimes, while others have accounted for the possibility that they are not directly visible. In the case of observable regimes, Zariphopoulou (1992) considers an infinite horizon investment-consumption model where the agent can invest her wealth in a stock and risk-free bond, with borrowing and stock short-selling constraints. In a similar regime switching framework, Sotomayor and Cadenillas (2009) study the infinite horizon problem of a risk averse investor maximizing regime dependent utility from terminal wealth and consumption. A different branch of literature has considered the case when regimes are hidden and need to be estimated from publicly available market information. Nagai and Runggaldier (2008) consider a finite horizon portfolio optimization problem, where a power investor allocates his wealth across money market account and stocks, whose price dynamics follow a diffusion process modulated by a hidden finite-state Markov process. Tamura and Watanabe (2011) extend the analysis to the case when the

Department of Applied Mathematics and Statistics, Johns Hopkins University, Baltimore, MD, 21218, USA (acappon1@jhu.edu). Proofs should be sent to agcappo@gmail.com.

Department of Statistics, Purdue University, West Lafayette, IN, 47907, USA (figueroa@purdue.edu). Research supported in part by the NSF Grant DMS-1149692.

Department of Mathematics, University of Bologna, Bologna, 40126, Italy (andrea.pascucci@unibo.it).

Address(es) of author(s) should be given

time horizon is infinite. Elliott and Siu (2011) study the optimal investment problem of an insurer when the model uncertainty is governed by a hidden Markov chain. Siu (2013) considers optimal investment problems in general non-Markovian hidden regime switching models. Sass and Haussmann (2004) consider a multi-stock market model, with stochastic interest rates and drift modulated by a hidden Markov chain. Combining appropriate Malliavin calculus and filtering results from hidden Markov models, they derive explicit representations of the optimal strategies. In a series of two papers, Fujimoto et al. (2013a) and Fujimoto et al. (2013b) consider a regime switching framework where logarithmic and power investors optimize their terminal utility by investing in stocks at random times due to liquidity constraints.

The literature surveyed above has considered markets consisting of securities carrying market, but not default, risk. In recent years, few studies have considered a portfolio optimization framework inclusive of defaultable securities. Kraft and Steffensen (2005) study optimal portfolio problems with defaultable assets within a Black-Cox framework. Kraft and Steffensen (2008) consider an investor who can allocate her wealth across multiple defaultable bonds, in a model where simultaneous defaults are allowed. In the same market model, Kraft and Steffensen (2009) define default as the beginning of financial distress, and discuss contagion effects on prices of defaultable bonds. Bielecki and Jang (2006) derive optimal investment strategies for a CRRA investor, allocating her wealth among a defaultable bond, risk-free bank account, and a stock. Bo et al. (2010) consider a portfolio optimization problem, where a logarithmic investor can choose a consumption rate, and invest her wealth across a defaultable perpetual bond, a stock, and a money market account. Jiao and Pham (2013) combine duality theory and dynamic programming to optimize the utility of a CRRA investor in a market consisting of a riskless bond and a stock subject to counterparty risk. Optimal investment under contagion risk has been considered by Bo and Capponi (2014), who construct an empirically motivated framework based on interacting intensity models, and analyze how contagion effects impact optimal allocation decisions due to abrupt changes in prices. A related study by Jiao et al. (2013) develop a portfolio framework where multiple default events can occur, and some of the securities may still be traded after default.

The first attempt at using regime switching within a portfolio optimization framework consisting of defaultable securities was done by Capponi and Figueroa-López (2014). Such a modeling choice is also empirically supported by a study of Giesecke et al. (2011), which identifies three credit regimes characterized by different levels of default intensity and recovery rates, via a historical analysis of the corporate bond market. Capponi and Figueroa-López (2014) consider an investor trading in a stock and defaultable security, whose price dynamics are modulated by an observable Markov chain. Using the HJB approach, they recover the optimal investment strategies as the unique solution to a coupled system of partial differential equations.

The present paper considers the case where regimes are hidden, so that the power investor must decide on the optimal allocation policy using only the observed market prices. This improves upon the realism of the model in Capponi and Figueroa-López (2014), given that in several circumstances market regimes such as inflation and recession, or credit regimes characterized by high or low credit spreads, are typically unobserved to investors. Moreover, the hidden regime feature requires a completely different analysis, and leads us to solving a partially observed stochastic control problem, where regime information must be inferred from an enlarged market filtration. The latter is composed *both* of a reference filtration generated by the observable security prices, *and* of a credit filtration tracking the occurrence of the default event. To the best of our knowledge, ours represents the first study in this direction.

We next describe our main contributions. First, by considering a portfolio optimization problem in a context of partial information with possibility of default, we advance earlier literature which has so far considered either one or the other aspect, but never both simultaneously. We construct an equivalent fully observed risk-sensitive control problem, where the new state is given by the regime filtered probabilities, generalizing the approach of Nagai and Runggaldier (2008) who do not deal with default event information. We use the filter probabilities to obtain the Hamilton-Jacobi-Bellman (HJB) equation for the dynamic optimization problem, which we separate it into coupled pre-default and post-default optimization subproblems. This is done using the projected filter process. We remark that the decomposition of a global optimal investment problem into sub-control problems in a progressively enlarged filtration has also been considered by Jiao et al. (2013) and Pham (2010). Their approach consists in first defining the sub-control problems in the reference market filtration exclusive of default event information, and then connecting them by assuming the existence of a conditional density on the default times, previously introduced in

El Karoui et al. (2010). Despite few similarities between ours and their approach arising from the fact that both consider a pre and post-default decomposition and solve backwards, there are also significant differences between the two approaches. We consider the wealth dynamics under the enlarged market filtration inclusive of default events and do not perform any pre-post default decomposition of the control problem at the level of the stochastic differential equation. It is only after deriving the HJB partial differential equations that the decomposition into pre and post-default PDEs naturally arises. Their approach instead exploits the exponential utility preference function of the investor to reduce the optimal investment problem to solving a recursive system of backward stochastic differential equations with respect to the default-free market filtration. A detailed analysis of these BSDEs including the possibility of jump times driven by Brownian motion is provided in Kharroubi and Lim (2013).

Secondly, the presence of default risk makes the HJB-PDE satisfied by the pre-default value function non-linear. There are two sources of nonlinearity, namely quadratic growth of the gradient and exponential nonlinearity. We first perform a suitable transformation yielding a parabolic PDE whose associated operator is linear in the gradient and matrix of second derivatives, but nonlinear in the solution. We then provide a rigorous analysis of the transformed PDE and prove the existence of a classical solution via a monotone iterative method. Since the nonlinear term is only locally, but not globally, Lipschitz continuous because the derivative explodes at zero, we also need to prove that the solution is bounded away from zero. In particular, we establish both a lower and upper bound for the solution, and prove $\mathcal{C}_P^{2,\alpha}$ regularity. We then use this result to prove a verification theorem establishing the correspondence between the solution to the PDE and the value function of the control problem. The proof of the theorem requires the development of a number of technical results, such as the guaranteed positivity of the filtering process. By contrast, the HJB-PDE corresponding to the post-default optimization problem can be linearized using a similar transformation to the one adopted by Nagai and Runggaldier (2008), and a unique classical solution can be guaranteed as shown, for instance, in Tamura and Watanabe (2011).

Thirdly, we provide a thorough comparative statics analysis to illustrate the impact of partial information on the optimal allocation decisions. We consider a square root investor and a two-states Markov chain. We find that the fraction of wealth invested in the stock increases as the filter probability of being in the regime with the highest growth rate increases. In order to be hedged against default, the investor shorts a higher number of units of defaultable security if the filter probability of staying in the highest default risk regime increases. Vice versa, when the probability of being in the safest regime increases, the investor increases his exposure to credit risk by shorting smaller amount of units of the defaultable security. If the regime is characterized by a sufficiently low level of default intensity, the square root investor may even go long credit, and purchase units of the defaultable security. We find that lower values of price volatility induce the investor to increase the fraction of wealth invested in the risky asset. More specifically, if the stock volatility is low, the filter gain coming from received observations is higher and the investor purchases increasingly more units as the stock volatility decreases. Similarly, for a sufficiently high probability of being in the high default risk regime, the investor shorts increasingly larger number of units of defaultable security as the volatility of the latter decreases. This reflects the risk averse nature of the investor, who wants to reduce his credit risk exposure more if the filter estimate becomes more accurate due to the higher informational gain from price observations. We also find that as observations become less informative due to higher price volatilities, the investor deposits a significant fraction of his wealth in the money market account. All this suggests that partial information tends to push the investor towards strategies reducing both market and credit risk exposure.

The rest of the paper is organized as follows. Section 2 defines the market model. Section 3 sets up the utility maximization problem. Section 4 derives the HJB equations corresponding to the risk sensitive control problem. Section 5 analyzes the solutions of the HJB-PDE equations. Section 6 develops a numerical analysis. Section 7 summarizes our main conclusions. Finally, two appendices present the main proofs of the paper.

2 The Market Model

Assume $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ is a complete filtered probability space, where \mathbb{P} is the historical probability measure, $\mathbb{G} := (\mathcal{G}_t)_{t \geq 0}$ is an enlarged filtration given by $\mathcal{G}_t := \mathcal{F}_t \vee \mathcal{H}_t$, where $(\mathcal{H}_t)_t$ is a filtration to be introduced below. We take the right continuous version of \mathbb{G} , i.e. \mathcal{G}_t is the smallest right-continuous filtration containing \mathbb{F} and \mathbb{H} , with $\mathcal{G}_t := \bigcap_{\epsilon > 0} (\mathcal{F}_{t+\epsilon} \vee \mathcal{H}_{t+\epsilon})$ (see also Belanger et al. (2004)).

Here, $\mathbb{F} := (\mathcal{F}_t)_t$ is a suitable filtration supporting a two dimensional Brownian motion $W_t = (W_t^{(1)}, W_t^{(2)})^\top$, where $^\top$ denotes the transpose. We also assume that the *hidden* states of the economy are modeled by a finite-state continuous-time Markov chain $X := \{X_t\}_{t \geq 0}$, which is adapted to \mathbb{G} and assumed to be independent of $W := \{W_t\}_{t \geq 0}$. Without loss of generality, the state space is identified by the set of unit vectors $\{e_1, e_2, \dots, e_N\}$, where $e_i = (0, \dots, 1, \dots, 0)^\top \in \mathbb{R}^N$. The following semi-martingale representation is well-known (cf. Elliott et al. (1994)):

$$X_t = X_0 + \int_0^t A(s)^\top X_s ds + \varphi(t), \quad (2.1)$$

where $\varphi(t) = (\varphi_1(t), \dots, \varphi_N(t))^\top$ is a \mathbb{R}^N -valued martingale under \mathbb{P} , and $A(t) := [\varpi_{i,j}(t)]_{i,j=1,\dots,N}$ is the so-called generator of the Markov process. Specifically, denoting $p_{i,j}(t, s) := \mathbb{P}(X_s = e_j | X_t = e_i)$, for $s \geq t$, and $\delta_{i,j} = \mathbf{1}_{i=j}$, we have that

$$\varpi_{i,j}(t) = \lim_{h \rightarrow 0} \frac{p_{i,j}(t, t+h) - \delta_{i,j}}{h};$$

cf. Bielecki and Rutkowski (2001). In particular, $\varpi_{i,i}(t) = -\sum_{j \neq i} \varpi_{i,j}(t)$. The following mild condition is also imposed:

$$\sup_{t \geq 0} \max_{i,j} \varpi_{i,j}(t) < \infty. \quad (2.2)$$

We denote by $p^\circ = (p^{\circ,1}, \dots, p^{\circ,N})$ the initial distribution on the Markov chain and, throughout the paper, assume that $p^{\circ,i} > 0$.

We consider a frictionless financial market consisting of three instruments: a risk-free bank account, a defaultable security, and a stock.

Risk-free bank account. The instantaneous market interest rate is assumed to be constant. The dynamics of the price process $\{B_t\}$, which describes the risk-free bank account, is given by

$$dB_t = rB_t dt, \quad B_0 = 1. \quad (2.3)$$

Stock Security. We assume that the appreciation rate $\{\mu_t\}$ of the stock depends on the economic regime X_t in the following way:

$$\mu_t := \langle \mu, X_t \rangle,$$

where $\mu = (\mu_1, \mu_2, \dots, \mu_N)$ is a vector with constant components denoting the values of the drift associated to the different economic regimes and where $\langle \cdot, \cdot \rangle$ denotes the standard inner product in \mathbb{R}^N . Under the historical measure, the stock dynamics is given by

$$dS_t = \mu_t S_t dt + \sigma S_t dW_t^{(1)}, \quad S_0 = s^\circ. \quad (2.4)$$

Defaultable Security. Before defining the vulnerable security considered in the present paper, we need to introduce the default process. Let τ be a nonnegative random variable, defined on $(\Omega, \mathcal{G}, \mathbb{P})$, representing the default time of the counterparty selling the security. Let $\mathcal{H}_t = \sigma(H_u : u \leq t)$ be the filtration generated by the *default process* $H_t := \mathbf{1}_{\tau \leq t}$. We use the canonical construction of the default time τ in terms of a given hazard process. The latter is defined by $\Gamma_t := -\ln(1 - \mathbb{E}^\mathbb{P}[H_t | \mathcal{F}_t])$ (see also definition 9.2.1 in Bielecki and Rutkowski (2001)), and postulated to have absolutely continuous sample paths with respect to the Lebesgue measure on \mathbb{R}^+ . In other words, it is assumed to admit the integral representation

$$\Gamma_t = \int_0^t h_u du$$

for a \mathbb{F} progressively measurable, nonnegative stochastic process $\{h_t\}_{t \geq 0}$, with integrable sample paths. The process h_t is referred to as the \mathbb{F} -hazard rate of τ , and will be specified later. We next give the details of the construction of the random time τ . We assume the existence of an exponential random variable χ defined on the probability space $(\Omega, \mathcal{G}, \mathbb{P})$, independent of the process $(X_t)_t$. We define τ by setting

$$\tau := \inf \left\{ t \in \mathbb{R}^+ : \int_0^t h_u du \geq \chi \right\}, \quad (2.5)$$

where we follow the convention $\inf \emptyset = +\infty$. It can be proven (see Bielecki and Rutkowski (2001), Section 6.5 for details) that

$$\xi_t := H_t - \int_0^t \bar{H}_u h_u du = H_t - \int_0^{t \wedge \tau} h_u du, \quad (2.6)$$

is a \mathbb{G} -martingale under \mathbb{P} , where $\bar{H}_u := 1 - H_u$ and $\bar{H}_u^- := \lim_{s \uparrow u} \bar{H}_s = \mathbf{1}_{\tau \geq u}$. Intuitively, Eq. (2.6) says that the single jump process needs to be compensated for default, prior to the occurrence of the event. As with the appreciation rate, we assume that the process h is driven by the hidden Markov chain as follows:

$$h_t := \langle h, X_t \rangle,$$

where $h = (h_1, \dots, h_N) \in (0, \infty)^N$ denotes the possible values that the default rate process can take depending on the economic regime in place. We model the pre-default dynamics of the defaultable security as

$$\frac{dP_t}{P_t} = a(t, X_t)dt + v dW_t^{(2)}, \quad (t < \tau), \quad P_0 = P^\circ, \quad (2.7)$$

where $P^\circ \in \mathbb{R}_+$ and $a : \mathbb{R}_+ \times \{e_1, \dots, e_N\} \rightarrow \mathbb{R}_+$ is a deterministic function. After default the security becomes worthless, i.e. $P_t := 0$ for any $t \geq \tau$ and, thus, $\{P_t\}_{t \geq 0}$ follows the dynamics:

$$dP_t = P_{t-} \left(a(t, X_t)dt + v dW_t^{(2)} - dH_t \right). \quad (2.8)$$

For future reference, we also impose the following mild technical assumption:

$$\int_0^T a^2(t, e_i)dt < \infty, \quad \text{for any } T > 0 \text{ and } i \in \{1, \dots, N\}. \quad (2.9)$$

Remark 2.1 As usual when dealing with hidden Markov models the volatility components are assumed to be constant, see for instance Nagai and Runggaldier (2008). If σ and v were not constant but consisting of distinct components depending on X , then the Markov chain $\{X_t\}_{t \geq 0}$ would become observable. This is because the quadratic variation of X_t would converge almost surely to the integrated volatility, see McKean (1965). Consequently, by inversion, the regime in place at time t would become known. Further, we notice that the choice of constant volatility might also provide a fairly good fit to market data when calibrating the hidden regime switching model to market prices. This has been empirically shown by Liechty and Roberts (2001) on data from the New York Merchantile stock exchange using Markov chain Monte-Carlo methods.

Remark 2.2 The specification given in (2.7) captures several relevant market models which have been considered in the literature:

1. First, the model (2.7) may be specialized to capture the pre-default dynamics of a defaultable stock. The latter is a widely used instrument in hybrid models of equity and credit. For instance, Linetsky (2006) and Carr *et al.* (2010) model the pre-bankruptcy risk-neutral dynamics of a defaultable stock as

$$dS_t = (r + h_t)S_t dt + v S_t dW_t,$$

where $\{W_t\}_{t \geq 0}$ is a Brownian driver and $\{h_t\}_{t \geq 0}$ is a stochastic (adapted) default intensity process. Clearly, such a specification is a special instance of (2.7), where we set $a(t, X_t) = r + h(X_t)$. The addition of the hazard rate in the drift ensures that the discounted stock price process is a martingale.

2. Secondly, the dynamics in Eq. (2.7) may be used to model the time evolution of prices of credit sensitive securities when an additive type of “microstructure or market-friction” noise is taken into account. In general, secondary market investors only observe market quotes for traded credit derivatives, such as spreads, at discrete points in time, e.g., at times $t_k = k\Delta$, $k = 0, \dots, N$, for a certain fixed time mesh $\Delta > 0$. The corresponding observed yield spreads z_{t_k} are then often modeled as $z_{t_k} = a_{t_k}(t_k, X_{t_k}) + \epsilon_k$, with an i.i.d. sequence $(\epsilon_k)_k$, independent of X , capturing microstructure noise due to illiquidity, transaction costs, or transmission “errors”. In that case, $a(t_k, X_{t_k})$ represents the underlying unobserved yield spread which follows an efficient arbitrage-free model of choice. Frey and Runggaldier (2011) argue that as the interarrival time $t_k - t_{k-1}$ between consecutive observations gets smaller, the cumulative log return process $z_t := \log(P_t/P_0)$ of the defaultable security converges, in law, to

$$\int_0^t a(s, X_s)ds + v W_t^{(2)}. \quad (2.10)$$

Again, the dynamics of (2.10) is in the form of our dynamics (2.7).

For future convenience, we introduce the two-dimensional observed pre-default log-price process $Y_t = (\log(S_t), \log(P_t))^\top$, whose dynamics is given by

$$dY_t = \vartheta(t, X_t)dt + \Sigma_Y dW_t, \quad (2.11)$$

where

$$\Sigma_Y := \begin{pmatrix} \sigma & 0 \\ 0 & \nu \end{pmatrix}, \quad \vartheta(t, X_t) := \left[\mu_t - \frac{\sigma^2}{2}, a(t, X_t) - \frac{\nu^2}{2} \right]^\top = \left[\langle \mu, X_t \rangle - \frac{\sigma^2}{2}, a(t, X_t) - \frac{\nu^2}{2} \right]^\top. \quad (2.12)$$

We also define two subfiltrations of \mathbb{G} , namely, the market filtration $\mathbb{G}^I := (\mathcal{G}_t^I)_{t \geq 0}$ where

$$\mathcal{G}_t^I := \mathcal{F}_t^I \vee \mathcal{H}_t, \quad \mathcal{F}_t^I := \sigma(S_u, P_u; u \leq t),$$

and the subfiltration $\mathbb{F}^X := (\mathcal{F}_t^X)_{t \geq 0}$, generated by the Markov chain $(X_t)_t$:

$$\mathcal{F}_t^X = \sigma(X_u; u \leq t).$$

Therefore, we may also write $\mathcal{G}_t = \mathcal{F}_t^X \vee \mathcal{G}_t^I$. From this, it is evident that while $(X_t)_{t \geq 0}$ is $(\mathcal{G}_t)_{t \geq 0}$ adapted, it is not $(\mathcal{G}_t^I)_{t \geq 0}$ adapted.

3 The Utility Maximization Problem

We consider an investor who wants to maximize her expected final utility during a trading period $[0, T]$, by dynamically allocating her financial wealth into (1) the risk-free bank account, (2) the stock, and (3) the defaultable security, as defined in the previous section. Let us denote by ν_t^B the number of shares of the risk-free bank account that the investor buys ($\nu_t^B > 0$) or sells ($\nu_t^B < 0$) at time t . Similarly, ν_t^S and ν_t^P denote the investor's portfolio positions in the stock and defaultable security at time t , respectively. The process $\nu := \{(\nu_t^B, \nu_t^S, \nu_t^P)\}_{t \geq 0}$ is called a *portfolio process*. We denote $V_t(\nu)$ the wealth of the portfolio process $\nu = (\nu^B, \nu^S, \nu^P)$ at time t , i.e.

$$V_t(\nu) = \nu_t^B B_t + \nu_t^S S_t + \nu_t^P \mathbf{1}_{\tau > t} P_t.$$

We require the processes ν_t^B, ν_t^S , and ν_t^P to be \mathcal{G}^I -predictable. The investor does not have intermediate consumption nor capital income to support her trading of financial assets and, hence, we also assume the following self-financing condition:

$$dV_t = \nu_t^B dB_t + \nu_t^S dS_t + \nu_t^P \mathbf{1}_{\tau > t} dP_t.$$

Let

$$\pi_t^B := \frac{\nu_t^B B_t}{V_{t-}(\nu)}, \quad \pi_t^S := \frac{\nu_t^S S_t}{V_{t-}(\nu)}, \quad \pi_t^P = \frac{\nu_t^P P_t}{V_{t-}(\nu)} \mathbf{1}_{\tau > t}, \quad (3.1)$$

if $V_{t-}(\nu) > 0$, while $\pi_t^B = \pi_t^P = \pi_t^S = 0$ when $V_{t-}(\nu) = 0$. The vector $\bar{\pi}_t := (\pi_t^B, \pi_t^S, \pi_t^P)$, called a *trading strategy*, represents the corresponding fractions of wealth invested in each asset at time t . Note that if $\bar{\pi} := (\bar{\pi}_t)_t$ is admissible (the precise definition will be given later), then the dynamics of the resulting wealth process in terms of $\bar{\pi}$ can be written as

$$dV_t = V_{t-} \left\{ \pi_t^B \frac{dB_t}{B_t} + \pi_t^S \frac{dS_t}{S_t} + \pi_t^P \frac{dP_t}{P_t} \right\}, \quad (3.2)$$

under the convention that $0/0 = 0$. The latter convention is needed to deal with the case when default has occurred ($t \geq \tau$), so that $P_t = 0$ and we have $\pi_t^P = 0$. Using that $\pi_t^B + \pi_t^P + \pi_t^S = 1$, and the corresponding dynamics of B_t, S_t and P_t , we can further rewrite the dynamics (3.2) as

$$\frac{dV_t^\pi}{V_{t-}^\pi} = rdt + \pi_t^S (\mu_t - r) dt + \pi_t^S \sigma dW_t^{(1)} + \pi_t^P (a(t, X_t) - r) dt + \pi_t^P \nu dW_t^{(2)}, \quad V_0^\pi = v, \quad (3.3)$$

for a given initial budget $v \in (0, \infty)$. Above, we use $\pi_t := (\pi_t^S, \pi_t^P)^\top$ to denote the time- t investment strategy only consisting of positions on the stock and defaultable security, and write V^π to emphasize the dependence of

the wealth process on the strategy π . The objective of the power investor is then to choose $\pi = (\pi^S, \pi^P)^\top$ so to maximize the expected utility from terminal wealth

$$J(v, \pi, T) := \frac{1}{\gamma} \mathbb{E}^\mathbb{P} [(V_T^\pi)^\gamma], \quad (3.4)$$

for a given fixed value of $\gamma \in (0, 1)$. By Itô's formula and Eq. (3.3), we readily obtain that $V_t^\gamma := (V_t^\pi)^\gamma$ follows the dynamics

$$\begin{aligned} dV_t^\gamma &= \gamma V_t^\gamma \left[r dt + \pi_t^S (\mu_t - r) dt + \pi_t^S \sigma dW_t^{(1)} + \pi_t^P (a(t, X_t) - r) dt + \pi_t^P v dW_t^{(2)} \right] \\ &\quad + \frac{1}{2} \gamma (\gamma - 1) V_t^\gamma [(\pi_t^S)^2 \sigma^2 dt + (\pi_t^P)^2 v^2 dt]. \end{aligned}$$

Next, recalling that $W_t := (W_t^{(1)}, W_t^{(2)})^\top$, $\pi_t = (\pi_t^S, \pi_t^P)^\top$, and the definition of Σ_Y given in (2.12), we may rewrite the above SDE as

$$dV_t^\gamma = V_t^\gamma [-\gamma \eta(t, X_t, \pi_t) dt + \gamma \pi_t^\top \Sigma_Y dW_t], \quad (3.5)$$

where

$$\eta(t, X_t, \pi_t) = -r + \pi_t^S (r - \langle \mu, X_t \rangle) + \pi_t^P (r - a(t, X_t)) + \frac{1 - \gamma}{2} \pi_t^\top \Sigma_Y^\top \Sigma_Y \pi_t. \quad (3.6)$$

It is then clear that the solution to the stochastic differential equation (3.5) with initial condition $V_0 = v$ is given by

$$V_t^\gamma = v^\gamma \exp \left(\gamma \int_0^t \pi_s^\top \Sigma_Y dW_s - \gamma \int_0^t \eta(s, X_s, \pi_s) ds - \frac{\gamma^2}{2} \int_0^t \pi_s^\top \Sigma_Y \Sigma_Y^\top \pi_s ds \right). \quad (3.7)$$

From Eq. (3.4) and (3.7), we can see that we need to solve a maximization problem with partial information since the regime X_t is not directly observable and investment strategies can only be based on past information of security prices. Our approach is to transform it into a fully observed risk sensitive control problem. Such a reduction is accomplished through two main steps. First, in Section 3.1 we show equivalence to a complete observation control problem with finite dimensional Markovian state. Then, in Section 3.2 we transform the complete observation control problem into a risk-sensitive stochastic control problem.

3.1 An Equivalent Formulation as a Complete Observation Control Problem

The objective of this section is to show how the partially observed control problem in (3.4) may be reduced to a complete observation control problem. This is accomplished by showing that the criterion (3.4) may be equivalently rewritten as an expectation, taken with respect to a suitably chosen probability measure, of an exponential function of the (observable) regime filtered probabilities. Next, we start developing the change to the new measure, chosen so that the underlying chain $(X_t)_t$ becomes independent of the investor filtration \mathbb{G}^I under such a measure. First, we introduce some needed notation and terminology. Given two semimartingales L and M , we denote by $[L]$ and $[L, M]$ the quadratic variation of L and the quadratic covariation of L and M , respectively. We also denote the stochastic exponential of L by $\mathcal{E}(L)$. If L is of the form $L_t = \int_0^t \theta_s^\top dY_s$, where Y_s is a \mathbb{R}^d -valued continuous Itô process, and $\{\theta_s\}_{s \geq 0}$ is \mathbb{G} predictable, then

$$\mathcal{E}_t(L) = \exp \left(\int_0^t \theta_u^\top dY_u - \frac{1}{2} \int_0^t \theta_u^\top \theta_u d[Y]_u \right). \quad (3.8)$$

If Z is of the form $Z_t = \int_0^t \iota_s d\xi_s$, where ξ_s has been defined in (2.6), and $\{\iota_s\}_{s \geq 0}$ is \mathbb{G} -predictable, with $\iota > -1$, then

$$\mathcal{E}_t(Z) = \exp \left(\int_0^t \log(1 + \iota_s) dH_s - \int_0^{t \wedge \tau} \iota_s h_s ds \right). \quad (3.9)$$

It is well known (see Bielecki and Rutkowski (2001), Section 4.3) that $R_t := \mathcal{E}_t(L) \mathcal{E}_t(Z)$ follows the SDE

$$R_t = 1 + \int_{(0, t]} R_{s-} (\theta_s^\top dY_s + \iota_s d\xi_s). \quad (3.10)$$

We now proceed to introduce the new measure $\hat{\mathbb{P}}$ on (Ω, \mathbb{G}) . Such a measure is defined in terms of its density process as follows:

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} \Big|_{\mathcal{G}_t} := \mathcal{E}_t \left(\int_0^\cdot -\vartheta(s, X_s)^\top \Sigma_Y^{-1} dW_s \right) \mathcal{E}_t \left(\int_0^\cdot \frac{1 - h_{s^-}}{h_{s^-}} d\xi_s \right) =: \rho_t^{(1)} \rho_t^{(2)}. \quad (3.11)$$

In particular, using Eqs. (3.8) and (3.9), $\rho_t^{(1)}$ and $\rho_t^{(2)}$ above are given by

$$\begin{aligned} \rho_t^{(1)} &= \exp \left(- \int_0^t \vartheta(s, X_s)^\top \Sigma_Y^{-1} dW_s - \frac{1}{2} \int_0^t \vartheta^\top (\Sigma_Y \Sigma_Y^\top)^{-1} \vartheta(s, X_s) ds \right), \\ \rho_t^{(2)} &= \exp \left(- \int_0^t \log(h_{u^-}) dH_u - \int_0^{t \wedge \tau} (1 - h_u) du \right) = h_{\tau^-}^{-1 \{ \tau \leq t \}} \exp \left(- \int_0^{t \wedge \tau} (1 - h_u) du \right). \end{aligned}$$

Moreover, from Eq. (3.10), the density process $\rho_t = \rho_t^{(1)} \rho_t^{(2)}$ admits the following representation

$$\rho_t = 1 + \int_0^t \rho_{s^-} \left(-\vartheta(s, X_s)^\top \Sigma_Y^{-1} dW_s + \frac{1 - h_{s^-}}{h_{s^-}} d\xi_s \right). \quad (3.12)$$

In order to show that $\hat{\mathbb{P}}$ is well-defined, one must verify that $\mathbb{E}^{\mathbb{P}}(\rho_T) = 1$. To this end, we use a general version of Novikov's condition, as proved in Protter and Shimbo (2008) (see Theorem 9 therein), which states that the stochastic exponential $\mathcal{E}(M)$ of a locally square integrable martingale M is a martingale on $[0, T]$ if

$$\mathbb{E}^{\mathbb{P}} \left[e^{\frac{1}{2} \langle M^c, M^c \rangle_T + \langle M^d, M^d \rangle_T} \right] < \infty, \quad (3.13)$$

where M^c and M^d are the continuous and purely discontinuous martingale parts of M . Here, $\langle M^c, M^c \rangle_T$ and $\langle M^d, M^d \rangle_T$ denote the compensators of the quadratic variations of M^c and M^d at time T , respectively (see Protter (2004), Page 70). From (3.12), $\rho_t = \mathcal{E}_t(M)$ with

$$M_t = - \int_0^t \vartheta(s, X_s)^\top \Sigma_Y^{-1} dW_s + \int_0^t \frac{1 - h_{s^-}}{h_{s^-}} d\xi_s.$$

Therefore, we have

$$\begin{aligned} \langle M^c, M^c \rangle_T &= \int_0^T \vartheta^\top(s, X_s) (\Sigma_Y \Sigma_Y^\top)^{-1} \vartheta(s, X_s) ds, \\ \langle M^d, M^d \rangle_T &= \int_0^T \left[\frac{(1 - h_s)^2}{h_s} \right] \bar{H}_s ds. \end{aligned}$$

Clearly, $\langle M^c, M^c \rangle_T$ is bounded in view of the condition (2.9). It remains to prove that $\langle M^d, M^d \rangle_T$ is also bounded. We have

$$\langle M^d, M^d \rangle_T = \int_0^T \left(\frac{1}{h_s} + h_s - 2 \right) \bar{H}_s ds \leq \int_0^T \left(\frac{1}{h_s} + h_s \right) ds.$$

Since $h_i > 0$ for all $i \in \{1, \dots, N\}$ and $h_s = \langle h, X_s \rangle$, we obtain that

$$\max_{i \in \{1, \dots, N\}} \left(\frac{1}{h_i} + h_i \right) < C,$$

for some constant $C > 0$. Thus, we conclude that $\langle M^d, M^d \rangle_T$ is also bounded.

Under the probability measure $\hat{\mathbb{P}}$, by Girsanov's theorem (see, e.g., Bielecki and Rutkowski (2001), Section 5.3), we have that

$$\hat{W}_t = W_t + \int_0^t \Sigma_Y^{-1} \vartheta(s, X_s) ds$$

is a Brownian motion, and

$$\hat{\xi}_t = \xi_t - \int_0^{t \wedge \tau} (1 - h_u) du = H_t - \int_0^{t \wedge \tau} du = H_t - \int_0^t \bar{H}_u^- du \quad (3.14)$$

is a \mathbb{G}^I -martingale. Note also that, from Eq. (2.11), the observed pre-default log-price process $Y_t = (\log(S_t), \log(P_t))^\top$ possesses the dynamics $dY_t = \Sigma_Y d\hat{W}_t$ under $\hat{\mathbb{P}}$. Furthermore, the inverse density process,

$$U_t := \frac{d\mathbb{P}}{d\hat{\mathbb{P}}}\bigg|_{\mathcal{G}_t},$$

can be written as $U_t = U_t^{(1)}U_t^{(2)}$, where

$$\begin{aligned} U_t^{(1)} &:= \exp\left(\int_0^t \vartheta(s, X_s)^\top \Sigma_Y^{-1} dW_s + \frac{1}{2} \int_0^t \vartheta^\top (\Sigma_Y \Sigma_Y^\top)^{-1} \vartheta(s, X_s) ds\right) \\ U_t^{(2)} &:= h_{\tau^-}^{\mathbf{1}_{\{\tau \leq t\}}} \exp\left(\int_0^{t \wedge \tau} (1 - h_u) du\right) = \mathcal{E}_t\left(\int_0^\cdot (h_{s^-} - 1) d\hat{\xi}_s\right). \end{aligned}$$

Using the previous probability measure $\hat{\mathbb{P}}$ together with the representation (3.7), Eq. (3.4) may be rewritten as

$$\begin{aligned} \frac{1}{\gamma} \mathbb{E}^{\mathbb{P}} [V_T^\gamma] &= \frac{v^\gamma}{\gamma} \mathbb{E}^{\hat{\mathbb{P}}} \left[e^{-\gamma \int_0^T \eta(s, X_s, \pi_s) ds + \gamma \int_0^T \pi_s^\top \Sigma_Y dW_s - \frac{\gamma^2}{2} \int_0^T \pi_s^\top \Sigma_Y \Sigma_Y^\top \pi_s ds} U_T \right] \\ &= \frac{v^\gamma}{\gamma} \mathbb{E}^{\hat{\mathbb{P}}} [L_T], \end{aligned} \quad (3.15)$$

where

$$L_t := \mathcal{E}_t\left(\int_0^\cdot Q(s, X_s, \pi_s)^\top \Sigma_Y d\hat{W}_s\right) U_t^{(2)} \exp\left(-\gamma \int_0^t \eta(s, X_s, \pi_s) ds\right), \quad (3.16)$$

$$Q(s, e_i, \pi_s) := (\Sigma_Y \Sigma_Y^\top)^{-1} \vartheta(s, e_i) + \gamma \pi_s = \left(\frac{1}{\sigma^2} \left(\mu_i - \frac{\sigma^2}{2}\right) + \gamma \pi_s^S, \frac{1}{v^2} \left(a(t, e_i) - \frac{v^2}{2}\right) + \gamma \pi_s^P\right)^\top. \quad (3.17)$$

Next, we proceed to give the filter probabilities and some useful related relationships, for which we first need to introduce some notation. Throughout, the unit N -simplex in \mathbb{R}^N is denoted by

$$\Delta_{N-1} = \{(d^1, d^2, \dots, d^N) : d^1 + d^2 + \dots + d^N = 1, d^i \geq 0, i = 1, \dots, N\}.$$

Let $g : D \rightarrow \mathbb{R}$, where $D = D_1 \times \{e_1, \dots, e_N\} \times D_2$, with D_1 and D_2 arbitrary, possibly empty, domains. The mapping $\hat{g} : D_1 \times \Delta_{N-1} \times D_2 \rightarrow \mathbb{R}$ is defined as

$$\hat{g}(y, d, z) = \sum_{i=1}^N g(y, e_i, z) d^i, \quad (3.18)$$

for each $y \in D_1$, $d \in \Delta_{N-1}$, and $z \in D_2$. Similarly, given a vector $l = (l_1, \dots, l_N) \in \mathbb{R}^N$, we define the associated mapping $\hat{l} : \Delta_{N-1} \rightarrow \mathbb{R}$ as

$$\hat{l}(d) := \sum_{i=1}^N l_i d^i. \quad (3.19)$$

Throughout, the filter probability that the regime X_t is e_i at time t , conditional on the filtration \mathcal{G}_t^I , is denoted by

$$p_t^i := \mathbb{P}(X_t = e_i | \mathcal{G}_t^I), \quad i = 1, \dots, N. \quad (3.20)$$

In particular, note the following useful relationships in terms of the transformations introduced in Eqs. (3.18)-(3.19):

$$\mathbb{E}^{\mathbb{P}} [g(y, X_t, z) | \mathcal{G}_t^I] = \hat{g}(y, p_t, z), \quad \mathbb{E}^{\mathbb{P}} [\langle l, X_t \rangle | \mathcal{G}_t^I] = \hat{l}(p_t), \quad (3.21)$$

where hereafter $p_t := (p_t^1, \dots, p_t^N)^\top$. We then have the following fundamental result, proven in Frey and Schmidt (2012):

Proposition 3.1 (Proposition 3.6 in Frey and Schmidt (2012)) *The normalized filter probabilities $p_t := (p_t^1, \dots, p_t^N)^\top$ are governed by the SDE*

$$\begin{aligned} dp_t^i &= \sum_{\ell=1}^N \varpi_{\ell,i}(t) p_t^\ell dt + p_t^i (\vartheta(t, e_i)^\top - \hat{\vartheta}(t, p_t)^\top) (\Sigma_Y \Sigma_Y^\top)^{-1} (dY_t - \hat{\vartheta}(t, p_t) dt) \\ &\quad + p_{t-}^i \frac{h_i - \hat{h}(p_{t-})}{\hat{h}(p_{t-})} \left(dH_t - \hat{h}(p_{t-}) \bar{H}_{t-} dt \right), \end{aligned} \quad (3.22)$$

with initial condition $p_0^i = p^{\circ,i}$.

Note that since $\hat{h}(p_t) \geq \min_i \{h_i\} > 0$, there is no singularity arising in the filtering equation (3.22). We remark that Frey and Runggaldier (2010), Section 4.1, also consider filter equations for finite-state Markov chains in the presence of multiple default events. However, they provide the dynamics of the unnormalized filter probabilities using a Zakai-type SDE, and then construct an algorithm to compute the filter probabilities.

Remark 3.2 Uniqueness of a strong solution to the system (3.22) was also established in Frey and Schmidt (2012) (see the discussion following Eqs. (A2)-(A-3) therein), building on results of Kliemann et al. (1990).

We are now ready to give the main result of this section. Define

$$\hat{L}_t = \mathcal{E}_t \left(\int_0^t \hat{Q}(s, p_s, \pi_s)^\top \Sigma_Y d\hat{W}_s \right) \mathcal{E}_t \left(\int_0^t (\hat{h}(p_{s-}) - 1) d\hat{\xi}_s \right) e^{-\gamma \int_0^t \hat{\eta}(s, p_s, \pi_s) ds}, \quad (3.23)$$

where, using the notation (3.18) and recalling the definitions of η and Q given in Eqs. (3.6) and (3.17), respectively,

$$\begin{aligned} \hat{\eta}(t, p_t, \pi_t) &= \sum_{i=1}^N \eta(t, e_i, \pi_t) p_t^i = -r + \pi_t^S (r - \hat{\mu}(p_t)) + \pi_t^P (r - \hat{a}(t, p_t)) + \frac{1-\gamma}{2} (\sigma^2 (\pi_t^S)^2 + v^2 (\pi_t^P)^2) \\ \hat{Q}(t, p_t, \pi_t) &= \sum_{i=1}^N Q(t, e_i, \pi_t) p_t^i = \left(\frac{1}{\sigma^2} \left(\hat{\mu}(p_t) - \frac{\sigma^2}{2} \right) + \gamma \pi_t^S, \frac{1}{v^2} \left(\hat{a}(t, p_t) - \frac{v^2}{2} \right) + \gamma \pi_t^P \right)^\top. \end{aligned}$$

Using the definitions of the stochastic exponentials given in (3.8) and (3.9), it follows from a direct application of Itô's formula that

$$d\hat{L}_t = \hat{L}_{t-} \left(\hat{Q}(t, p_t, \pi_t)^\top dY_t + (\hat{h}(p_{t-}) - 1) d\hat{\xi}_t \right) - \gamma \hat{\eta}(t, p_t, \pi_t) \hat{L}_t dt. \quad (3.24)$$

Then, we have the following crucial result, whose proof is reported in Appendix A.

Proposition 3.3 *It holds that*

$$J(v, \pi, T) = \frac{v^\gamma}{\gamma} \mathbb{E}^{\hat{\mathbb{P}}} \left[\hat{L}_T \right]. \quad (3.25)$$

The representation in (3.25) establishes the correspondence between the original partially observed control problem (criterion (3.4) depending on the hidden state X_t) and a complete observation control problem (criterion (3.25) depending on the observed (filter) probabilities p_t).

3.2 The Risk-Sensitive Control Problem

The objective of this section is to show how the complete observation control problem (3.25) may be reduced to a risk-sensitive stochastic control problem. Such a representation proves to be useful for analyzing the control problem via the HJB approach in the next section. The reduction is obtained building on the approach of Nagai and Runggaldier (2008) who do not consider the defaultable security. Next, we develop the change to the new measure $\tilde{\mathbb{P}}$, so to write the criterion (3.25) in the risk sensitive form. The measure change from $\hat{\mathbb{P}}$ to $\tilde{\mathbb{P}}$ is defined via its Radon-Nikodym density as follows:

$$\frac{d\tilde{\mathbb{P}}}{d\hat{\mathbb{P}}} \Big|_{\mathcal{G}_t^I} := \zeta_t := \mathcal{E}_t \left(\int_0^t \hat{Q}(s, p_s, \pi_s)^\top \Sigma_Y d\hat{W}_s \right) \mathcal{E}_t \left(\int_0^t (\hat{h}(p_{s-}) - 1) d\hat{\xi}_s \right). \quad (3.26)$$

Note that the probability measure $\tilde{\mathbb{P}}$ depends, through ζ_t , on the strategy π_t . Hence, in order for $\tilde{\mathbb{P}}$ to be a probability measure, we need to require that the set of admissible strategies satisfies the condition

$$\mathbb{E}^{\tilde{\mathbb{P}}}[\zeta_T] = \mathbb{E}^{\mathbb{P}}[\rho_T \zeta_T] = 1. \quad (3.27)$$

In order to impose (3.27), we again use the general Novikov's condition (3.13). In this case, it is easy to check $\rho_t \zeta_t = \mathcal{E}_t(M)$ with

$$M_t = \int_0^t \left(\hat{Q}(s, p_s, \pi_s)^\top \Sigma_Y - \vartheta(s, X_s)^\top \Sigma_Y^{-1} \right) dW_s + \int_0^t \frac{\hat{h}(p_{s-}) - h_{s-}}{h_{s-}} d\xi_s,$$

and, thus,

$$\begin{aligned} \langle M^c, M^c \rangle_T &\leq 2 \int_0^T \hat{Q}^\top \Sigma_Y \Sigma_Y^\top \hat{Q}(s, p_s, \pi_s) ds + 2 \int_0^T \vartheta^\top (\Sigma_Y \Sigma_Y)^{-1} \vartheta(s, X_s) ds \\ \langle M^d, M^d \rangle_T &= \int_0^{\tau \wedge T} \left(\frac{\hat{h}(p_s) - h_s}{h_s} \right)^2 h_s ds. \end{aligned}$$

The second term in the expression of $\langle M^c, M^c \rangle_T$ is uniformly bounded in view of the condition (2.9), while $\langle M^d, M^d \rangle_T$ is also bounded since the integrand therein is bounded by $2 \max_i \{h_i^2\} / \min_i \{h_i\}$. Therefore, we only need to require that $\mathbb{E}^{\mathbb{P}} \left[e^{\frac{1}{2} \int_0^T \hat{Q}^\top \Sigma_Y \Sigma_Y^\top \hat{Q}(s, p_s, \pi_s) ds} \right] < \infty$, for which it suffices that $\pi = (\pi^S, \pi^P)^\top$ meets the integrability condition:

$$\mathbb{E}^{\mathbb{P}} \left[e^{\frac{\sigma^2 \gamma^2}{2} \int_0^T (\pi^S)^2 ds + \frac{v^2 \gamma^2}{2} \int_0^T (\pi^P)^2 ds} \right] < \infty. \quad (3.28)$$

Once we have established conditions for the validity of the probability transformation (3.26), we can then apply Girsanov's theorem (cf. Bielecki and Rutkowski (2001)) to conclude that

$$\tilde{W}_t = \hat{W}_t - \int_0^t \Sigma_Y^\top \hat{Q}(s, p_s, \pi_s) ds$$

is a \mathbb{G}^J -Brownian motion under $\tilde{\mathbb{P}}$, while

$$\tilde{\xi}_t = \hat{\xi}_t - \int_0^{t \wedge \tau} (\hat{h}(p_s) - 1) ds = H_t - \int_0^{t \wedge \tau} \hat{h}(p_s) ds = H_t - \int_0^t \hat{h}(p_s) \bar{H}_{s-} ds, \quad (3.29)$$

is a \mathbb{G}^J martingale under $\tilde{\mathbb{P}}$. It then follows immediately that

$$J(v, \pi, T) = \frac{v^\gamma}{\gamma} \mathbb{E}^{\tilde{\mathbb{P}}} \left[\hat{L}_T \zeta_T^{-1} \right] = \frac{v^\gamma}{\gamma} \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{-\gamma \int_0^T \hat{\eta}(s, p_s, \pi_s) ds} \right], \quad (3.30)$$

where the dynamics of p_t^i in Eq. (3.22) may be rewritten under the measure $\tilde{\mathbb{P}}$ as

$$\begin{aligned} dp_t^i &= p_t^i \left(\vartheta(t, e_i)^\top - \hat{\vartheta}(t, p_t)^\top \right) \Sigma_Y^{-1} d\tilde{W}_t \\ &+ \left(\sum_{\ell=1}^N \varpi_{\ell, i}(t) p_t^\ell + \gamma p_t^i \left(\vartheta(t, e_i)^\top - \hat{\vartheta}(t, p_t)^\top \right) \pi_t \right) dt + p_t^i \frac{h_i - \hat{h}(p_{t-})}{\hat{h}(p_{t-})} d\tilde{\xi}_t, \quad t \geq 0. \end{aligned} \quad (3.31)$$

Hence, the overall conclusion is that the original problem is reduced to a risk sensitive control problem of the form:

$$\sup_{\pi} J(v; \pi; T) = \frac{v^\gamma}{\gamma} \sup_{\pi} \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{-\gamma \int_0^T \hat{\eta}(s, p_s, \pi_s) ds} \right], \quad (3.32)$$

where the maximization is done across suitable strategies $(\pi_t)_t$. We shall specify later on the precise class of trading strategies π on which the portfolio optimization problem is defined.

Remark 3.4 As customary with Markovian optimal control problems, we will solve the risk-sensitive control problem starting at time $t = 0$ by embedding it into a dynamical control problem starting at any time $t \in [0, T]$. Roughly speaking, the latter problem can be seen as the original problem (3.32) but starting at time $t \in (0, T]$ instead of 0. In order to formally define the dynamical problem, we consider a family of SDEs indexed by t of the form:

$$dp_s^{t,i} = p_s^{t,i} \left(\vartheta(s, e_i)^\top - \hat{\vartheta}(s, p_s^t)^\top \right) \Sigma_Y^{-1} d\tilde{W}_s^t + \left(\sum_{\ell=1}^N \varpi_{\ell,i}(s) p_s^{t,\ell} + \gamma p_s^{t,i} \left(\vartheta(s, e_i)^\top - \hat{\vartheta}(s, p_s^t)^\top \right) \pi_s^t \right) ds + p_s^{t,i} \frac{h_i - \hat{h}(p_{s-}^t)}{\hat{h}(p_{s-}^t)} d\tilde{\xi}_s^t, \quad s \in (t, T], \quad (3.33)$$

with initial condition $p_t^{t,i} = p^{\circ,i}$, defined on a suitable space $(\Omega^t, \mathcal{G}^t, \mathbb{G}^t, \tilde{\mathbb{P}}^t)$, equipped with a Wiener process $\{\tilde{W}_s^t\}_{s \geq t}$ starting at t and, an independent, one-point counting process $\{H_s^t\}_{s \geq t}$ such that $H_t^t = z^\circ \in \{0, 1\}$ and

$$\tilde{\xi}_s^t := H_s^t - \int_t^s \hat{h}(p_u^t) \bar{H}_{u-}^t du, \quad s \geq t, \quad (3.34)$$

is a $\tilde{\mathbb{P}}^t$ -martingale. Hereafter, the construction of the process $\{p_s^t\}_{s \in [t, T]} := \{(p_s^{t,1}, \dots, p_s^{t,N})^\top\}_{s \in [t, T]}$ is carried out in a similar way as the construction of the solution to (3.31). Concretely, start defining the process $\{p_s^t\}_{s \geq 0} := \{(p_s^{t,1}, \dots, p_s^{t,N})^\top\}_{s \geq 0}$ via the representation (3.20), which, analogously to $\{p_s\}_{s \geq 0}$ follows the SDE (3.22), is the solution of a system of SDE's of the form:

$$dp_s^{t,i} = \sum_{\ell=1}^N \varpi_{\ell,i}(s) p_s^{t,\ell} ds + p_s^{t,i} (\vartheta(s, e_i)^\top - \hat{\vartheta}(s, p_s^t)^\top) (\Sigma_Y \Sigma_Y^\top)^{-1} (dY_s^t - \hat{\vartheta}(s, p_s^t) ds) + p_s^{t,i} \frac{h_i - \hat{h}(p_{s-}^t)}{\hat{h}(p_{s-}^t)} (dH_s^t - \hat{h}(p_{s-}^t) \bar{H}_{s-}^t ds), \quad s \geq t, \quad (3.35)$$

on a probability space $(\Omega^t, \mathcal{G}^t, \mathbb{G}^t, \mathbb{P}^t)$. Here, $Y_s^t := (\log(S_s^t), \log(P_s^t))^\top$, $s \geq t$, with $\{S_s^t\}_{s \geq t}$ and $\{P_s^t\}_{s \geq t}$ defined analogously to (2.4) and (2.8):

$$dS_s^t = \langle \mu, X_s^t \rangle S_s^t ds + \sigma S_s^t dW_s^{(1,t)}, \quad s > t, \quad S_t^t = S^\circ, \\ dP_s^t = P_{s-}^t \left(a(s, X_s^t) ds + v dW_s^{(2,t)} - dH_s^t \right), \quad s > t, \quad P_t^t = P^\circ.$$

The hidden Markov chain $\{X_s^t\}_{s \geq t}$ has initial distribution $p_t^{t,i} = \mathbb{P}^t(X_t^t = e_i) = p^{\circ,i}$ and generator $A(s) := [\varpi_{i,j}(s)]_{i,j=1, \dots, N}$, $s \geq t$. Once we have defined the process $\{p_s^t\}_{s \geq t}$, we proceed to define $\tilde{\mathbb{P}}^t$ in terms of a suitable trading strategy $\{\pi_s^t\}_{s \in [t, T]} := \{(\pi_s^{t,S}, \pi_s^{t,P})^\top\}_{s \in [t, T]}$ analogously to $\tilde{\mathbb{P}}$, and processes $\{\tilde{W}^t\}_{s \geq t}$ and $\{\tilde{\xi}^t\}_{s \geq t}$ analogously to \tilde{W} and $\tilde{\xi}$ so that, under $\tilde{\mathbb{P}}^t$, the process $p^t = (p^{t,1}, \dots, p^{t,N})^\top$ satisfies (3.33). Note that the existence of the measure transformation $\tilde{\mathbb{P}}^t$ and, hence, of the solution to the SDE (3.33) is guaranteed provided that π^t satisfies the analogous of (3.28):

$$\mathbb{E}^{\mathbb{P}^t} \left[e^{\frac{\sigma^2 \gamma^2}{2} \int_t^T (\pi_s^{t,S})^2 ds + \frac{v^2 \gamma^2}{2} \int_t^T (\pi_s^{t,P})^2 ds} \right] < \infty. \quad (3.36)$$

4 HJB formulation

This section is devoted to formulating the HJB equation. Given that the filter probability process $p_s = (p_s^1, \dots, p_s^N)$ is degenerate in \mathbb{R}^N , we consider the projected $N - 1$ dimensional process

$$\tilde{p}_s := (\tilde{p}_s^1, \dots, \tilde{p}_s^{N-1})^\top := (p_s^1, \dots, p_s^{N-1})^\top,$$

as opposed to the actual filtering process. Next, we rewrite the problem (3.30) in terms of the above process, which now lies in the space

$$\tilde{\Delta}_{N-1} = \{(d^1, \dots, d^{N-1}) : d^1 + \dots + d^{N-1} < 1, d^i > 0\},$$

in view of the Lemma B.1 below. Let us start with some notation needed to write the SDE of \tilde{p} in matrix form. First, similarly to (3.18) and (3.19), given a vector $l = (l_1, \dots, l_N) \in \mathbb{R}^N$ and a function $g : D \rightarrow \mathbb{R}$,

where $D = D_1 \times \{e_1, \dots, e_N\} \times D_2$, with D_1 and D_2 arbitrary, possibly empty domains, define the mappings $\tilde{g} : D_1 \times \tilde{\Delta}_{N-1} \times D_2 \rightarrow \mathbb{R}$ and $\tilde{l} : \tilde{\Delta}_{N-1} \rightarrow \mathbb{R}$ as follows:

$$\tilde{g}(y, d, z) = g(y, e_N, z) + \sum_{i=1}^{N-1} [g(y, e_i, z) - g(y, e_N, z)] d^i, \quad \tilde{l}(d) := l_N + \sum_{i=1}^{N-1} [l_i - l_N] d^i, \quad (4.1)$$

for $d = (d^1, \dots, d^{N-1}) \in \tilde{\Delta}_{N-1}$ and $y \in D_1, z \in D_2$. The following relationships are useful in what follows. For $y \in D_1, z \in D_2$,

$$\hat{g}(y, p_t, z) = \tilde{g}(y, \tilde{p}_t, z), \quad \hat{l}(p_t) = \tilde{l}(\tilde{p}_t). \quad (4.2)$$

Throughout, the projection of a vector $l = (l_1, \dots, l_N)$ on the first $N - 1$ coordinates is denoted by $l' := (l_1, \dots, l^{N-1})$. Similarly, for a given matrix B , we use B' to denote the projection on the submatrix consisting of the first $N - 1$ columns. Hence,

$$\vartheta(t)' := (\vartheta(t, e_1), \dots, \vartheta(t, e_{N-1})) = \begin{pmatrix} \mu_1 - \frac{\sigma^2}{2}, & \dots, & \mu_{N-1} - \frac{\sigma^2}{2} \\ a(t, e_1) - \frac{v^2}{2}, & \dots, & a(t, e_{N-1}) - \frac{v^2}{2} \end{pmatrix}.$$

We use $\text{Diag}(\mathbf{b})$ to denote the diagonal matrix, whose i^{th} diagonal element is the i^{th} component of the vector \mathbf{b} . Further, let $\beta_\varpi(t, \tilde{p}_t)$ be the $(N - 1) \times 1$ vector defined by

$$\beta_\varpi(t, \tilde{p}_t) = \left(\varpi_{N,1}(t) + \sum_{\ell=1}^{N-1} [\varpi_{\ell,1}(t) - \varpi_{N,1}(t)] \tilde{p}_t^\ell, \dots, \varpi_{N,N-1}(t) + \sum_{\ell=1}^{N-1} [\varpi_{\ell,N-1}(t) - \varpi_{N,N-1}(t)] \tilde{p}_t^\ell \right)^\top. \quad (4.3)$$

Finally, we also use $\mathbf{1}$ to denote the $N - 1$ dimensional column vector whose entries are all ones.

In what follows, we work with the collection of processes $\{p_s^t\}_{t \leq s \leq T} = \{(p_s^{t,1}, \dots, p_s^{t,N})^\top\}_{t \leq s \leq T}$ constructed on a suitable probability space $(\Omega^t, \mathcal{G}^t, \mathbb{G}^t, \mathbb{P}^t)$ as described in Remark 3.4. Using Eq. (3.33), the dynamics of the vector process $\tilde{p}_s^t := (\tilde{p}_s^{t,1}, \dots, \tilde{p}_s^{t,N-1}) := (p_s^{t,1}, \dots, p_s^{t,N-1})$ under $\tilde{\mathbb{P}}^t$ may be rewritten as

$$\begin{aligned} d\tilde{p}_s^t &= \text{Diag}(\tilde{p}_s^t) \left(\vartheta(s)' - \mathbf{1} \tilde{\vartheta}(s, \tilde{p}_s^t) \right)^\top \Sigma_Y^{-1} d\tilde{W}_s^t + \beta_\varpi(s, \tilde{p}_s^t) ds \\ &\quad + \gamma \text{Diag}(\tilde{p}_s^t) \left(\vartheta(s)' - \mathbf{1} \tilde{\vartheta}(s, \tilde{p}_s^t) \right)^\top \pi_s^t ds + \text{Diag}(\tilde{p}_{s-}^t) \frac{1}{\tilde{h}(\tilde{p}_{s-}^t)} \left(h' - \mathbf{1} \tilde{h}(\tilde{p}_{s-}^t) \right) d\tilde{\xi}_s^t, \quad t < s \leq T, \\ \tilde{p}_t^t &= \tilde{p}^\circ, \end{aligned}$$

where the initial value of the process is $\tilde{p}^\circ := (\tilde{p}^{\circ,1}, \dots, \tilde{p}^{\circ,N-1})^\top = (p^{\circ,1}, \dots, p^{\circ,N-1})^\top$. Next, let us define

$$\begin{aligned} \kappa(s, \tilde{p}_s^t) &:= \text{Diag}(\tilde{p}_s^t) \left(\vartheta(s)' - \mathbf{1} \tilde{\vartheta}(s, \tilde{p}_s^t) \right)^\top \Sigma_Y^{-1} \\ \beta_\gamma(s, \tilde{p}_s^t, \pi_s^t) &:= \beta_\varpi(s, \tilde{p}_s^t) + \gamma \kappa(s, \tilde{p}_s^t) \Sigma_Y^\top \pi_s^t, \\ \varrho(\tilde{p}_{s-}^t) &:= \text{Diag}(\tilde{p}_{s-}^t) \frac{1}{\tilde{h}(\tilde{p}_{s-}^t)} \left(h' - \mathbf{1} \tilde{h}(\tilde{p}_{s-}^t) \right). \end{aligned}$$

Then, the dynamics of $\tilde{p}_s^t = (\tilde{p}_s^{t,1}, \dots, \tilde{p}_s^{t,N-1})^\top$ for $s \in [t, T]$, under the probability measure $\tilde{\mathbb{P}}^t$, is given by

$$d\tilde{p}_s^t = \beta_\gamma(s, \tilde{p}_s^t, \pi_s^t) ds + \kappa(s, \tilde{p}_s^t) d\tilde{W}_s^t + \varrho(\tilde{p}_{s-}^t) d\tilde{\xi}_s^t \quad (t < s \leq T), \quad \tilde{p}_t^t = \tilde{p}^\circ \in \tilde{\Delta}_{N-1}. \quad (4.4)$$

A similar expression may be written for the vector $\tilde{p}_s := (\tilde{p}_s^1, \dots, \tilde{p}_s^{N-1})^\top := (p_s^1, \dots, p_s^{N-1})^\top$ solving Eq. (3.31), which lives in the ‘‘real world’’ space $(\Omega, \mathcal{G}, \mathbb{G}, \mathbb{P})$ and starts at time 0.

Note that π^t affects the evolution of $\{\tilde{p}_s^t\}_{t \leq s \leq T}$ through the drift β_γ , and also through the measure $\tilde{\mathbb{P}}^t$ due to an admissibility constraint analogous to (3.27). As explained in Remark 3.4, the condition (3.27) is satisfied provided that (3.36) holds true. In light of these observations, the following class of admissible controls is natural.

Definition 4.1 The class of admissible strategies $\mathcal{A}(t, T; \tilde{p}^\circ, z^\circ)$ consists of locally bounded feedback trading strategies

$$\pi_s^t := (\pi_s^{t,S}, \pi_s^{t,P}) = (\pi^{t,S}(s, \tilde{p}_{s-}^t, H_s^t), \pi^{t,P}(s, \tilde{p}_{s-}^t, H_s^t))$$

for $t < s \leq T$ and $\pi_t^t := (\pi_t^{t,S}, \pi_t^{t,P}) = (\pi^{t,S}(t, \tilde{p}_t^t, H_t^t), \pi^{t,P}(t, \tilde{p}_t^t, H_t^t)) = (\pi^S(t, \tilde{p}^\circ, z^\circ), \pi^P(t, \tilde{p}^\circ, z^\circ))$, satisfying

$$\mathbb{E}^{\mathbb{P}^t} \left[\exp \left(\frac{\sigma^2 \gamma^2}{2} \int_t^T (\pi^{t,S}(s, \tilde{p}_{s-}^t, H_s^t))^2 ds + \frac{v^2 \gamma^2}{2} \int_t^T (\pi^{t,P}(s, \tilde{p}_{s-}^t, H_s^t))^2 ds \right) \right] < \infty. \quad (4.5)$$

so to guarantee that the measure change defined by (3.26) is well defined.

Let us now define the dynamic programming problem associated with our original utility maximization problem. For each $t \in [0, T)$, $\tilde{p}^\circ \in \tilde{\Delta}_{N-1}$, $z^\circ \in \{0, 1\}$, and Markov strategy $\pi^t \in \mathcal{A}(t, T; \tilde{p}^\circ, z^\circ)$, we set

$$G(t, \tilde{p}^\circ, z^\circ, \pi^t) := \mathbb{E}^{\tilde{\mathbb{P}}^t} \left[e^{-\gamma \int_t^T \tilde{\eta}(s, \tilde{p}_s^t, \pi_s^t) ds} \right], \quad (4.6)$$

where we recall that by construction $H_t^t = z^\circ$, \tilde{p}^t is given as in (4.4), and $\tilde{\eta}$ is defined from η in accordance to (4.1) as

$$\tilde{\eta}(s, \tilde{p}_s^t, \pi_s^t) = \eta(s, e_N, \pi_s^t) + \sum_{i=1}^{N-1} (\eta(s, e_i, \pi_s^t) - \eta(s, e_N, \pi_s^t)) \tilde{p}_s^{t,i}.$$

Next, we define the value function

$$w(t, \tilde{p}^\circ, z^\circ) := \sup_{\pi^t \in \mathcal{A}(t, T; \tilde{p}^\circ, z^\circ)} \log(G(t, \tilde{p}^\circ, z^\circ, \pi^t)). \quad (4.7)$$

The crucial step to link the above dynamic programming problem with our original problem is outlined next:

$$\begin{aligned} \sup_{\pi \in \mathcal{A}(0, T; \tilde{p}^\circ, z^\circ)} J(v; \pi; T) &= \frac{v^\gamma}{\gamma} \sup_{\pi \in \mathcal{A}(0, T; \tilde{p}^\circ, z^\circ)} \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{-\gamma \int_0^T \hat{\eta}(s, p_s, \pi_s) ds} \right] \\ &= \frac{v^\gamma}{\gamma} \sup_{\pi \in \mathcal{A}(0, T; \tilde{p}^\circ, z^\circ)} \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{-\gamma \int_0^T \tilde{\eta}(s, \tilde{p}_s, \pi_s) ds} \right] \\ &= \frac{v^\gamma}{\gamma} \sup_{\pi^0 \in \mathcal{A}(0, T; \tilde{p}^\circ, z^\circ)} \mathbb{E}^{\tilde{\mathbb{P}}^0} \left[e^{-\gamma \int_0^T \tilde{\eta}(s, \tilde{p}_s^0, \pi_s^0) ds} \right] \\ &= \frac{v^\gamma}{\gamma} e^{w(0, \tilde{p}^\circ, 1, \dots, \tilde{p}^\circ, N-1, z^\circ)}, \end{aligned}$$

where the first and second equalities follow from Eqs. (3.30) and (4.2), while the third equality follows from the uniqueness of the strong solution to the system (3.35), which can be established as noticed in Remark 3.2.

We now proceed to derive the HJB equation corresponding to the value function in Eq. (4.7). Before doing so, we need to compute the generator \mathcal{L} of the Markov process $s \in [t, T] \rightarrow (s, \tilde{p}_s^t, H_s^t)$. This is done in the following lemma. Below and hereafter, we use $\mathbf{y} \cdot \mathbf{h}$ to denote componentwise multiplication of two vectors \mathbf{y} and \mathbf{h} .

Lemma 4.2 Let $(\tilde{p}_s^t)_{s \in [t, T]}$ be the process in (4.4) with π^t of the form

$$\pi_s^t := \pi(s, \tilde{p}_{s-}^t, H_s^t) \quad (t < s \leq T), \quad \pi_t^t := \pi(t, \tilde{p}^\circ, z^\circ),$$

for a suitable function $\pi(s, \tilde{p}, z)$ such that (4.4) admits a unique strong solution. Then, for any $f(t, \tilde{p}, z)$ such that $f(t, \tilde{p}, 1)$ and $f(t, \tilde{p}, 0)$ are both $C^{1,2}$ -functions, we have

$$f(s, \tilde{p}_s^t, H_s^t) = f(t, \tilde{p}^\circ, z^\circ) + \int_t^s \tilde{\mathcal{L}}f(u, \tilde{p}_u^t, H_u^t) du + \tilde{M}_s(f), \quad s \in (t, T], \quad (4.8)$$

where, denoting $\nabla_{\tilde{p}} f(t, \tilde{p}, z) := (\frac{\partial f}{\partial \tilde{p}^1}, \dots, \frac{\partial f}{\partial \tilde{p}^{N-1}})$, $f_t(t, \tilde{p}, z) := \frac{\partial f}{\partial t}$, and $D^2 f := \left[\frac{\partial^2 f}{\partial \tilde{p}^i \partial \tilde{p}^j} \right]_{i,j=1}^{N-1}$ and recalling the notation $\tilde{h}(\tilde{p}) := h^N + \sum_{i=1}^{N-1} (h_i - h_N) \tilde{p}^i$ and $h' := (h_1, \dots, h_{N-1})^\top$,

$$\begin{aligned} \tilde{\mathcal{L}}f(t, \tilde{p}, z) &:= f_t(t, \tilde{p}, z) + \nabla_{\tilde{p}} f \beta_\gamma(t, \tilde{p}, \pi(t, \tilde{p}, z)) + \frac{1}{2} \text{tr}(\kappa \kappa^\top D^2 f(t, \tilde{p}, z)) \\ &\quad + (1-z) \left(f \left(t, \frac{1}{\tilde{h}(\tilde{p})} (\tilde{p} \cdot h'), 1 \right) - f(t, \tilde{p}, 0) \right) \tilde{h}(\tilde{p}). \end{aligned}$$

Moreover, the $\tilde{\mathbb{P}}^t$ -local martingale component is

$$\tilde{M}_s(f) = \int_t^s \nabla_{\tilde{p}} f \kappa(u, \tilde{p}_u^t) d\tilde{W}_u^t + \int_t^s \left(f \left(u, \frac{1}{\tilde{h}(\tilde{p}_{u-}^t)} (\tilde{p}_{u-}^t \cdot h'), 1 \right) - f(u, \tilde{p}_{u-}^t, 0) \right) d\tilde{\xi}_u^t. \quad (4.9)$$

Proof For simplicity, throughout the proof we drop the superscript t in the processes \tilde{p}^t , π^t , \tilde{W}^t , $\tilde{\xi}^t$, and H^t . Let $\tilde{p}^{c,i}$ denote the continuous component of \tilde{p}^i , determined by the first two terms on the right-hand side of Eq. (4.4). Using Itô's formula, we have

$$\begin{aligned} f(s, \tilde{p}_s, H_s) &= f(t, \tilde{p}_t, H_t) + \int_t^s f_u(u, \tilde{p}_u, H_u) du + \sum_{i=1}^{N-1} \int_t^s \frac{\partial f}{\partial \tilde{p}^i} d\tilde{p}_u^{c,i} + \frac{1}{2} \sum_{i,j=1}^{N-1} \int_t^s \frac{\partial^2 f}{\partial \tilde{p}^i \partial \tilde{p}^j} d\langle \tilde{p}^{c,i}, \tilde{p}^{c,j} \rangle_u \\ &\quad + \sum_{t < u \leq s} (f(u, \tilde{p}_u, H_u) - f(u, \tilde{p}_{u-}, H_{u-})). \end{aligned} \quad (4.10)$$

Note that the size of the jump of \tilde{p}_τ^i at the default time τ is given by

$$\tilde{p}_\tau^i - \tilde{p}_{\tau-}^i = \tilde{p}_{\tau-}^i \frac{h_i - \tilde{h}(\tilde{p}_{\tau-})}{\tilde{h}(\tilde{p}_{\tau-})}, \quad (4.11)$$

thus implying that $\tilde{p}_\tau^i = \tilde{p}_{\tau-}^i h_i / \tilde{h}(\tilde{p}_{\tau-})$ and $\tilde{p}_\tau = (1/\tilde{h}(\tilde{p}_{\tau-}))(\tilde{p}_{\tau-} \cdot h')$. For $t < \tau \leq s$, this leads to

$$\begin{aligned} \sum_{t < u \leq s} (f(u, \tilde{p}_u, H_u) - f(u, \tilde{p}_{u-}, H_{u-})) &= \left[f \left(\tau, \frac{1}{\tilde{h}(\tilde{p}_{\tau-})} (\tilde{p}_{\tau-} \cdot h'), 1 \right) - f(\tau, \tilde{p}_{\tau-}, 0) \right] (H_s - H_t) \\ &= \int_t^s \left(f \left(u, \frac{1}{\tilde{h}(\tilde{p}_{u-})} (\tilde{p}_{u-} \cdot h'), 1 \right) - f(u, \tilde{p}_{u-}, 0) \right) dH_u \\ &= \int_t^s \left(f \left(u, \frac{1}{\tilde{h}(\tilde{p}_{u-})} (\tilde{p}_{u-} \cdot h'), 1 \right) - f(u, \tilde{p}_{u-}, 0) \right) (d\tilde{\xi}_u + \bar{H}_u - \tilde{h}(\tilde{p}_{u-}) du), \end{aligned}$$

where in the last equality we have used Eq. (3.29) and the fact that $\hat{h}_u = \sum_{i=1}^N h_i p_u^i = h_N + \sum_{i=1}^{N-1} (h_i - h_N) p_u^i = \tilde{h}(\tilde{p}_u)$. From this, we deduce that Eq. (4.10) may be rewritten as

$$\begin{aligned} f(s, \tilde{p}_s, H_s) &= f(t, \tilde{p}_t, H_t) + \int_t^s f_u(u, \tilde{p}_u, H_u) du + \int_t^s \nabla_{\tilde{p}} f \beta_\gamma(u, \tilde{p}_u, \pi_u) du + \frac{1}{2} \sum_{i,j=1}^{N-1} \int_t^s (\kappa \kappa^\top)_{ij} \frac{\partial^2 f}{\partial \tilde{p}^i \partial \tilde{p}^j}(u, \tilde{p}_u, H_u) du \\ &\quad + \int_t^s \nabla_{\tilde{p}} f \kappa(u, \tilde{p}_u) d\tilde{W}_u + \int_t^s \left(f \left(u, \frac{1}{\tilde{h}(\tilde{p}_{u-})} (\tilde{p}_{u-} \cdot h'), 1 \right) - f(u, \tilde{p}_{u-}, 0) \right) d\tilde{\xi}_u \\ &\quad + \int_t^s \left(f \left(u, \frac{1}{\tilde{h}(\tilde{p}_u)} (\tilde{p}_u \cdot h'), 1 \right) - f(u, \tilde{p}_u, 0) \right) (1 - H_u) \tilde{h}(\tilde{p}_u) du, \end{aligned} \quad (4.12)$$

which proves the lemma.

We are now ready to derive the HJB equation associated to the control problem. We first obtain it based on standard heuristic arguments, and then in the subsequent section we provide rigorous verification theorems for the solution. In light of the dynamic programming principle, we expect that, for any $s \in (t, T]$,

$$w(t, \tilde{p}^\circ, z^\circ) = \sup_{\pi^t \in \mathcal{A}(t, T; \tilde{p}^\circ, z^\circ)} \log \mathbb{E}^{\tilde{\mathbb{P}}^t} \left[e^{w(s, \tilde{p}_s^t, H_s^t) - \gamma \int_t^s \tilde{\eta}(u, \tilde{p}_u^t, \pi_u^t) du} \right]. \quad (4.13)$$

with $\tilde{p}_t^t = \tilde{p}^\circ$, and $H_t^t = z^\circ$. Next, define $\varepsilon(s, \tilde{p}, z) = e^{w(s, \tilde{p}, z)}$ and note that, in light of Lemma 4.2,

$$\varepsilon(s, \tilde{p}_s^t, H_s^t) = \varepsilon(t, \tilde{p}^\circ, z^\circ) + \int_t^s \tilde{\mathcal{L}}\varepsilon(u, \tilde{p}_u^t, H_u^t) du + \tilde{M}_s(\varepsilon),$$

where the last term $\widetilde{M}_s(\varepsilon)$ represents the local martingale component of $\varepsilon(s, \widetilde{p}_s^t, H_s^t)$. Plugging the previous equation into (4.13), we expect the following relation to hold:

$$0 = \sup_{\pi^t \in \mathcal{A}(t, T; \widetilde{p}^\circ, z^\circ)} \mathbb{E}^{\widetilde{p}^t} \left[\varepsilon(t, \widetilde{p}^\circ, H^\circ) \left(e^{-\gamma \int_t^s \widetilde{\eta}(u, \widetilde{p}_u^t, \pi_u^t) du} - 1 \right) + e^{-\gamma \int_t^s \widetilde{\eta}(u, \widetilde{p}_u^t, \pi_u^t) du} \int_t^s \widetilde{\mathcal{L}}\varepsilon(u, \widetilde{p}_u^t, H_u^t) du \right],$$

assuming that the local martingale component is a true martingale. Dividing by $s - t$ and taking the limit of the above expression as $s \rightarrow t$ leads us to the HJB equation:

$$0 = \sup_{\pi} \left[\left(\widetilde{\mathcal{L}} - \gamma \widetilde{\eta}(t, \widetilde{p}^\circ, \pi) \right) \varepsilon(t, \widetilde{p}^\circ, z^\circ) \right]. \quad (4.14)$$

Let us write (4.14) in terms of w . To this end, let us denote the differential component of $\widetilde{\mathcal{L}}$ as $\widetilde{\mathcal{D}}$; i.e.,

$$\widetilde{\mathcal{D}}f(t, \widetilde{p}, z) := f_t(t, \widetilde{p}, z) + \nabla_{\widetilde{p}} f(t, \widetilde{p}, z) \beta_\gamma(t, \widetilde{p}, \pi(t, \widetilde{p}, z)) + \frac{1}{2} \text{tr}(\kappa \kappa^\top D^2 f(t, \widetilde{p}, z)).$$

Then, we note that

$$\begin{aligned} \widetilde{\mathcal{L}}\varepsilon(t, \widetilde{p}, z) &= \widetilde{\mathcal{D}}\varepsilon(t, \widetilde{p}, z) + (1 - z) \widetilde{h}(\widetilde{p}) \left(e^{w(t, \frac{1}{h(\widetilde{p})} \widetilde{p} \cdot h', 1)} - e^{w(t, \widetilde{p}, 0)} \right) \\ &= e^{w(t, \widetilde{p}, z)} \left(\widetilde{\mathcal{D}}w + \frac{1}{2} \|\nabla_{\widetilde{p}} w \kappa\|^2 + (1 - z) \widetilde{h}(\widetilde{p}) \left[e^{w(t, \frac{1}{h(\widetilde{p})} \widetilde{p} \cdot h', 1) - w(t, \widetilde{p}, 0)} - 1 \right] \right). \end{aligned} \quad (4.15)$$

Thus, Eq. (4.14) takes the form:

$$0 = \sup_{\pi} \left[\varepsilon(t, \widetilde{p}^\circ, z^\circ) \left(\widetilde{\mathcal{D}}w + \frac{1}{2} \|\nabla_{\widetilde{p}} w \kappa\|^2 + (1 - z^\circ) \widetilde{h}(\widetilde{p}^\circ) \left[e^{w(t, \frac{1}{h(\widetilde{p}^\circ)} \widetilde{p}^\circ \cdot h', 1) - w(t, \widetilde{p}^\circ, 0)} - 1 \right] - \gamma \widetilde{\eta}(t, \widetilde{p}^\circ, z^\circ, \pi) \right) \right]. \quad (4.16)$$

In order to get a more explicit form, let us recall that

$$\eta(t, e_i, \pi) = -r + \pi^S (r - \langle \mu, e_i \rangle) + \pi^P (r - a(t, e_i)) + \frac{1 - \gamma}{2} \pi^\top \Sigma_Y^\top \Sigma_Y \pi,$$

and note that

$$\begin{aligned} \widetilde{\eta}(t, \widetilde{p}, \pi) &= \eta(t, e_N, \pi(t, \widetilde{p}, z)) + \sum_{i=1}^{N-1} (\eta(t, e_i, \pi(t, \widetilde{p}, z)) - \eta(t, e_N, \pi(t, \widetilde{p}, z))) \widetilde{p}^i \\ &= -r + \pi^S (r - \widetilde{\mu}(\widetilde{p})) + \pi^P (r - \widetilde{a}(t, \widetilde{p})) + \frac{1 - \gamma}{2} \pi^\top \Sigma_Y^\top \Sigma_Y \pi. \end{aligned} \quad (4.17)$$

We can now rewrite Eq. (4.16) as

$$\begin{aligned} \frac{\partial w}{\partial t} + \frac{1}{2} \text{tr}(\kappa \kappa^\top D^2 w) + \frac{1}{2} (\nabla_{\widetilde{p}} w) \kappa \kappa^\top (\nabla_{\widetilde{p}} w)^\top + \gamma r + (1 - z^\circ) \widetilde{h}(\widetilde{p}^\circ) \left[e^{w(t, \frac{1}{h(\widetilde{p}^\circ)} \widetilde{p}^\circ \cdot h', 1) - w(t, \widetilde{p}^\circ, 0)} - 1 \right] \\ + \sup_{\pi} \left\{ (\nabla_{\widetilde{p}} w) \beta_\gamma - \gamma \pi^S (r - \widetilde{\mu}(\widetilde{p}^\circ)) - (1 - z^\circ) \gamma \pi^P (r - \widetilde{a}(t, \widetilde{p}^\circ)) - \frac{1}{2} \gamma (1 - \gamma) \pi^\top \Sigma_Y^\top \Sigma_Y \pi \right\} = 0, \end{aligned} \quad (4.18)$$

with terminal condition $w(T, \widetilde{p}^\circ, z^\circ) = 0$ and where all the derivatives of w are evaluated at $(t, \widetilde{p}^\circ, z^\circ)$.

Depending on whether or not default has occurred, we will have two separate optimization problems to solve. Indeed, after default has occurred, the investor cannot invest in the defaultable security and only allocates his wealth in the stock and risk-free asset. The next section analyzes in detail the two cases.

5 Solution to the Optimal Control Problem

We analyze the control problems developed in the previous section. We first decompose it into two related optimization subproblems: the post and the pre-default problems. As we will demonstrate, in order to solve the pre-default optimization subproblem, we need the solution of the post-default one. Before proceeding further, we recall some functional spaces which will be needed for the following proofs. We set $C_P^{2,\alpha}$ as the set of functions locally in $C_P^{2,\alpha} \left((0, T) \times \tilde{\Delta}_{N-1} \right) \cap C \left([0, T] \times \tilde{\Delta}_{N-1} \right)$, where we recall that for a given domain D of \mathbb{R}^{N-1} and $\alpha \in (0, 1)$, the parabolic Hölder space $C_P^{2,\alpha}(D)$ is defined by the following norms

$$\|\psi\|_{C_P^{2,\alpha}(D)} = \|\psi\|_{C_P^\alpha(D)} + \|\partial_t \psi\|_{C_P^\alpha(D)} + \sum_{i=1}^{N-1} \|\partial_{\tilde{p}_i} \psi\|_{C_P^\alpha(D)} + \sum_{i,j=1}^{N-1} \|\partial_{\tilde{p}_i \tilde{p}_j} \psi\|_{C_P^\alpha(D)} \quad (5.1)$$

with

$$\|\psi\|_{C_P^\alpha(D)} := \sup_{(t,\tilde{p}) \in D} |\psi(t,\tilde{p})| + \sup_{\substack{(t,\tilde{p}), (t',\tilde{p}') \in D \\ (t,\tilde{p}) \neq (t',\tilde{p}')}} \frac{|\psi(t,\tilde{p}) - \psi(t',\tilde{p}')|}{(|\tilde{p} - \tilde{p}'|^2 + |t - t'|)^{\frac{\alpha}{2}}}.$$

Further, we denote

$$\|\psi\|_{C_P^{1,\alpha}(D)} = \|\psi\|_{C_P^\alpha(D)} + \sum_{i=1}^{N-1} \|\partial_{\tilde{p}_i} \psi\|_{C_P^\alpha(D)}. \quad (5.2)$$

5.1 Post Default Optimization Problem

Assume that default has already occurred; i.e., we are at a time t so that $\tau < t$. In particular, this means that $\pi_t^P = 0$. Let us denote by $\underline{w}(t, \tilde{p}) := w(t, \tilde{p}, 1)$ the value function in the post-default optimization problem. Then, we may rewrite Eq. (4.18) as follows:

$$0 = \underline{w}_t + \frac{1}{2} \text{tr}(\underline{\kappa} \underline{\kappa}^\top D^2 \underline{w}) + \frac{1}{2} (\nabla_{\tilde{p}} \underline{w}) \underline{\kappa} \underline{\kappa}^\top (\nabla_{\tilde{p}} \underline{w})^\top + \gamma r \\ + \sup_{\pi^S} \left[(\nabla_{\tilde{p}} \underline{w}) \underline{\beta}_\gamma - \gamma \pi^S (r - \tilde{\mu}(\tilde{p})) - \frac{\sigma^2}{2} \gamma (1 - \gamma) (\pi^S)^2 \right], \quad (5.3)$$

$$\underline{w}(T, \tilde{p}) = 0. \quad (5.4)$$

Here, $\underline{\kappa}(\tilde{p})$ is a $(N-1) \times 1$ vector determined by the first column of $\kappa(t, \tilde{p})$ (the second column of $\kappa(t, \tilde{p})$ consists of all zeros). Concretely,

$$\underline{\kappa}(\tilde{p}) := \left[\text{Diag}(\tilde{p}) \left(\vartheta^\top - \mathbf{1} \vartheta(\tilde{p}) \right) \right] \frac{1}{\sigma}, \quad (5.5)$$

where $\vartheta = (\mu_1 - \frac{1}{2}\sigma^2, \dots, \mu_{N-1} - \frac{1}{2}\sigma^2)$ is the first row of $\vartheta(t)'$ (the second row consists of all zeros) and, correspondingly, $\vartheta(\tilde{p}) = \tilde{\mu}(\tilde{p}) - \frac{1}{2}\sigma^2$ is a scalar with $\tilde{\mu}(\tilde{p}) = \mu_N + \sum_{i=1}^{N-1} (\mu_i - \mu_N) \tilde{p}^i$. Similarly, $\underline{\beta}_\gamma(t, \tilde{p}, \pi)$ in (5.3) is defined as

$$\underline{\beta}_\gamma(t, \tilde{p}, \pi) := \beta_\varpi(t, \tilde{p}) + \gamma \sigma \pi^S \underline{\kappa}(\tilde{p}),$$

where we recall that $\beta_\varpi(t, \tilde{p})$ has been defined in Eq. (4.3). It can easily be checked that the maximizer of Eq. (5.3) is given by

$$\pi^S(t, \tilde{p}) = \frac{1}{\sigma^2(1-\gamma)} \{ \tilde{\mu}(\tilde{p}) - r + \sigma (\nabla_{\tilde{p}} \underline{w}(t, \tilde{p})) \underline{\kappa}(\tilde{p}) \}. \quad (5.6)$$

Plugging the maximizer (5.6) in (5.3), we obtain

$$\underline{w}_t + \frac{1}{2} \text{tr}(\underline{\kappa} \underline{\kappa}^\top D^2 \underline{w}) + \frac{1}{2(1-\gamma)} (\nabla_{\tilde{p}} \underline{w}) \underline{\kappa} \underline{\kappa}^\top (\nabla_{\tilde{p}} \underline{w})^\top + (\nabla_{\tilde{p}} \underline{w}) \underline{\Phi} + \underline{\Psi} = 0, \quad (5.7)$$

$$\underline{w}(T, \tilde{p}) = 0, \quad (5.8)$$

where

$$\begin{aligned}\underline{\Phi}(t, \tilde{p}) &= \beta_{\omega}(t, \tilde{p}) + \frac{\gamma}{1-\gamma} \frac{\tilde{\mu}(\tilde{p}) - r}{\sigma} \underline{\kappa}(\tilde{p}), \\ \underline{\Psi}(\tilde{p}) &= \gamma r + \frac{\gamma}{2(1-\gamma)} \left(\frac{\tilde{\mu}(\tilde{p}) - r}{\sigma} \right)^2.\end{aligned}$$

Next we state, without proof, a useful result as a lemma.

Lemma 5.1 [Theorem 3.1 in Tamura and Watanabe (2011)] For any $T \geq 0$, there exists a classical solution \underline{w} that solves the Cauchy problem (5.7)-(5.8).

Remark 5.2 Tamura and Watanabe (2011) show the existence of a classical solution on the extension of the simplex to \mathbb{R}^{N-1} . Then, they prove that if $\tilde{p} \in \tilde{\Delta}_{N-1}$, the solution coincides with the solution to the Cauchy problem (5.7)-(5.8). It is well known from standard results, see Friedman (1964), Theorem 1, pag.92, that such a solution is $\mathcal{C}_P^{2,\alpha}$. We will use this fact in our subsequent proofs.

We then have

Theorem 5.3 The following assertions hold true:

- (1) The solution $\underline{w}(t, \tilde{p})$ coincides with the value function $w(t, \tilde{p}, 1)$ introduced in (4.7).
- (2) The optimal feedback control $\{\pi_s^S\}_{s \in [t, T]}$, denoted by $\tilde{\pi}_s^S$, can be written as $\tilde{\pi}_s^S = \tilde{\pi}^S(s, \tilde{p}_s^t)$ with

$$\tilde{\pi}^S(s, \tilde{p}) := \frac{1}{\sigma^2(1-\gamma)} (\tilde{\mu}(\tilde{p}) - r + \sigma \nabla_{\tilde{p}} \underline{w}(s, \tilde{p}) \underline{\kappa}(\tilde{p})). \quad (5.9)$$

Moreover, the feedback trading strategy $\tilde{\pi}_s := (\tilde{\pi}_s^S, \tilde{\pi}_s^P)^\top$, $\tilde{\pi}_s^P := 0$, is admissible; i.e., it satisfies the conditions of Definition 4.1.

The proof of Theorem 5.3 is reported in Appendix B. For now, let us mention a few useful remarks about the solution \underline{w} of (5.7)-(5.8). The existence of the solution \underline{w} follows from the Feynman-Kac formula as outlined in, e.g., the proof of Theorem 3.1 in Tamura and Watanabe (2011) (see also Nagai and Runggaldier (2008)). More specifically, the idea therein is to transform the problem into a linear PDE via the Cole-Hopf transformation:

$$\underline{\psi}(t, \tilde{p}) = e^{\frac{1}{1-\gamma} \underline{w}(t, \tilde{p})}. \quad (5.10)$$

Then, it follows that $\underline{w}(t, \tilde{p})$ solves Eq. (5.7)-(5.8) if and only if $\underline{\psi}(t, \tilde{p})$ solves the linear PDE

$$\begin{aligned}\frac{\partial \underline{\psi}}{\partial t} + \frac{1}{2} \text{tr}(\underline{\kappa} \underline{\kappa}^\top D^2 \underline{\psi}) + \underline{\Phi}^\top \nabla_{\tilde{p}} \underline{\psi} + \frac{\underline{\Psi}}{1-\gamma} \underline{\psi} &= 0, \\ \underline{\psi}(T, \tilde{p}) &= 1.\end{aligned} \quad (5.11)$$

5.2 Pre Default Optimization Problem

Assume that $\tau > t$, i.e. default has not occurred by time t . Let us denote by $\bar{w}(t, \tilde{p}) := w(t, \tilde{p}, 0)$ the value function corresponding to the pre-default optimization problem. Then, we may rewrite Eq. (4.18) as

$$\begin{aligned}\bar{w}_t + \frac{1}{2} \text{tr}(\bar{\kappa} \bar{\kappa}^\top D^2 \bar{w}) + \frac{1}{2} (\nabla_{\tilde{p}} \bar{w}) \bar{\kappa} \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top + \\ \sup_{\pi = (\pi^S, \pi^P)} \left\{ (\nabla_{\tilde{p}} \bar{w}) \bar{\beta}_\gamma - \gamma \pi^S (r - \tilde{\mu}(\tilde{p})) - \gamma \pi^P (r - \tilde{a}(t, \tilde{p})) - \frac{1}{2} \gamma (1-\gamma) \pi^\top \Sigma_Y^\top \Sigma_Y \pi \right\} \\ + \tilde{h}(\tilde{p}) \left[e^{\underline{w}\left(t, \frac{1}{\tilde{h}(\tilde{p})} \tilde{p} \cdot h'\right) - \bar{w}(t, \tilde{p})} - 1 \right] + \gamma r = 0, \\ \bar{w}(T, \tilde{p}) = 0.\end{aligned} \quad (5.12)$$

Above, $\bar{\kappa}(t, \tilde{p})$ is a $(N-1) \times 2$ matrix given by

$$\bar{\kappa}(t, \tilde{p}) := \kappa(t, \tilde{p}) = [\text{Diag}(\tilde{p}) (\bar{\vartheta}(t)^\top - \mathbf{1}\bar{\vartheta}(t, \tilde{p})^\top)] (\Sigma_Y \Sigma_Y^\top)^{-1} \Sigma_Y,$$

with $\bar{\vartheta}(t) = \vartheta(t)'$ being a $2 \times (N-1)$ matrix and $\tilde{\vartheta}(t, \tilde{p}) := \bar{\vartheta}(t, \tilde{p})$. Further,

$$\bar{\beta}_\gamma(t, \tilde{p}, \pi) := \beta_\gamma(t, \tilde{p}, \pi) = \beta_\omega(t, \tilde{p}) + \gamma \bar{\kappa}(t, \tilde{p}) \Sigma_Y^\top \pi.$$

It is important to point out the explicit appearance of the solution \underline{w} to the HJB post default equation in the PDE (5.12) satisfied by the pre-default HJB equation \bar{w} . This establishes the required relationship between pre and post-default optimization subproblems.

Next, define

$$\Upsilon(t, \tilde{p}) = (r - \tilde{\mu}(\tilde{p}), r - \tilde{a}(t, \tilde{p}))^\top.$$

Then, we can rewrite Eq. (5.12) as

$$\begin{aligned} & \bar{w}_t + \frac{1}{2} \text{tr}(\bar{\kappa} \bar{\kappa}^\top D^2 \bar{w}) + \frac{1}{2} (\nabla_{\tilde{p}} \bar{w}) \bar{\kappa} \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top + \\ & \sup_{\pi} \left\{ (\nabla_{\tilde{p}} \bar{w}) \bar{\beta}_\gamma - \gamma \pi^\top \Upsilon - \frac{1}{2} \gamma (1 - \gamma) \pi^\top \Sigma_Y^\top \Sigma_Y \pi \right\} + \tilde{h}(\tilde{p}) \left[e^{\frac{w}{\bar{\kappa}(\tilde{p})} \tilde{p} \cdot h'} - \bar{w}(t, \tilde{p}) - 1 \right] + \gamma r = 0 \quad (5.13) \\ & \bar{w}(T, \tilde{p}) = 0. \end{aligned}$$

Using the first order condition, we obtain that the maximal point π^* is the solution of the following equation:

$$\gamma \Sigma_Y \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top - \gamma \Upsilon - \gamma (1 - \gamma) \Sigma_Y^\top \Sigma_Y \pi^* = 0.$$

Solving the previous equation for π^* yields:

$$\pi^* = \frac{1}{1 - \gamma} (\Sigma_Y^\top \Sigma_Y)^{-1} (-\Upsilon + \Sigma_Y \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top). \quad (5.14)$$

After plugging π^* into (5.13), and performing algebraic simplifications (see Appendix B for details), we obtain

$$\bar{w}_t + \frac{1}{2} \text{tr}(\bar{\kappa} \bar{\kappa}^\top D^2 \bar{w}) + \frac{1}{2(1 - \gamma)} (\nabla_{\tilde{p}} \bar{w}) \bar{\kappa} \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top + (\nabla_{\tilde{p}} \bar{w}) \bar{\Phi} + \tilde{h}(\tilde{p}) e^{\frac{w}{\bar{\kappa}(\tilde{p})} \tilde{p} \cdot h'} e^{-\bar{w}(t, \tilde{p})} + \bar{\Psi} = 0, \quad (5.15)$$

$$\bar{w}(T, \tilde{p}) = 0, \quad (5.16)$$

where

$$\bar{\Phi}(t, \tilde{p}) = \beta_\omega(t, \tilde{p}) - \frac{\gamma}{1 - \gamma} \bar{\kappa} \Sigma_Y^{-1} \Upsilon(t, \tilde{p}),$$

$$\bar{\Psi}(t, \tilde{p}) = \frac{1}{2} \frac{\gamma}{1 - \gamma} \Upsilon^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Upsilon(t, \tilde{p}) + \gamma r - \tilde{h}(\tilde{p}).$$

Next, we prove the existence of a classical solution to the above Cauchy problem. We first perform a similar transformation as in the post-default case, and obtain that the function $\bar{w}(t, \tilde{p})$ solves the problem (5.15) if and only if the function

$$\bar{\psi}(t, \tilde{p}) = e^{\frac{1}{1 - \gamma} \bar{w}(t, \tilde{p})}, \quad (5.17)$$

solves the Cauchy problem

$$\begin{aligned} & \bar{\psi}_t(t, \tilde{p}) + \frac{1}{2} \text{tr}(\bar{\kappa} \bar{\kappa}^\top D^2 \bar{\psi}(t, \tilde{p})) + \bar{\Phi}(t, \tilde{p}) \nabla_{\tilde{p}} \bar{\psi}(t, \tilde{p}) + \bar{\Psi}(t, \tilde{p}) \frac{\bar{\psi}(t, \tilde{p})}{1 - \gamma} + \tilde{h}(\tilde{p}) e^{\frac{w}{\bar{\kappa}(\tilde{p})} \tilde{p} \cdot h'} \frac{\bar{\psi}^\gamma(t, \tilde{p})}{1 - \gamma} = 0 \\ & \bar{\psi}(T, \tilde{p}) = 1. \quad (5.18) \end{aligned}$$

Notice that problem (5.18) is *non-linear*. Hence, proving the existence of a classical solution is not as direct as in the post-default case where the transformed HJB-PDE given by Eq. (5.11) turned out to be linear. We establish this result by applying a monotone iterative method used in Di Francesco et al. (2007) for the study of obstacle problems for American options. There are, however, significant differences between the two problems, mainly arising from the appearance of the non-linear term $\bar{\psi}^\gamma$ in our PDE (5.18). This term is not globally Lipschitz continuous, while all PDE coefficients in Di Francesco et al. (2007) satisfy this condition. For this reason, it is crucial for us to prove that $\bar{\psi}$ is bounded away from zero, while for their obstacle problem Di Francesco et al. (2007) only need to show that the solution is bounded, i.e. $|\bar{\psi}(t, \tilde{p})| \leq ce^{ct}$ for some positive constant c . Below, we make use of the parabolic Hölder space $C_P^{2, \alpha}$ defined by the norm (5.1). We then have

Theorem 5.4 *Problem (5.18) admits a classical solution $\bar{\psi} \in \mathcal{C}_P^{2,\alpha}$ for any $\alpha \in (0,1)$. Moreover, there exists a constant $C \geq 1$, only depending on the L^∞ -norms of the coefficients of the PDE, such that, for each (t, \tilde{p}) ,*

$$\frac{1}{C} \leq \bar{\psi}(t, \tilde{p}) \leq e^{CT}. \quad (5.19)$$

The proof of Theorem 5.4 is reported in Appendix B. Let us remark that also

$$\bar{w}(t, \tilde{p}) \in \mathcal{C}_P^{2,\alpha} \quad (5.20)$$

in view of the relation (5.17) and the estimate (5.19), yielding that \bar{w} has the same properties of $\bar{\psi}$ in the previous theorem. The following result shows a verification theorem for the pre default optimization problem.

Theorem 5.5 *Suppose that the conditions of Theorem 5.3 are satisfied and, in particular, let $\underline{w} \in \mathcal{C}_P^{2,\alpha}$ be the solution of (5.7) with terminal condition (5.8). Additionally, let $\bar{w} \in \mathcal{C}_P^{2,\alpha}$ be the solution to the Cauchy problem (5.15)-(5.16) established in Theorem 5.4. Then, the following assertions hold true:*

- (1) *The solution $\bar{w}(t, \tilde{p})$ coincides with the optimal value function $w(t, \tilde{p}, 0)$ introduced in (4.7).*
- (2) *The optimal feedback controls $\{\pi_s^t\}_{s \in [t, T]} := \{(\pi_s^S, \pi_s^P)^\top\}_{s \in [t, T]}$, denoted by $\tilde{\pi} := (\tilde{\pi}^S, \tilde{\pi}^P)^\top$, can be written as $\tilde{\pi}_s^S = \tilde{\pi}^S(s, \tilde{p}_{s-}^t, H_{s-}^t)$ and $\tilde{\pi}_s^P = \tilde{\pi}^P(s, \tilde{p}_{s-}^t, H_{s-}^t)$ with*

$$(\tilde{\pi}^S(s, \tilde{p}, 0), \tilde{\pi}^P(s, \tilde{p}, 0))^\top := \frac{1}{1-\gamma} (\Sigma_Y^\top \Sigma_Y)^{-1} (-\Upsilon(s, \tilde{p}) + \Sigma_Y \bar{\kappa}^\top \nabla_{\tilde{p}} \bar{w}(s, \tilde{p})^\top), \quad (5.21)$$

$$(\tilde{\pi}^S(s, \tilde{p}, 1), \tilde{\pi}^P(s, \tilde{p}, 1))^\top := \left(\frac{1}{\sigma^2(1-\gamma)} (\tilde{\mu}(\tilde{p}) - r + \sigma \bar{\kappa}^\top \nabla_{\tilde{p}} \underline{w}(s, \tilde{p})^\top), 0 \right)^\top. \quad (5.22)$$

The proof of Theorem 5.5 is reported in Appendix B.

6 Numerical Analysis

We provide a numerical analysis of the optimal strategies and value functions derived in the previous sections. We set $N = 2$, i.e. we consider two regimes, thus the vector $\tilde{p} := p$ becomes one dimensional with p denoting the filter probability that the Markov chain is in regime “1”. Unless otherwise specified, throughout the section we use the following benchmark parameters: $\sigma = 0.4$, $v = 0.5$, $r = 0.03$, $\mu_1 = 0.5$, $\mu_2 = 0.2$, $h_1 = 1$, $h_2 = 0.2$, $\varpi_{11} = -0.5$, and $\varpi_{22} = -1$. We fix the time horizon to $T = 10$. We set $\gamma = 0.5$, i.e. we consider a square root investor. We describe the numerical setup in Section 6.1, and give an economic analysis of the strategies in Section 6.2.

6.1 Setup

Since the solution to the pre-default HJB-PDE depends on the solution to the post-default HJB PDE, we first need to solve Eq. (5.11). In case of two regimes, the PDE (5.11) becomes two dimensional with $t \in \mathbb{R}^+$ and $0 \leq p \leq 1$. More specifically, Eq. (5.11) reduces to

$$\begin{aligned} \frac{\partial \underline{\psi}(t, p)}{\partial t} + \frac{1}{2} \underline{\kappa}(t, p)^2 \frac{\partial^2 \underline{\psi}(t, p)}{\partial p^2} + \underline{\Phi}(t, p) \frac{\partial \underline{\psi}(t, p)}{\partial p} + \underline{\Psi}(t, p) \frac{\underline{\psi}(t, p)}{1-\gamma} &= 0, \\ \underline{\psi}(T, p) &= 1, \end{aligned}$$

where

$$\begin{aligned} \underline{\kappa}(t, p) &= \sigma^{-1} p (\mu_1 - (\mu_1 p + \mu_2 (1-p))) \\ \beta_\varpi(t, p) &= \varpi_{11} p + \varpi_{21} (1-p) = \varpi_{11} p - \varpi_{22} (1-p) \\ \underline{\Phi}(t, p) &= \beta_\varpi(t, p) + \frac{\gamma}{1-\gamma} \frac{\mu_1 p + (1-p)\mu_2 - r}{\sigma} \underline{\kappa}(t, p) \\ \underline{\Psi}(t, p) &= \gamma r + \frac{\gamma}{2(1-\gamma)} \left(\frac{\mu_1 p + (1-p)\mu_2 - r}{\sigma} \right)^2 \end{aligned} \quad (6.1)$$

We numerically solve the above derived PDE using a standard Crank-Nicolson method. From the transformation (5.10), we obtain that the post-default value function is given by $\underline{w}(t, p) = (1 - \gamma) \log(\underline{\psi}(t, p))$. The latter is then used into the pre-default PDE, computed as described next. On (t, p) , with $t \in \mathbb{R}^+$, and $0 \leq p \leq 1$, the PDE (5.18) satisfied by $\bar{\psi}(t, p)$ reduces to

$$\frac{\partial \bar{\psi}(t, p)}{\partial t} + \frac{1}{2} \bar{\kappa}(t, p) \bar{\kappa}(t, p)^\top \frac{\partial^2 \bar{\psi}(t, p)}{\partial p^2} + \bar{\Phi}(t, p) \frac{\partial \bar{\psi}}{\partial p} + \bar{\Psi}(t, p) \frac{\bar{\psi}(t, p)}{1 - \gamma} + (h_2 + (h_1 - h_2)p) e^{\underline{w}\left(t, \frac{ph_1}{h_2 + (h_1 - h_2)p}\right)} \frac{\bar{\psi}(t, p)^\gamma}{1 - \gamma} = 0,$$

$$\bar{\psi}(T, p) = 1,$$

where

$$\bar{\kappa}(t, p) = p(1 - p) \left[\frac{\mu_1 - \mu_2}{\sigma}, \frac{a(t, e_1) - a(t, e_2)}{v} \right] =: [\bar{\kappa}_{11}(t, p), \bar{\kappa}_{12}(t, p)]$$

$$\bar{\Phi}(t, p) = \varpi_{21} + (\varpi_{11} - \varpi_{21})p + \frac{\gamma}{1 - \gamma} \left(\frac{\bar{\kappa}_{11}(t, p)}{\sigma} (\tilde{\mu}(p) - r) + \frac{\bar{\kappa}_{12}(t, p)}{v} (\tilde{a}(t, p) - r) \right)$$

$$\bar{\Psi}(t, p) = \frac{1}{2} \frac{\gamma}{1 - \gamma} \left(\frac{(\tilde{\mu}(p) - r)^2}{\sigma^2} + \frac{(\tilde{a}(t, p) - r)^2}{v^2} \right) + \gamma r - h_2 + (h_1 - h_2)p.$$

In the following analysis, we set $a(t, e_1) = -(r + h_1)$, and $a(t, e_2) = -(r + h_2)$. In agreement with the notation in Section 2, $a(t, e_1)$ and $a(t, e_2)$ denote the risk adjusted returns of the defaultable security when the current regime is “1”, respectively “2”. Similarly to the post-default case, we employ a standard Crank-Nicolson method to solve the above derived nonlinear PDE. The solution to the pre-default PDE is then obtained as $\underline{w}(t, p) = (1 - \gamma) \log(\underline{\psi}(t, p))$.

6.2 Analysis of Strategies

Figure 6.1 shows that the stock investment strategy is increasing in the filter probability of the hidden chain being in the first regime. This happens because under our parameter choices, the growth rate of the stock is higher in regime “1”, while the volatility stays unchanged. Consequently, if the filter estimate of being in the most profitable regime gets higher, the risk averse investor would prefer to shift a larger amount of wealth in the stock. On the other hand, as the probability of being in regime “1” increases, the risk averse investor shorts a higher amount of the defaultable security. This happens because the default intensity in regime “1” is the highest, and thus the risk averse investor wishing to decrease his exposure to default risk goes short in the vulnerable security. Notice also the key role played by the stock volatility σ . As the volatility gets lower, the investor shorts more units of the vulnerable security and invest the resulting proceeds in the stock security. Indeed, since the stock volatility is low while the default risk unchanged, the risk averse investor prefers to invest larger fraction of wealth in the stock security, and does so by raising cash via short-selling of the vulnerable security. The latter action also results in him having reduced exposure to credit risk. Since all model parameters depend on time only through the underlying hidden regime, investment strategies are not very sensitive to passage of time.

From figure 6.2, we can also see that both the pre-default and post-default value functions are decreasing in time, and increasing in the filter probability p . Moreover, as the filter probability of being in regime “1” increases, the investor extracts more utility given that he realizes higher gains by simultaneously shorting the vulnerable security and purchasing the stock security.

The investor increases the amount of shorted units of the vulnerable security if the filter probability of staying in the regime characterized by high default intensity gets larger. By contrast, if the filter probability of staying in the regime with low default intensity increases he takes higher credit risk exposure by purchasing more bond units. This is clearly illustrated in the left panel of figure 6.3. Moreover, when $h_1 < h_2$ smaller amount of units of the vulnerable security are shorted if the probability p of staying in the low risk regime gets higher. However, when the default intensity h_1 in regime “1” exceeds the default intensity h_2 in regime “2” ($h_2 = 0.2$), the opposite effect is observed. As h_1 gets higher, a larger amount of units of the vulnerable security are shorted if the filter probability of staying in regime “1” increases. We also notice that, ceteris paribus, the pre-default value function is increasing in h_1 for a fixed p value, and also, for a given value of h_1 the pre-default value function increases in p .

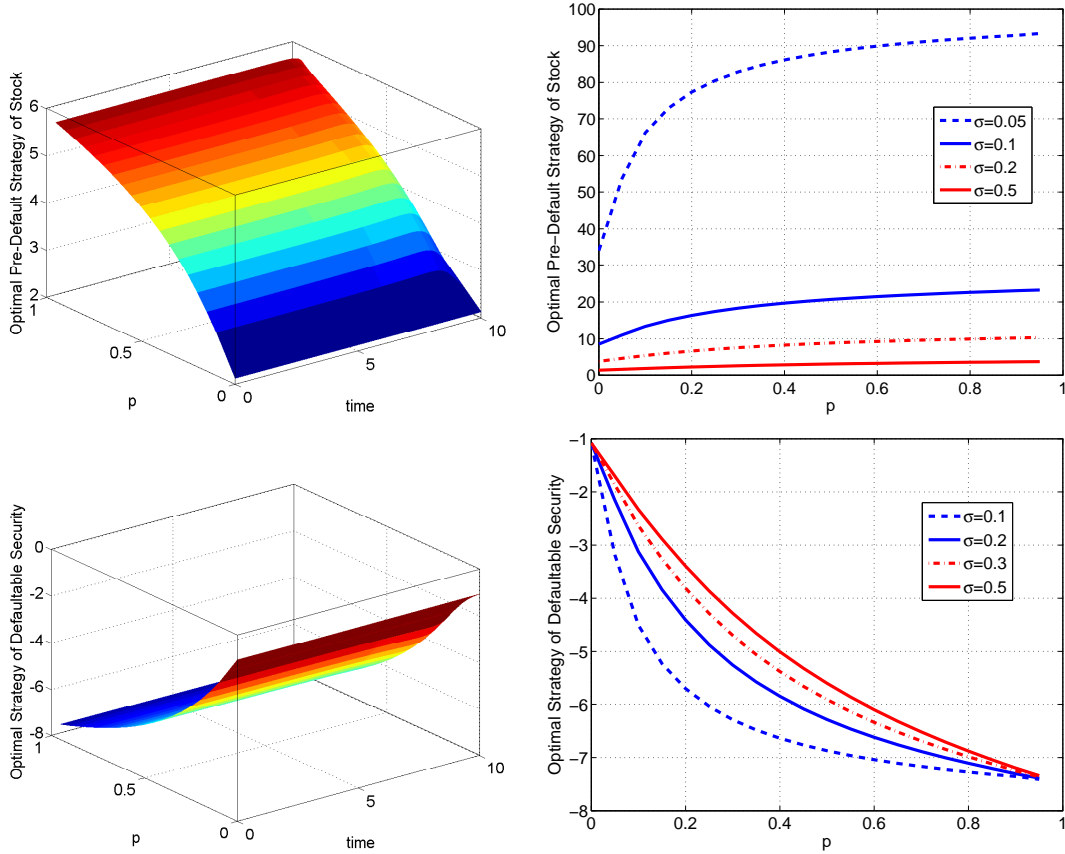


Fig. 6.1 The top panels report the stock investment strategy. The bottom panels report the investment strategy in the vulnerable security. In the right panels, we set $t = 0$.

We next analyze the dependence of the optimal stock and vulnerable security investment strategies on the volatilities σ and v . Figure 6.4 shows that when the stock volatility is low, the investor puts a large fraction of wealth in the stock security. This happens because the filter gain coming from receiving stock observations is the highest. Since observations are more informative, the risk averse investor deposits a larger fraction of wealth in the stock, especially if the filter probability of being in the high growth regime (regime “1” with $\mu_1 = 0.5$) is high. As the volatility gets larger, stock price observations become less informative leading the investor to decrease the wealth fraction invested in stock. When the volatility exceeds a certain threshold, regardless of the filter probability the investor always puts a small amount of wealth in the stock. We also notice that a similar role is played by the volatility v of the defaultable security. From the right panel of figure 6.4, we notice that when v is low, i.e. price observations of the vulnerable security are very informative, the investor wants to reduce more his exposure to default risk. Hence, he shorts more units of the vulnerable security especially if the filter probability of being in the highest credit risk regime (regime “1” with $h_1 = 0.05$) is large. This reflects the risk averse nature of the investor who dislikes default risk and uncertainty. As for the stock, when v gets larger the investment strategy in the defaultable security becomes less sensitive to the filter probability and for large values of v the investor may even find it optimal to purchase the defaultable security. This happens when the potential loss incurred by the investor when he is long credit and default occurs (hence making the vulnerable security worthless) is outweighed by the risk adjusted return resulting from holding the defaultable security.

We conclude by relating partial to full information settings. As price volatilities become smaller, the regime switching model becomes more observable. This is because price observations become more informative and allow the investor to build more accurate estimates of the regime in place. Consequently, the above analysis outlines the important role played by regime uncertainty in determining the optimal strategies of risk averse investors. Compared to the case of fully observed regimes studied in Capponi and Figueroa-López (2014), the presence of incomplete information induces the risk averse investor to decrease the wealth amount invested in the risky securities. As

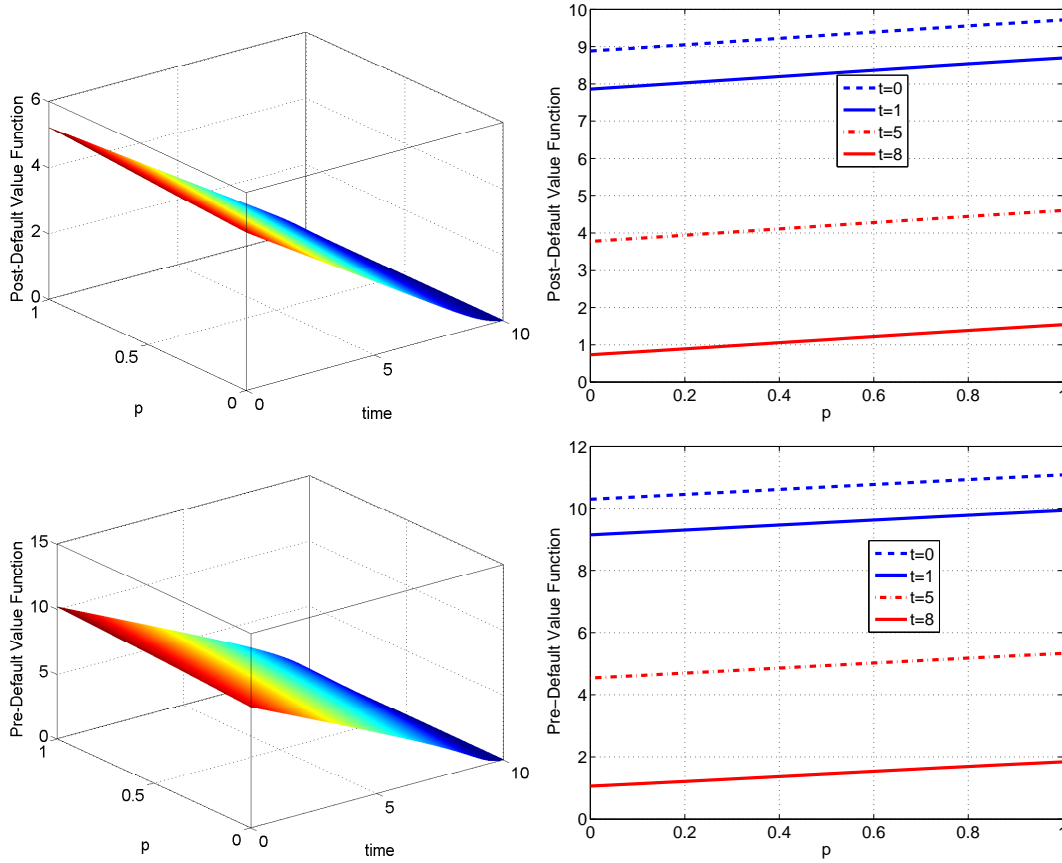


Fig. 6.2 The pre-default and post-default value functions plotted versus time t and filter probability p .

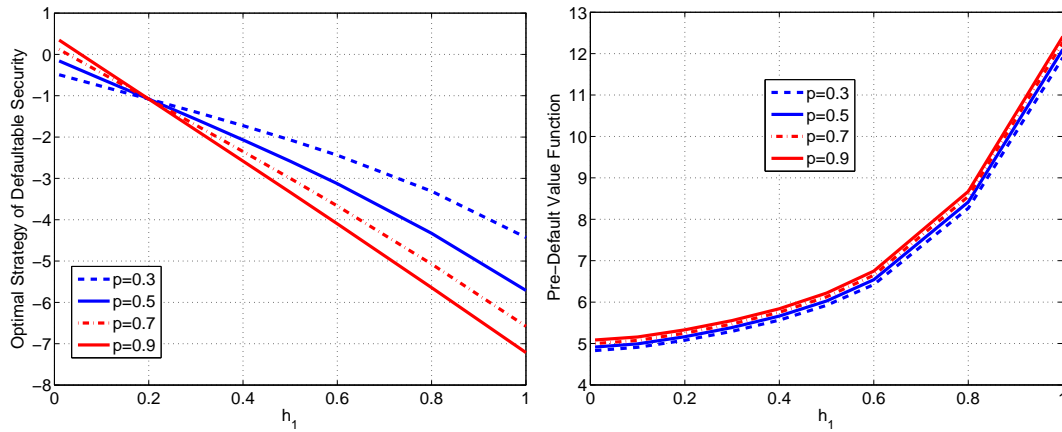


Fig. 6.3 The left panel reports the optimal investment strategy in the defaultable security. The right panel reports the pre-default value function. We set $t = 0$.

clearly illustrated in figure 6.4, when the price volatilities are sufficiently high ($\sigma \approx 0.8$ for the stock and $v \approx 0.6$ for the defaultable security), the investor deposits almost entire amount of wealth in the money market account.

7 Conclusions

We have studied the optimal investment problem of a power investor in an economy consisting of a defaultable security, a stock, and a money market account. The price processes of these securities are assumed to have drift coefficients and default intensities modulated by a hidden Markov chain. We have reduced the partially observed

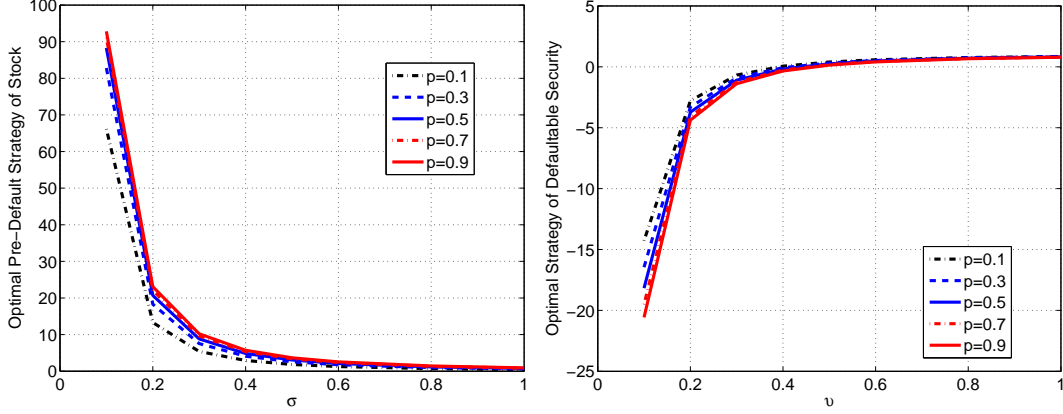


Fig. 6.4 The left panel reports the dependence of the optimal stock investment strategy on σ . The right panel reports the dependence of the optimal investment strategy in the vulnerable security on v . We set $t = 0$. In the right panel, we set $h_1 = 0.05$ and $h_2 = 0.01$.

stochastic control problem to a risk sensitive one, where the state is given by the filtered regime probabilities. The conditioning filtration, generated by the prices of the stock and of the defaultable security, and by the indicator of default occurrence, is driven both by a Brownian component and by a pure jump martingale. The filter has been used to derive the HJB partial differential equation corresponding to the risk sensitive control problem. We have split the latter into a pre-default and a post-default dynamic programming subproblem. The HJB PDE corresponding to the post-default value function can be transformed to a linear parabolic PDE, for which existence and uniqueness of a classical solution can be guaranteed. By contrast, the HJB PDE corresponding to the pre-default value function has exponential nonlinearity and quadratic gradient growth. We have provided a detailed mathematical analysis of such PDE and established the existence of a classical solution with $\mathcal{C}_P^{2,\alpha}$ regularity. We have then proven verification theorems establishing the correspondence between the PDE solutions and the value functions of the control problem. Our study has been complemented with a thorough numerical analysis illustrating the role of regime uncertainty, default risk, and price volatilities on the optimal allocation decisions and value functions.

Acknowledgments

The authors gratefully acknowledge two anonymous reviewers and Wolfgang Runggaldier for providing constructive and insightful comments which improved significantly the quality of the manuscript. Agostino Capponi would also like to thank Ramon Van Handel for very useful discussions and insights provided in the original model setup.

A Proofs related to Section 3

Lemma A.1 *Let*

$$q_t^i = \mathbb{E}^{\hat{\mathbb{P}}} \left[L_t \mathbf{1}_{\{X_t = e_i\}} \middle| \mathcal{G}_t^I \right]. \quad (\text{A.1})$$

Then, the dynamics of $(q_t^i)_{t \geq 0}$, $i = 1, \dots, N$, under the measure $\hat{\mathbb{P}}$, is given by the following system of stochastic differential equations (SDE):

$$\begin{aligned} dq_t^i &= \sum_{\ell=1}^N \varpi_{\ell,i}(t) q_t^\ell dt + q_t^i Q^\top(t, e_i, \pi_t) \Sigma_Y d\hat{W}_t + q_{t-}^i (h_i - 1) d\hat{\xi}_t - \gamma \eta(t, e_i, \pi_t) q_t^i dt, \\ q_0^i &= p_0^i. \end{aligned} \quad (\text{A.2})$$

Proof Let us introduce the following notation

$$H_t^i := \mathbf{1}_{\{X_t = e_i\}}.$$

Note that $X_t = (H_t^1, \dots, H_t^N)^\top$ and, from (2.1),

$$H_t^i = H_0^i + \int_0^t \sum_{\ell=1}^N \varpi_{\ell,i}(s) H_s^\ell ds + \varphi_i(t). \quad (\text{A.3})$$

From Eq. (3.16) and (3.10), we deduce that, under $\hat{\mathbb{P}}$,

$$dL_t = L_{t-}(h_t - 1)d\hat{\xi}_t + L_t Q^\top(t, X_t, \pi_t) dY_t - L_t \gamma \eta(t, X_t, \pi_t) dt$$

which yields that

$$[L, H^i]_t = \int_0^t L_{s-} Q^\top(s, X_s, \pi_s) d[Y, H^i]_s + \int_0^t L_{s-}(h_s - 1)d[\hat{\xi}, H^i]_s.$$

As $(Y_t)_{t \geq 0}$ and $(H_t)_{t \geq 0}$ are independent of $(X_t)_{t \geq 0}$ (and, hence, of H^i), under $\hat{\mathbb{P}}$ it holds (see also Wong and Hajek (1985)) that, $\hat{\mathbb{P}}$ almost surely,

$$[Y, H^i]_s = [\hat{\xi}, H^i]_s = 0, \text{ for all } s \geq 0.$$

Thus, applying Itô's formula, we obtain

$$\begin{aligned} L_t H_t^i &= H_0^i + \int_0^t H_{s-}^i dL_s + \int_0^t L_{s-} dH_s^i \\ &= H_0^i + \int_0^t H_{s-}^i L_s Q^\top(s, X_s, \pi_s) dY_s + \int_0^t H_{s-}^i L_{s-}(h_{s-} - 1)d\hat{\xi}_s \\ &\quad - \int_0^t H_{s-}^i L_s \gamma \eta(s, X_s, \pi_s) ds + \int_0^t L_s \sum_{\ell=1}^N \varpi_{\ell, i}(s) H_s^\ell ds + \int_0^t L_{s-} d\varphi_i(s) \end{aligned} \quad (\text{A.4})$$

Since $(\varphi_i(t))_{t \geq 0}$ is a $((\mathcal{F}_t^X)_{t \geq 0}, \hat{\mathbb{P}})$ -martingale, and \mathcal{G}_T^I is independent of \mathcal{F}_T^X under $\hat{\mathbb{P}}$, we have that $\mathbb{E}^{\hat{\mathbb{P}}} \left[\int_0^t L_{s-} d\varphi_i(s) | \mathcal{G}_t^I \right] = 0$. Therefore, taking \mathcal{G}_t^I conditional expectations in Eq. (A.4), we obtain

$$\begin{aligned} \mathbb{E}^{\hat{\mathbb{P}}} \left[L_t H_t^i | \mathcal{G}_t^I \right] &= 1 + \int_0^t \mathbb{E}^{\hat{\mathbb{P}}} \left[L_s H_s^i Q^\top(s, e_i, \pi_s) | \mathcal{G}_s^I \right] dY_s + \int_0^t \mathbb{E}^{\hat{\mathbb{P}}} \left[L_{s-} H_{s-}^i (h_{s-} - 1) | \mathcal{G}_s^I \right] d\hat{\xi}_s \\ &\quad - \int_0^t \mathbb{E}^{\hat{\mathbb{P}}} \left[H_s^i L_s \gamma \eta(s, e_i, \pi_s) | \mathcal{G}_s^I \right] ds + \int_0^t \mathbb{E}^{\hat{\mathbb{P}}} \left[\sum_{\ell=1}^N \varpi_{\ell, i}(s) L_s H_s^\ell | \mathcal{G}_s^I \right] ds, \end{aligned} \quad (\text{A.5})$$

where we have used that, if φ_i is \mathbb{G} -predictable then (see, for instance, Wong and Hajek (1985), Ch. 7, Lemma 3.2)

$$\begin{aligned} \mathbb{E}^{\hat{\mathbb{P}}} \left[\int_0^t \varphi_s L_{s-} dY_s | \mathcal{G}_t^I \right] &= \int_0^t \mathbb{E}^{\hat{\mathbb{P}}} \left[\varphi_s L_{s-} | \mathcal{G}_s^I \right] dY_s \\ \mathbb{E}^{\hat{\mathbb{P}}} \left[\int_0^t \varphi_s L_{s-} d\hat{\xi}_s | \mathcal{G}_t^I \right] &= \int_0^t \mathbb{E}^{\hat{\mathbb{P}}} \left[\varphi_s L_{s-} | \mathcal{G}_s^I \right] d\hat{\xi}_s \\ \mathbb{E}^{\hat{\mathbb{P}}} \left[\int_0^t \varphi_s L_{s-} ds | \mathcal{G}_t^I \right] &= \int_0^t \mathbb{E}^{\hat{\mathbb{P}}} \left[\varphi_s L_{s-} | \mathcal{G}_s^I \right] ds \end{aligned}$$

Observing that under $\hat{\mathbb{P}}$ $dY_t = \Sigma_Y d\hat{W}_t$, using that $Q(t, e_i, \pi_t)$ and $\eta(t, e_i, \pi_t)$ are $(\mathcal{G}_t^I)_{t \geq 0}$ adapted, and that the Markov-chain generator $A(t)$ is deterministic, we obtain Eq. (A.2) upon taking the differential of Eq. (A.5).

Lemma A.2 *The following identities hold*

$$q_t^i = \hat{L}_t p_t^i \quad (\text{A.6})$$

$$p_t^i = \frac{q_t^i}{\sum_{j=1}^N q_t^j} \quad (\text{A.7})$$

where q_t^i , \hat{L}_t , and p_t^i are defined, respectively, by (A.1), (3.23), and (3.20).

Proof We start establishing the relation (A.6) by comparing the dynamics of q_t^i and of $\hat{L}_t p_t^i$. The dynamics of q_t^i is known from Lemma A.1 and given in Eq. (A.2). Next, we derive the dynamics of $\hat{L}_t p_t^i$. We have

$$d(\hat{L}_t p_t^i) = \hat{L}_{t-} dp_t^i + p_{t-}^i d\hat{L}_t + d \left[\hat{L}, p^i \right]_t.$$

From Eq. (3.24) and (3.22), we obtain

$$\begin{aligned} d \left[\hat{L}, p^i \right]_t &= p_t^i \hat{L}_t \hat{\vartheta}^\top(t, p_t) \Sigma_Y \Sigma_Y^{-1} \left(\vartheta(t, e_i) - \hat{\vartheta}(t, p_t) \right) dt + p_t^i \hat{L}_t \gamma \pi_t^\top \left(\vartheta(t, e_i) - \hat{\vartheta}(t, p_t) \right) dt \\ &\quad + \left(\hat{h}_{t-} - 1 \right) \frac{h_i - \hat{h}_{t-}}{\hat{h}_{t-}} \hat{L}_{t-} p_{t-}^i dH_t. \end{aligned} \quad (\text{A.8})$$

Using the above equations, along with (3.22), we obtain

$$\begin{aligned}
d(\hat{L}_t p_t^i) &= \hat{L}_t \left(\sum_{\ell=1}^N \varpi_{\ell,i}(t) p_t^\ell dt \right) + \hat{L}_t p_t^i \left(\vartheta^\top(t, e_i) - \hat{\vartheta}^\top(t, p_t) \right) \left(\Sigma_Y \Sigma_Y^\top \right)^{-1} \left(dY_t - \hat{\vartheta}(t, p_t) dt \right) \\
&+ \hat{L}_t p_t^i \frac{h_i - \hat{h}_t^-}{\hat{h}_t^-} \left(dH_t - \hat{h}_t^- \bar{H}_t^- dt \right) + p_t^i \hat{L}_t \hat{Q}^\top(t, p_t, \pi_t) dY_t - p_t^i \hat{L}_t \gamma \hat{\eta}(t, p_t, \pi_t) dt \\
&+ p_t^i \hat{L}_t^- (\hat{h}_t^- - 1) (dH_t - \bar{H}_t^- dt) + (\hat{h}_t^- - 1) \frac{h_i - \hat{h}_t^-}{\hat{h}_t^-} \hat{L}_t^- p_t^i dH_t \\
&+ p_t^i \hat{L}_t \hat{\vartheta}^\top(t, p_t) \left(\Sigma_Y \Sigma_Y^\top \right)^{-1} \left(\vartheta(t, e_i) - \hat{\vartheta}(t, p_t) \right) dt + \gamma p_t^i \hat{L}_t \pi_t^\top \left(\vartheta(t, e_i) - \hat{\vartheta}(t, p_t) \right) dt.
\end{aligned} \tag{A.9}$$

Next, observe that

$$\begin{aligned}
\hat{L}_t p_t^i \left(\vartheta^\top(t, e_i) - \hat{\vartheta}^\top(t, p_t) \right) \left(\Sigma_Y \Sigma_Y^\top \right)^{-1} \left(dY_t - \hat{\vartheta}(t, p_t) dt \right) + p_t^i \hat{L}_t \hat{Q}^\top(t, p_t, \pi_t) dY_t = \\
\hat{L}_t p_t^i \hat{Q}^\top(t, e_i, \pi_t) dY_t - \hat{L}_t p_t^i \left(\vartheta^\top(t, e_i) - \hat{\vartheta}^\top(t, p_t) \right) \left(\Sigma_Y \Sigma_Y^\top \right)^{-1} \hat{\vartheta}(t, p_t) dt
\end{aligned} \tag{A.10}$$

Moreover,

$$\eta(t, e_i, \pi_t) - \hat{\eta}(t, p_t, \pi_t) = \pi_t^\top \left(\hat{\vartheta}(t, p_t) - \vartheta(t, e_i) \right). \tag{A.11}$$

Using relations (A.10), and (A.11), along with straightforward simplifications, we may simplify Eq. (A.9) to

$$d(\hat{L}_t p_t^i) = \left(\sum_{\ell=1}^N \varpi_{\ell,i}(t) \hat{L}_t p_t^\ell dt \right) + \hat{L}_t p_t^i \hat{Q}^\top(t, e_i, \pi_t) dY_t - \gamma \hat{L}_t p_t^i \eta(t, e_i, \pi_t) dt + \hat{L}_t^- p_t^i (h_i - 1) d\hat{\xi}_t. \tag{A.12}$$

Using that $dY_t = \Sigma_Y d\hat{W}_t$, we have that the equality (A.6) holds via a direct comparison of equations (A.12) and (A.2).

Next, we establish (A.7). Using Eq. (A.6) and that $\sum_{i=1}^N p_t^i = 1$, we deduce that

$$d \left(\sum_{i=1}^N q_t^i \right) = d \left(\sum_{i=1}^N \hat{L}_t p_t^i \right) = d\hat{L}_t$$

hence obtaining that $\sum_{i=1}^N q_t^i = \hat{L}_t$. Using again (A.6), this gives

$$p_t^i = \frac{q_t^i}{\hat{L}_t} = \frac{q_t^i}{\sum_{j=1}^N q_t^j}$$

This completes the proof.

Proof of Proposition 3.3.

Proof

Using Eq. (3.15), (A.1), and the relation (A.6) established in Lemma A.2, we have that

$$\begin{aligned}
J(v, \pi, T) &= \frac{1}{\gamma} \mathbb{E}^\mathbb{P} [V_T^\gamma] = \frac{v^\gamma}{\gamma} \mathbb{E}^{\hat{\mathbb{P}}} [L_T] = \frac{v^\gamma}{\gamma} \mathbb{E}^{\hat{\mathbb{P}}} \left[\mathbb{E}^{\hat{\mathbb{P}}} [L_T | \mathcal{G}_T^I] \right] \\
&= \frac{v^\gamma}{\gamma} \sum_{i=1}^N \mathbb{E}^{\hat{\mathbb{P}}} \left[\mathbb{E}^{\hat{\mathbb{P}}} [L_T \mathbf{1}_{\{X_T = e_i\}} | \mathcal{G}_T^I] \right] = \frac{v^\gamma}{\gamma} \sum_{i=1}^N \mathbb{E}^{\hat{\mathbb{P}}} [q_T^i] \\
&= \frac{v^\gamma}{\gamma} \sum_{i=1}^N \mathbb{E}^{\hat{\mathbb{P}}} [\hat{L}_T p_T^i] = \frac{v^\gamma}{\gamma} \mathbb{E}^{\hat{\mathbb{P}}} [\hat{L}_T],
\end{aligned}$$

thus proving the statement.

B Proofs related to Section 5

We start with a Lemma, which will be needed in the section where the verification theorem is proven.

Lemma B.1 *For any $T > 0$ and $i \in \{1, \dots, N\}$, it holds that*

- (1) $\mathbb{P}(p_t^i > 0, \text{ for all } t \in [0, T]) = 1.$
- (2) $\mathbb{P}(p_t^i < 1, \text{ for all } t \in [0, T]) = 1.$

Proof Define $\varsigma = \inf\{t : p_t^i = 0\} \wedge T$. If p_t^i can hit zero, then $\mathbb{P}(p_\varsigma^i = 0) > 0$. Recall that $p_t^i = \frac{q_t^i}{\sum_j q_t^j}$ from Eq. (A.7), hence $p_\varsigma^i = \frac{q_\varsigma^i}{\sum_j q_\varsigma^j}$, where the equality

$$q_\varsigma^i = \mathbb{E}^{\tilde{\mathbb{P}}} \left[L_\varsigma \mathbf{1}_{X_\varsigma = e_i} | \mathcal{G}_\varsigma^I \right]$$

is true by the optional projection property, see Rogers and Williams (2006). Define the two dimensional (observed) log-price process $Y_t = (\log(S_t), \log(P_t))^\top$. As $q_\varsigma^i = \mathbb{E}^{\tilde{\mathbb{P}}} [L_\varsigma \mathbf{1}_{X_\varsigma = e_i} | \mathcal{G}_\varsigma^I]$, and using that $L_\varsigma > 0$, we can choose a modification $g(Y, H, X_\varsigma)$ of $\mathbb{E}^{\tilde{\mathbb{P}}} [L_\varsigma | \mathcal{G}_\varsigma^I, X_\varsigma]$ such that $g > 0$, and, for each e_i , $g(Y, H, e_i)$ is \mathcal{G}_ς^I -measurable. By the tower property

$$q_\varsigma^i = \mathbb{E}^{\tilde{\mathbb{P}}} [g(Y, H, X_\varsigma) \mathbf{1}_{X_\varsigma = e_i} | \mathcal{G}_\varsigma^I] = g(Y, H, e_i) \tilde{\mathbb{P}}(X_\varsigma = e_i | \mathcal{G}_\varsigma^I) = g(Y, H, e_i) \mathbb{P}(X_t = e_i) \Big|_{t=\varsigma},$$

where the first equality follows because ς is \mathcal{G}_ς^I -measurable and the last two equalities because X is independent of \mathcal{G}^I under $\tilde{\mathbb{P}}$. As $\mathbb{P}(X_t = e_i) > 0$ and $g > 0$, we get that $q_\varsigma^i > 0$ a.s, which contradicts that $\mathbb{P}(p_\varsigma^i = 0) > 0$. This proves the first statement in the Lemma. Next, we notice that

$$\mathbb{P}(p_t^i = 0, \text{ for some } t \in [0, T]) = 1 - \mathbb{P}(p_t^i > 0, \text{ for all } t \in [0, T]) = 0,$$

where the last equality follows from the first statement. This immediately yields the second statement.

Proof of Eq. (5.15)

Proof Let us first analyze the first term in the sup of Eq. (5.13), i.e. $\beta_\gamma^\top \nabla_{\tilde{p}} w$. For brevity, we use $\beta_\omega := \beta_\omega(t, \tilde{p}, 0)$. By definition of β_γ , and using the maximizer $\pi_t := \pi_t^*$ in (5.14), we have

$$\begin{aligned} \beta_\gamma^\top &= \beta_\omega^\top + \gamma \pi^\top \Sigma_Y \bar{\kappa}^\top = \beta_\omega^\top + \frac{\gamma}{1-\gamma} \left(\Sigma_Y \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top - \Upsilon \right)^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Sigma_Y \bar{\kappa}^\top \\ &= \beta_\omega^\top + \frac{\gamma}{1-\gamma} (\nabla_{\tilde{p}} \bar{w}) \bar{\kappa} \Sigma_Y^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Sigma_Y \bar{\kappa}^\top - \frac{\gamma}{1-\gamma} \Upsilon^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Sigma_Y \bar{\kappa}^\top \end{aligned} \quad (\text{B.1})$$

Further, again using the expression for $\pi = \pi^*$, the second term in the sup is equal to

$$\begin{aligned} -\gamma \pi^\top \Upsilon &= -\frac{\gamma}{1-\gamma} \left(-\Upsilon + \Sigma_Y \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top \right)^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Upsilon \\ &= \frac{\gamma}{1-\gamma} \Upsilon^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Upsilon - \frac{\gamma}{1-\gamma} (\nabla_{\tilde{p}} \bar{w}) \bar{\kappa} \Sigma_Y^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Upsilon. \end{aligned} \quad (\text{B.2})$$

The third term in the sup may be simplified as

$$\begin{aligned} &-\frac{1}{2} \frac{\gamma}{1-\gamma} \left(-\Upsilon + \Sigma_Y \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top \right)^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \left(-\Upsilon + \Sigma_Y \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top \right) \\ &= -\frac{1}{2} \frac{\gamma}{1-\gamma} \Upsilon^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Upsilon + \frac{1}{2} \frac{\gamma}{1-\gamma} \Upsilon^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Sigma_Y \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top \\ &+ \frac{1}{2} \frac{\gamma}{1-\gamma} (\nabla_{\tilde{p}} \bar{w}) \bar{\kappa} \Sigma_Y^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Upsilon - \frac{1}{2} \frac{\gamma}{1-\gamma} (\nabla_{\tilde{p}} \bar{w}) \bar{\kappa} \Sigma_Y^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Sigma_Y \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top. \end{aligned} \quad (\text{B.3})$$

Using Eq. (B.1), (B.2), and (B.3), we obtain that

$$\begin{aligned} &\sup_\pi \left\{ \beta_\gamma^\top (\nabla_{\tilde{p}} \bar{w})^\top - \gamma \pi_t^\top \Upsilon - \frac{1}{2} \gamma (1-\gamma) \pi_t^\top \Sigma_Y^\top \Sigma_Y \pi_t \right\} \\ &= \beta_\omega^\top (\nabla_{\tilde{p}} \bar{w})^\top + \frac{1}{2} \frac{\gamma}{1-\gamma} (\nabla_{\tilde{p}} \bar{w}) \bar{\kappa} \Sigma_Y^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Sigma_Y \bar{\kappa}^\top (\nabla_{\tilde{p}} \bar{w})^\top + \frac{1}{2} \frac{\gamma}{1-\gamma} \Upsilon^\top (\Sigma_Y^\top \Sigma_Y)^{-1} \Upsilon - \frac{\gamma}{1-\gamma} (\nabla_{\tilde{p}} \bar{w}) \bar{\kappa} \Sigma_Y^\top \Upsilon, \end{aligned}$$

and therefore, after re-arrangement, we obtain Eq. (5.15).

Proof of Theorem 5.3.

Proof In order to ease the notational burden, throughout the proof we will write \tilde{p} for \tilde{p}° , \tilde{p}_s for \tilde{p}_s^t , π for π^t , $\tilde{\mathbb{P}}$ for $\tilde{\mathbb{P}}^t$, \mathbb{P} for \mathbb{P}^t , \tilde{W} for \tilde{W}^t , X for X^t , and \mathcal{G}_s^I for $\mathcal{G}_s^{t,I}$. Let us first remark that

$$\mathbb{P} \left(\tilde{p}_s \in \tilde{\Delta}_{N-1}, t \leq s \leq T \right) = 1. \quad (\text{B.4})$$

Indeed, set $\tilde{p}_s^N = 1 - \sum_{j=1}^{N-1} \tilde{p}_s^j$ and recall from Remark 3.4 that the process \tilde{p}_s^i is given by

$$\tilde{p}_s^i := \mathbb{P} \left(X_s = e_i | \mathcal{G}_s^I \right), \quad (t \leq s \leq T, i = 1, \dots, N). \quad (\text{B.5})$$

Therefore, using Lemma B.1, we deduce that all the \tilde{p}^i , with $i = 1, \dots, N$, remain positive in $[t, T]$, a.s., and, hence, (B.4) is satisfied.

Next, we prove that the feedback trading strategy $\tilde{\pi}_s := (\tilde{\pi}_s^S, \tilde{\pi}_s^P)^\top$, $\tilde{\pi}_s^P := 0$, is admissible; i.e.,

$$\mathbb{E}^\mathbb{P} \left[\exp \left(\frac{\sigma^2 \gamma^2}{2} \int_t^T \left(\tilde{\pi}^S(s, \tilde{p}_s) \right)^2 ds \right) \right] < \infty. \quad (\text{B.6})$$

We have that (B.6) follows from Eq. (B.4) and the fact that $(\tilde{\pi}^S(s, \tilde{p}))^2$ is uniformly bounded on $[0, T] \times \tilde{\Delta}_{N-1}$. To see the latter property, note that

$$\sup_{(s, \tilde{p}) \in [0, T] \times \tilde{\Delta}_{N-1}} \left(\tilde{\pi}^S(s, \tilde{p}) \right)^2 \leq \frac{2}{\sigma^4(1-\gamma)^2} \sup_{(s, \tilde{p}) \in [0, T] \times \tilde{\Delta}_{N-1}} (\tilde{\mu}(\tilde{p}) - r)^2 + \frac{2}{\sigma^2(1-\gamma)^2} \sup_{(s, \tilde{p}) \in [0, T] \times \tilde{\Delta}_{N-1}} (\nabla_{\tilde{p}} \underline{w}(s, \tilde{p}) \underline{\kappa}(\tilde{p}))^2.$$

The first term on the right hand side is clearly bounded since $|\tilde{\mu}(\tilde{p})| \leq \max_i |\mu_i|$ for any $\tilde{p} \in \tilde{\Delta}_{N-1}$. For the second term, using the definition of $\underline{\kappa}$ given in Eq. (5.5), we have

$$\sup_{(s, \tilde{p}) \in [0, T] \times \tilde{\Delta}_{N-1}} (\nabla_{\tilde{p}} \underline{w}(s, \tilde{p}) \underline{\kappa}(\tilde{p}))^2 = \frac{1}{\sigma^2} \sup_{(s, \tilde{p}) \in [0, T] \times \tilde{\Delta}_{N-1}} \left(\sum_{j=1}^{N-1} \partial_{\tilde{p}^j} \underline{w}(s, \tilde{p}) \tilde{p}^j \left(\mu_j - \sum_{i=1}^N \mu_i \tilde{p}^i \right) \right)^2, \quad (\text{B.7})$$

where $\tilde{p}^N := 1 - \sum_{i=1}^{N-1} \tilde{p}^i$. The last expression is bounded since each $\partial_{\tilde{p}^j} \underline{w}(s, \tilde{p})$ is bounded on $[0, T] \times \tilde{\Delta}_{N-1}$ in view of Lemma 5.1 and Remark 5.2, where it is shown $C_P^{2, \alpha}$ regularity for $\underline{w}(s, \tilde{p})$, hence bounded first and second order space derivatives on $\tilde{\Delta}_{N-1}$.

Now, fix an arbitrary feedback control $\pi_s^S := \pi^S(s, \tilde{p}_s)$ such that $(\pi^S, \pi^P) \in \mathcal{A}(t, T; \tilde{p}, 1)$, where $\pi_s^P \equiv 0$ and $\mathcal{A}(t, T; \tilde{p}, 1)$ is defined as in Definition 4.1, and define the process

$$M_s^{\pi^S} := e^{-\gamma \int_t^s \underline{\eta}(u, \tilde{p}_u, \pi_u^S) du} e^{\underline{w}(s, \tilde{p}_s)}, \quad (t \leq s \leq T),$$

where

$$\underline{\eta}(u, \tilde{p}, \pi^S) = \underline{\eta}(u, \tilde{p}, (\pi^S, 0)^\top) = -r + \pi^S(r - \tilde{\mu}(\tilde{p})) + \frac{1-\gamma}{2} \sigma^2 (\pi^S)^2. \quad (\text{B.8})$$

In what follows, we write for simplicity M^π for M^{π^S} and π for π^S . Note that the process $\{M_s^\pi\}_{t \leq s \leq T}$ is uniformly bounded. Indeed, (B.8) is convex in π^S and by minimizing it over π^S , it follows that, for any $\tilde{p} \in \tilde{\Delta}_{N-1}$,

$$-\underline{\eta}(t, \tilde{p}, \pi) \leq r + \frac{(\tilde{\mu}(\tilde{p}) - r)^2}{2(1-\gamma)\sigma^2} \leq r + \frac{(\max_i \mu_i^2 + r^2)}{(1-\gamma)\sigma^2} < \infty.$$

Therefore, since $\underline{w} \in C([0, T] \times \tilde{\Delta}_{N-1})$, there exists a constant $K < \infty$ for which

$$M_s^\pi = e^{-\gamma \int_t^s \underline{\eta}(u, \tilde{p}_u, \pi_u) du} e^{\underline{w}(s, \tilde{p}_s)} \leq K e^{\gamma \|\underline{\eta}\|_\infty (T-t)} =: A < \infty. \quad (\text{B.9})$$

We prove the result through the following steps:

(i) Define the process $\mathcal{Y}_s = e^{\underline{w}(s, \tilde{p}_s)}$. By Itô's formula and the generator formula (4.8) with $f(s, \tilde{p}) = e^{\underline{w}(s, \tilde{p})}$,

$$\begin{aligned} M_s^\pi &= M_t^\pi + \int_t^s e^{-\gamma \int_t^u \underline{\eta}(r, \tilde{p}_r, \pi_r) dr} d\mathcal{Y}_u - \gamma \int_t^s \underline{\eta}(u, \tilde{p}_u, \pi_u) e^{-\gamma \int_t^u \underline{\eta}(r, \tilde{p}_r, \pi_r) dr} \mathcal{Y}_u du \\ &= M_t^\pi + \int_t^s M_u^\pi \left(\frac{\partial \underline{w}}{\partial u} + \frac{1}{2} \text{tr}(\underline{\kappa} \underline{\kappa}^\top D^2 \underline{w}) + \frac{1}{2} (\nabla_{\tilde{p}} \underline{w}) \underline{\kappa} \underline{\kappa}^\top (\nabla_{\tilde{p}} \underline{w})^\top + (\nabla_{\tilde{p}} \underline{w}) \underline{\beta}_\gamma - \gamma \underline{\eta} \right) du + \int_t^s M_u^\pi \nabla_{\tilde{p}} \underline{w} \underline{\kappa} d\tilde{W}_u^{(1)}. \end{aligned}$$

Using the expression of $\underline{\eta}$ in (B.8) and some rearrangement, we may write M^π as

$$M_s^\pi = M_t^\pi + \int_t^s M_u^\pi R(u, \tilde{p}_u, \pi_u) du + \int_t^s M_u^\pi \nabla_{\tilde{p}} \underline{w} \underline{\kappa} d\tilde{W}_u^{(1)}$$

with

$$R(u, \tilde{p}, \pi) = \underline{w}_u + \frac{1}{2} \text{tr}(\underline{\kappa} \underline{\kappa}^\top D^2 \underline{w}) + \frac{1}{2} (\nabla_{\tilde{p}} \underline{w}) \underline{\kappa} \underline{\kappa}^\top (\nabla_{\tilde{p}} \underline{w})^\top + \gamma r + (\nabla_{\tilde{p}} \underline{w}) \underline{\beta}_\gamma - \gamma \pi (r - \tilde{\mu}(\tilde{p})) - \frac{\sigma^2}{2} \gamma (1-\gamma) \pi^2. \quad (\text{B.10})$$

Clearly, $R(u, \tilde{p}, \pi)$ is a concave function in π since $R_{\pi\pi} = -\sigma^2 \gamma (1-\gamma) < 0$. If we maximize $R(u, \tilde{p}, \pi)$ as a function of π for each (u, \tilde{p}) , we find that the optimum is given by (5.9). Upon substituting (5.9) into (B.10), we get that

$$R(u, \tilde{p}, \pi) \leq R(u, \tilde{p}, \tilde{\pi}^S(u, \tilde{p})) = \underline{w}_u + \frac{1}{2} \text{tr}(\underline{\kappa} \underline{\kappa}^\top D^2 \underline{w}) + \frac{1}{2(1-\gamma)} (\nabla_{\tilde{p}} \underline{w}) \underline{\kappa} \underline{\kappa}^\top (\nabla_{\tilde{p}} \underline{w})^\top + (\nabla_{\tilde{p}} \underline{w}) \underline{\Phi} + \underline{\Psi} = 0,$$

where the last equality follows from Eq. (5.7). Therefore, we get the inequality

$$\mathbb{E}^{\tilde{\mathbb{P}}} [M_T^\pi] \leq M_t^\pi + \mathbb{E}^{\tilde{\mathbb{P}}} \left[\int_t^T M_u^\pi \nabla_{\tilde{p}} \underline{w} \underline{\kappa} d\tilde{W}_u^{(1)} \right],$$

with equality if $\pi = \tilde{\pi}^S$. From (5.5), it is easy to check that $\sup_{\tilde{p} \in \tilde{\Delta}_{N-1}} \|\underline{\kappa}(\tilde{p})\|^2 \leq 2 \max_i \{\mu_i\} / \sigma$. Then, since the partial derivatives $\partial_{\tilde{p}^j} \underline{w}(s, \tilde{p})$ are uniformly bounded on $[0, T] \times \tilde{\Delta}_{N-1}$ (see also the argument after Eq. (B.7)), (B.9) implies that

$$\sup_{t \leq u \leq T} |M_u^\pi \nabla_{\tilde{p}} \underline{w} \underline{\kappa}|^2 \leq A \sup_{t \leq u \leq T} \|\underline{\kappa}(\tilde{p}_u)\|^2 \sup_{t \leq u \leq T} \|\nabla_{\tilde{p}} \underline{w}(u, \tilde{p}_u)\|^2 \leq B,$$

for some non-random constant $B < \infty$. We conclude that

$$\mathbb{E}^{\tilde{\mathbb{P}}} [M_T^\pi] \leq M_t^\pi = e^{\underline{w}(t, \tilde{p}_t)} = e^{\underline{w}(t, \tilde{p})}, \quad (\text{B.11})$$

with equality if $\pi = \tilde{\pi}^S$.

(ii) For simplicity, let us write $\tilde{\pi}_s := \tilde{\pi}^S(s, \tilde{p}_s)$. First, note that from the fact that we have equality in (B.11) when $\pi = \tilde{\pi}$,

$$e^{\underline{w}(t, \tilde{p})} = \mathbb{E}^{\tilde{\mathbb{P}}} [M_T^{\tilde{\pi}}] = \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{-\gamma \int_t^T \underline{\eta}(u, \tilde{p}_u, \tilde{\pi}_u) du} e^{\underline{w}(T, \tilde{p}_T)} \right] = \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{-\gamma \int_t^T \underline{\eta}(u, \tilde{p}_u, \tilde{\pi}_u) du} \right]. \quad (\text{B.12})$$

Similarly, for every feedback control $\pi_s = \pi(s, \tilde{p}_s)$ such that $(\pi, 0) \in \mathcal{A}(t, T; \tilde{p}, 1)$,

$$\mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{-\gamma \int_t^T \underline{\eta}(u, \tilde{p}_u, \pi_u) du} \right] = \mathbb{E}^{\tilde{\mathbb{P}}} [M_T^\pi] \leq e^{\underline{w}(t, \tilde{p})} = \mathbb{E}^{\tilde{\mathbb{P}}} \left[e^{-\gamma \int_t^T \underline{\eta}(u, \tilde{p}_u, \tilde{\pi}_u) du} \right],$$

where the inequality in the previous equation comes from (B.11) and the last equality therein follows from (B.12). The previous relationships show the optimality of $\tilde{\pi}$ and prove the assertions (1) and (2).

Proof of Theorem 5.4.

Proof For brevity, define the operator

$$\mathcal{B} = \partial_t + \frac{1}{2} \text{tr} \left(\bar{\kappa} \bar{\kappa}^\top D^2 \right) + \nabla_{\tilde{p}} \bar{\Phi}$$

and denote by

$$H(t, \tilde{p}, u) = -\tilde{h}(\tilde{p}) e^{\frac{w}{h} \left(t, \frac{1}{h(\tilde{p})} \tilde{p} \cdot h' \right)} \frac{u^\gamma}{1-\gamma}, \quad u \in \mathbb{R}_+,$$

the non-linear term of the PDE (5.18). Notice that since $\tilde{h} > 0$ by construction, then $H \leq 0$. Moreover, $u \mapsto H(t, \tilde{p}, u)$ is smooth and Lipschitz continuous on $[\bar{c}, +\infty[$ for any $\bar{c} > 0$, uniformly w.r.t. (t, \tilde{p}) . We set

$$\bar{\psi}_0(t, \tilde{p}) = e^{c(T-t)}, \quad t \in [0, T],$$

where c is a suitably large positive constant such that

$$cu + H(t, \tilde{p}, u) - \frac{\bar{\Psi}(t, \tilde{p})}{1-\gamma} u \geq 0, \quad \text{for any } (t, \tilde{p}) \in (0, T) \times \tilde{\Delta}_{N-1} \text{ and } u \geq 1. \quad (\text{B.13})$$

Then we define recursively the sequence $(\bar{\psi}_j)_{j \in \mathbb{N}}$ by

$$\begin{cases} \left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma} \right) \bar{\psi}_j - \lambda \bar{\psi}_j = H(\cdot, \cdot, \bar{\psi}_{j-1}) - \lambda \bar{\psi}_{j-1}, \\ \bar{\psi}_j(T, \cdot) = 1, \end{cases} \quad (\text{B.14})$$

where λ is the Lipschitz constant of $u \mapsto H(\cdot, \cdot, u)$ on $[\bar{c}, +\infty[$ and \bar{c} is the strictly positive constant defined as

$$\bar{c} = e^{-\frac{T}{1-\gamma}} \|\bar{\Psi}\|_\infty. \quad (\text{B.15})$$

Let us recall that the *linear problem* (B.14) has a classical solution in $C_P^{2,\alpha}$ whose existence can be proven as in Lemma 5.1, see also the following Remark 5.2. Next we prove by induction that

i) $(\bar{\psi}_j)$ is a decreasing sequence, that is

$$\bar{\psi}_{j+1} \leq \bar{\psi}_j, \quad j \geq 0; \quad (\text{B.16})$$

ii) $(\bar{\psi}_j)$ is uniformly strictly positive and in particular

$$\bar{\psi}_{j+1} \geq \bar{c}, \quad j \geq 0, \quad (\text{B.17})$$

with \bar{c} as in (B.15).

First, we observe that

$$\bar{\psi}_0 \geq 1, \quad \text{and} \quad \left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma} \right) \bar{\psi}_0 = \left(-c + \frac{\bar{\Psi}}{1-\gamma} \right) \bar{\psi}_0. \quad (\text{B.18})$$

Next we prove (B.16)-(B.17) for $j = 0$: by (B.18) and (B.13) we have

$$\begin{cases} \left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma} - \lambda \right) (\bar{\psi}_1 - \bar{\psi}_0) = H(\cdot, \cdot, \bar{\psi}_0) + c\bar{\psi}_0 - \frac{\bar{\Psi}}{1-\gamma} \bar{\psi}_0 \geq 0, \\ (\bar{\psi}_1 - \bar{\psi}_0)(T, \tilde{p}) = 0. \end{cases} \quad (\text{B.19})$$

where the inequality above follows from the fact that c is chosen as in (B.13), and $\bar{\psi}_0 \geq 1$ as observed in Eq. (B.18). Since the process \tilde{p}_t never reaches the boundary of the simplex by Lemma B.1, it follows from the Feynman-Kac representation theorem (or, equivalently, the maximum principle) that $\bar{\psi}_1 \leq \bar{\psi}_0$: indeed we have

$$(\bar{\psi}_1 - \bar{\psi}_0)(t, \tilde{p}) = \mathbb{E}^{\tilde{\mathbb{P}}} \left[- \int_t^T e^{-\int_t^s \left(\frac{1}{1-\gamma} \bar{\Psi}(r, \tilde{p}_r) - \lambda \right) dr} \left(H(s, \tilde{p}_s, \bar{\psi}_0) + c\bar{\psi}_0(s, \tilde{p}_s) - \frac{1}{1-\gamma} \bar{\Psi}(s, \tilde{p}_s) \bar{\psi}_0(s, \tilde{p}_s) \right) \Big|_{\tilde{p}_t = \tilde{p}} \right] \leq 0 \quad (\text{B.20})$$

where the last inequality follows directly from the inequality in (B.19). This proves (B.16) when $j = 0$. Using the recursive definition (B.14), along with the fact that $H \leq 0$, $\lambda > 0$ and inequality (B.20), we obtain

$$\left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma}\right) \bar{\psi}_1 = H(\cdot, \cdot, \bar{\psi}_0) + \lambda(\bar{\psi}_1 - \bar{\psi}_0) \leq 0. \quad (\text{B.21})$$

Then (B.17) with $j = 0$ follows again from the Feynman-Kac theorem: indeed by (B.21) we have

$$\begin{aligned} \bar{\psi}_1(t, \bar{p}) &= \mathbb{E}^{\bar{p}} \left[- \int_t^T e^{-\int_t^s \frac{1}{1-\gamma} \bar{\Psi}(r, \bar{p}_r) dr} (H(s, \bar{p}_s, \bar{\psi}_0) + \lambda(\bar{\psi}_1 - \bar{\psi}_0)(s, \bar{p}_s)) \Big| \bar{p}_t = \bar{p} \right] + \mathbb{E}^{\bar{p}} \left[e^{\frac{1}{1-\gamma} \int_t^T \bar{\Psi}(s, \bar{p}_s) ds} \Big| \bar{p}_t = \bar{p} \right] \\ &\geq e^{-\frac{T}{1-\gamma} \|\bar{\Psi}\|_\infty}, \end{aligned} \quad (\text{B.22})$$

where the last inequality follows from the positivity of the first expectation above guaranteed by (B.21).

Next we assume the inductive hypothesis to hold,

$$\bar{c} \leq \bar{\psi}_j \leq \bar{\psi}_{j-1} \quad (\text{B.23})$$

and prove (B.16)-(B.17). Recalling that λ is the Lipschitz constant of $u \mapsto H(\cdot, \cdot, u)$ on $[\bar{c}, +\infty[$, by (B.23) we have

$$\begin{cases} \left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma} - \lambda\right) (\bar{\psi}_{j+1} - \bar{\psi}_j) = H(\cdot, \cdot, \bar{\psi}_j) - H(\cdot, \cdot, \bar{\psi}_{j-1}) - \lambda(\bar{\psi}_j - \bar{\psi}_{j-1}) \geq 0, \\ (\bar{\psi}_{j+1} - \bar{\psi}_j)(T, \bar{p}) = 0. \end{cases}$$

Thus (B.16) follows from the Feynman-Kac theorem using the same procedure as in (B.19) and (B.20). Moreover we have

$$\left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma}\right) \bar{\psi}_{j+1} = H(\cdot, \cdot, \bar{\psi}_j) + \lambda(\bar{\psi}_{j+1} - \bar{\psi}_j) \leq 0, \quad (\text{B.24})$$

where the inequality above follows by (B.16) and using that $H \leq 0$ and $\lambda > 0$. Then, as in (B.22), we have that (B.17) follows from the Feynman-Kac theorem.

In conclusion, for $j \in \mathbb{N}$, we have

$$\bar{c} \leq \bar{\psi}_{j+1} \leq \bar{\psi}_j \leq \bar{\psi}_0. \quad (\text{B.25})$$

Now the thesis follows by proceeding as in the proof of Theorem 3.3 in Di Francesco et al. (2007). Indeed let us denote by $\bar{\psi}$ the pointwise limit of $(\bar{\psi}_j)$ as $j \rightarrow +\infty$: since $\bar{\psi}_j$ is a solution of (B.14) and by the uniform estimate (B.25), we can apply standard a priori Morrey-Sobolev-type estimates (see, Theorems 2.1 and 2.2 in Di Francesco et al. (2007)) to conclude that, for any $\alpha \in]0, 1[$, $\|\bar{\psi}_j\|_{C_P^{1,\alpha}((0,T) \times \tilde{\Delta}_{N-1})}$ is bounded by a constant only dependent on \mathcal{B} , α and λ . Hence by the classical Schauder interior estimate (see, for instance, Theorem 2.3 in Di Francesco et al. (2007)), we deduce that $\|\bar{\psi}_j\|_{C_P^{2,\alpha}((0,T) \times \tilde{\Delta}_{N-1})}$ is bounded uniformly in $j \in \mathbb{N}$. It follows that $(\bar{\psi}_j)_{j \in \mathbb{N}}$ admits a subsequence (denoted by itself) that converges in $C^{2,\alpha}$. Thus passing at limit in (B.14) as $j \rightarrow \infty$, we have

$$\left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma}\right) \bar{\psi} = H(\cdot, \cdot, \bar{\psi}), \quad \text{in } (0, T) \times \tilde{\Delta}_{N-1},$$

and $\bar{\psi}(T, \cdot) = 1$.

Finally, in order to prove that $\bar{\psi} \in C((0, T] \times \tilde{\Delta}_{N-1})$, we use the standard argument of barrier functions. We recall that w is a barrier function for the operator $\left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma}\right)$, on the domain $(0, T] \times \tilde{\Delta}_{N-1}$, at the point (T, \bar{p}) if $w \in C^2(V \cap ((0, T] \times \tilde{\Delta}_{N-1}))$ where V is a neighborhood of (T, \bar{p}) and we have

- i) $\left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma}\right) w \leq -1$ in $V \cap ((0, T) \times \tilde{\Delta}_{N-1})$;
- ii) $w > 0$ in $V \cap ((0, T) \times \tilde{\Delta}_{N-1}) \setminus \{(T, \bar{p})\}$ and $w(T, \bar{p}) = 0$.

Next we fix $\bar{p} \in \tilde{\Delta}_{N-1}$: following Friedman (1964) Chap.3 Sec.4, it is not difficult to check that

$$w(t, \bar{p}) = (|\bar{p} - \bar{x}|^2 + c_1(T-t)) e^{c_2(T-t)},$$

is a barrier at (T, \bar{p}) provided that c_1, c_2 are sufficiently large. Then we set

$$v^\pm(t, \bar{p}) = 1 \pm kw(t, \bar{p})$$

where k is a suitably large positive constant, independent of j , such that

$$\left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma}\right) (\bar{\psi}_j - v^+) \geq H(\cdot, \cdot, \bar{\psi}_{j-1}) - \lambda(\bar{\psi}_{j-1} - \bar{\psi}_j) - \frac{\bar{\Psi}}{1-\gamma} - k \left(\mathcal{B} + \frac{\bar{\Psi}}{1-\gamma}\right) w \geq 0,$$

and $\bar{\psi}_j \leq v^+$ on $\partial(V \cap ((0, T) \times \tilde{\Delta}_{N-1}))$. The maximum principle yields $\bar{\psi}_j \leq v^+$ on $V \cap ((0, T) \times \tilde{\Delta}_{N-1})$; analogously we have $\bar{\psi}_j \geq v^-$ on $V \cap ((0, T) \times \tilde{\Delta}_{N-1})$, and letting $j \rightarrow \infty$ we get

$$1 - kw(t, \bar{p}) \leq \bar{\psi}(t, \bar{p}) \leq 1 + kw(t, \bar{p}), \quad (t, \bar{p}) \in V \cap ((0, T) \times \tilde{\Delta}_{N-1}).$$

Therefore we deduce that

$$\lim_{(t, \bar{p}) \rightarrow (T, \bar{p})} \bar{\psi}(t, \bar{p}) = 1$$

which concludes the proof.

Proof of Theorem 5.5

Proof As in the proof of Theorem 5.3, to ease the notational burden we will write \tilde{p} for \tilde{p}° , \tilde{p}_s for \tilde{p}_s^t , π for π^t , \tilde{W} for \tilde{W}^t , $\tilde{\mathbb{P}}$ for $\tilde{\mathbb{P}}^t$, \mathbb{P} for \mathbb{P}^t , and \mathcal{G}_s^I for $\mathcal{G}_s^{t,I}$. Similarly to the proof of the post default verification theorem, it is easy to see that the trading strategy $\tilde{\pi}_s := (\tilde{\pi}_s^S, \tilde{\pi}_s^P)^\top = \left(\tilde{\pi}^S(s, \tilde{p}_{s-}, H_{s-}^t), \tilde{\pi}^P(s, \tilde{p}_{s-}, H_{s-}) \right)^\top$, as defined from equations (5.21)-(5.22), is admissible; i.e., satisfies (4.5). This essentially follows from the condition (2.9) and the fact that both $\underline{w}(s, \tilde{p})$ and $\bar{w}(s, \tilde{p})$ belong to $\mathcal{C}_P^{2,\alpha}$, hence their first and second order space derivatives are bounded on $[0, T] \times \tilde{\Delta}_{N-1}$. Here, it is also useful to recall that $\mathbb{P}(\tilde{p}_s \in \tilde{\Delta}_{N-1}, t \leq s \leq T) = 1$ as shown in the proof of Theorem 5.3.

Next, for a fixed feedback control $\pi_s := (\pi_s^S, \pi_s^P) := (\pi^S(s, \tilde{p}_{s-}, H_{s-}), \pi^P(s, \tilde{p}_{s-}, H_{s-}))$ such that $(\pi^S, \pi^P) \in \tilde{\mathcal{A}}(t, T; \tilde{p}, 0)$, define the process

$$M_s^\pi := e^{-\gamma \int_t^s \tilde{\eta}(u, \tilde{p}_u, \pi_u) du} e^{w(s, \tilde{p}_s, H_s)}, \quad (t \leq s \leq T), \quad (\text{B.26})$$

where $w(s, \tilde{p}, z) := (1-z)\bar{w}(s, \tilde{p}) + z\underline{w}(s, \tilde{p})$ and $\tilde{\eta}$ is defined as in Eq. (4.17). Note that $\tilde{\eta}$ can be written as

$$\tilde{\eta}(t, \tilde{p}, \pi) = -r + \pi^S(r - \tilde{\mu}(\tilde{p})) + \frac{1-\gamma}{2}\sigma^2(\pi^S)^2 + \pi^P(r - \tilde{a}(t, \tilde{p})) + \frac{1-\gamma}{2}v^2(\pi^P)^2,$$

and, thus, $-\tilde{\eta}$ is concave in π . This in turn implies that there exists a nonrandom constant $A < \infty$ such that

$$0 < M_s^\pi \leq A < \infty, \quad t \leq s \leq T, \quad (\text{B.27})$$

since $\underline{w}, \bar{w} \in C([0, T] \times \tilde{\Delta}_{N-1})$. We prove the result through the following two steps:

(i) Define the processes $\mathcal{Y}_s = e^{w(s, \tilde{p}, H_s)}$ and $\mathcal{U}_s = e^{-\gamma \int_t^s \tilde{\eta}(u, \tilde{p}_u, \pi_u) du}$. By Itô's formula, the generator formula (4.8) with $f(s, \tilde{p}, z) = e^{w(s, \tilde{p}, z)}$, and the same arguments as those used to derive (4.15),

$$\begin{aligned} M_s^\pi &= M_t^\pi + \int_t^s \mathcal{U}_u - d\mathcal{Y}_u - \gamma \int_t^s \tilde{\eta}(u, \tilde{p}_u, \pi_u) \mathcal{U}_u \mathcal{Y}_u du \\ &= M_t^\pi + \int_t^s M_u^\pi \left[\frac{\partial w}{\partial u} + \frac{1}{2} \text{tr}(\kappa \kappa^\top D^2 w) + \frac{1}{2} (\nabla_{\tilde{p}} w) \kappa \kappa^\top (\nabla_{\tilde{p}} w)^\top + (\nabla_{\tilde{p}} w) \beta_\gamma \right. \\ &\quad \left. + (1 - H_u) \tilde{h}(\tilde{p}_u) \left(e^{\frac{w}{h} \left(u, \frac{1}{h(\tilde{p}_u)} \tilde{p}_u \cdot h' \right) - \bar{w}(u, \tilde{p}_u)} - 1 \right) - \gamma \tilde{\eta} \right] du + \mathcal{M}_s^c + \mathcal{M}_s^d, \end{aligned}$$

where

$$\mathcal{M}_s^c := \int_t^s M_u^\pi \nabla_{\tilde{p}} w \kappa(u, \tilde{p}_u) d\tilde{W}_u, \quad \mathcal{M}_s^d := \int_t^s \mathcal{U}_u - \left(e^{\frac{w}{h} \left(u, \frac{1}{h(\tilde{p}_u)} \tilde{p}_u \cdot h' \right) - \bar{w}(u, \tilde{p}_u)} - e^{\bar{w}(u, \tilde{p}_u)} \right) d\tilde{\xi}_u. \quad (\text{B.28})$$

Using the expression of η in Eq. (4.17), and similar arguments to those used to derive (4.17), we may write M^π as

$$M_s^\pi = M_t^\pi + \int_t^s M_u^\pi R(u, \tilde{p}_u, \pi_u, H_u) du + \mathcal{M}_s^c + \mathcal{M}_s^d$$

with

$$\begin{aligned} R(u, \tilde{p}, \pi, z) &= \frac{\partial w}{\partial u} + \frac{1}{2} \text{tr}(\kappa \kappa^\top D^2 w) + \frac{1}{2} (\nabla_{\tilde{p}} w) \kappa \kappa^\top (\nabla_{\tilde{p}} w)^\top + \gamma r + (1-z) \tilde{h}(\tilde{p}) \left[e^{\frac{w}{h} \left(u, \frac{1}{h(\tilde{p})} \tilde{p} \cdot h' \right) - \bar{w}(u, \tilde{p})} - 1 \right] \\ &\quad + z \left((\nabla_{\tilde{p}} \underline{w}) \beta_\gamma - \gamma \pi^S(r - \tilde{\mu}(\tilde{p})) - \frac{\gamma(1-\gamma)}{2} \sigma^2(\pi^S)^2 \right) \\ &\quad + (1-z) \left((\nabla_{\tilde{p}} \bar{w}) \beta_\gamma - \gamma \pi^P(r - \tilde{a}(t, \tilde{p})) - \frac{\gamma(1-\gamma)}{2} v^2(\pi^P)^2 \right) \end{aligned} \quad (\text{B.29})$$

Clearly, $R(u, \tilde{p}, \pi, z)$ is a concave function in π for each (u, \tilde{p}, z) . Furthermore, this function reaches its maximum at $\tilde{\pi}(u, \tilde{p}, z) = (\tilde{\pi}^S(u, \tilde{p}, z), \tilde{\pi}^P(u, \tilde{p}, z))$ as defined in the statement of the theorem. Upon substituting this maximum into (B.29) and rearrangements similar to those leading to (5.7) and (5.15) (depending on whether $z = 1$ or $z = 0$), we get

$$R(u, \tilde{p}, \pi, z) \leq R(u, \tilde{p}, \tilde{\pi}(u, \tilde{p}, z), z) = 0,$$

in light of the corresponding equations (5.7) and (5.15). Therefore, we get the inequality

$$\mathbb{E}^{\tilde{\mathbb{P}}} [M_T^\pi] \leq M_t^\pi + \mathbb{E}^{\tilde{\mathbb{P}}} [\mathcal{M}_T^c + \mathcal{M}_T^d],$$

with equality if $\pi = \tilde{\pi}$. Note that $\mathbb{E}^{\tilde{\mathbb{P}}} [\mathcal{M}_T^c] = 0$ since it is possible to find a nonrandom constant B such that

$$\sup_{t \leq u \leq T} |M_u^\pi \nabla_{\tilde{p}} w \kappa(u, \tilde{p}_u)|^2 \leq A \sup_{t \leq u \leq T} \|\kappa(u, \tilde{p}_u)\|^2 \sup_{t \leq u \leq T} \|\nabla_{\tilde{p}} w(u, \tilde{p}_u)\|^2 \leq B,$$

in view of (B.27) and the fact that the partial derivatives of \underline{w} and \bar{w} are uniformly bounded on $[0, T] \times \tilde{\Delta}_{N-1}$. The latter statement follows from the fact that both \underline{w} and \bar{w} are $\mathcal{C}_P^{2,\alpha}$ on $\tilde{\Delta}_{N-1}$ in light of Lemma 5.1 and Theorem 5.4. To deal with \mathcal{M}^d , note that since

$w, \bar{w} \in C([0, T] \times \tilde{\Delta}_{N-1})$ and $\{\mathcal{U}_s\}_{t \leq s \leq T}$ is uniformly bounded (due to the fact that $-\tilde{\eta}$ is concave), we have that the integrand of the second integral in (B.28) is uniformly bounded and, thus, $\mathbb{E}^{\tilde{\mathbb{P}}}[\mathcal{M}_T^d] = 0$ as well. The two previous facts, together with the initial conditions $H_t = 0$ and $\tilde{p}_t = \tilde{p}$, lead to

$$\mathbb{E}^{\tilde{\mathbb{P}}}[M_T^\pi] \leq M_t^\pi = e^{w(t, \tilde{p}_t, H_t)} = e^{w(t, \tilde{p}, 0)} = e^{\bar{w}(t, \tilde{p})}, \quad (\text{B.30})$$

with equality if $\pi = \tilde{\pi}$.

(ii) The rest of the proof is similar to the post default case. Concretely, using the fact that we have equality in (B.30) when $\pi = \tilde{\pi}$,

$$e^{\bar{w}(t, \tilde{p})} = \mathbb{E}^{\tilde{\mathbb{P}}}[M_T^{\tilde{\pi}}] = \mathbb{E}^{\tilde{\mathbb{P}}}\left[e^{-\gamma \int_t^T \tilde{\eta}(u, \tilde{p}_u, \tilde{\pi}_u) du} e^{w(T, \tilde{p}_T, H_T)}\right] = \mathbb{E}^{\tilde{\mathbb{P}}}\left[e^{-\gamma \int_t^T \tilde{\eta}(u, \tilde{p}_u, \tilde{\pi}_u) du}\right], \quad (\text{B.31})$$

since $w(T, \tilde{p}_T, H_T) := (1 - H_T)\bar{w}(T, \tilde{p}_T) + H_T w(T, \tilde{p}_T) \equiv 0$. Also, from (B.30), for every feedback control $\pi_s = \pi(s, \tilde{p}_s, H_s) \in \mathcal{A}(t, T; \tilde{p}, 0)$,

$$\mathbb{E}^{\tilde{\mathbb{P}}}\left[e^{-\gamma \int_t^T \tilde{\eta}(u, \tilde{p}_u, \pi_u) du}\right] = \mathbb{E}^{\tilde{\mathbb{P}}}[M_T^\pi] \leq M_t^\pi = e^{\bar{w}(t, \tilde{p})} = \mathbb{E}^{\tilde{\mathbb{P}}}\left[e^{-\gamma \int_t^T \tilde{\eta}(u, \tilde{p}_u, \tilde{\pi}_u) du}\right],$$

where the last equality above follows from (B.31). This proves the assertions (1) and (2).

References

- Belanger, A., Shreve, S. and Wong, D. A general framework for pricing credit risk. *Mathematical Finance* 14(3), 317–350, 2004.
- Bielecki, T., and Jang, I., Portfolio optimization with a defaultable security. *Asia-Pacific Financial Markets* 13(2), 113-127, 2006.
- Bielecki, T., and Rutkowski, M. *Credit Risk: Modelling, Valuation and Hedging*, Springer, New York, NY, 2001.
- Bo, L., A. Capponi. Optimal investment in credit derivatives portfolio under contagion risk. *Mathematical Finance*. Forthcoming, 2014.
- Bo, L., Wang, Y., and Yang, X., An optimal portfolio problem in a defaultable market. *Advances in Applied Probability* 42(3), 689-705, 2010.
- Capponi, A., and Figueroa-López. Dynamic Portfolio Optimization with a Defaultable Security and Regime-Switching Markets. *Mathematical Finance* 24(2), 207–249, 2014.
- Carr, P., Linetsky, V., and Mendoza-Arriaga, R. Time-Changed Markov Processes in Unified Credit-Equity Modeling. *Mathematical Finance* 20(4), 527–569, 2010.
- Di Francesco, M., Pascucci, A., and Polidoro, S. The obstacle problem for a class of hypoelliptic ultraparabolic equations. *Proceedings of The Royal Society of London. Series A. Mathematical, Physical and Engineering Sciences*, 464, 2089, 155-176, 2008.
- El Karoui, N., Jeanblanc, M. and Jiao, Y. What happens after a default: the conditional density approach. *Stochastics Processes and their Applications* 120(7), 1011-1032, 2010.
- Elliott, R. J., Aggoun, L., and Moore, J. B. *Hidden Markov models: estimation and control*. Berlin Heidelberg New York: Springer, 1994.
- Elliott, R.J. and Siu, T.K. A Hidden Markov Model for Optimal Investment of An Insurer with Model Uncertainty. *International Journal of Robust and Nonlinear Control* 22(7), 778807, 2012.
- Frey, R., and Runggaldier, W.J. Credit Risk and Incomplete Information: a Nonlinear-Filtering Approach. *Finance and Stochastics* 14, 495-526, 2010.
- Frey, R., and Runggaldier, W.J. Nonlinear Filtering in Models for Interest-Rate and Credit Risk. In: *The Oxford Handbook on Nonlinear Filtering* (D.Crisan, B.Rozovski, eds.), Oxford University Press, 923-959, 2011.
- Frey, R., and Schmidt, T. Pricing and Hedging of Credit Derivatives via the Innovations Approach to Nonlinear Filtering. *Finance and Stochastics* 16(1), 105-133, 2012.
- Friedman, A. *Partial differential equations of parabolic type*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.
- Fujimoto, K., Nagai, H., and Runggaldier, W.J. Expected log-utility maximization under incomplete information and with Cox-process observations. *Asia-Pacific Financial Markets* 21(1), 35-66, 2014.
- Fujimoto, K., Nagai, H., and Runggaldier, W.J. Expected power-utility maximization under incomplete information and with Cox-process observations. *Applied Mathematics and Optimization* 67, 33-72, 2013.
- Giesecke, K., Longstaff, F., Schaefer, S., and Strebulaev, I. Corporate Bond Default Risk: A 150-Year Perspective. *Journal of Financial Economics* 102(2), 233-250, 2011.
- Jiao, Y., Kharroubi, I., and Pham, H. Optimal investment under multiple defaults risk: a BSDE-decomposition approach. *Annals of Applied Probability* 23(2), 455-491, 2013.
- Jiao, Y., and Pham, H. Optimal investment with counterparty risk: a default density approach, *Finance and Stochastics* 15(4), 725-753, 2011.
- Karatzas, I., and Shreve, S. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, 1988.
- Kharroubi, I., and Lim, T. Progressive enlargement of filtrations and Backward SDEs with jumps. *Journal of Theoretical Probability*, <http://dx.doi.org/10.1007/s10959-012.0428-1>, 2012.
- Kliemann, W., Koch, G., and Marchetti, F. On the unnormalized solution of the filtering problem with counting process observations. *IEEE Trans. Inf. Theory* 36, 1415-1425, 1990.
- Kraft, H., and Steffensen, M., Portfolio problems stopping at first hitting time with application to default risk. *Mathematical Methods of Operations Research* 63, 123-150, 2005.
- Kraft, H., and Steffensen, M., Asset allocation with contagion and explicit bankruptcy procedures. *Journal of Mathematical Economics* 45, 147-167, 2009.
- Kraft, H., and Steffensen, M., How to invest optimally in corporate bonds. *Journal of Economic Dynamics and Control* 32, 348-385, 2008.

- Liechty, J., and Roberts, G. Markov Chains Monte-Carlo Methods for Switching Diffusion Models. *Biometrika*, 88(2), 299–315, 2001.
- Linetsky, V. Pricing Equity Derivatives subject to Bankruptcy. *Mathematical Finance* 16(2), 255–282, 2006.
- McKean, H. P. A free boundary problem for the heat equation arising from a problem in mathematical economics, *Ind. Management Rev.*, 6 (1965), pp. 3239.
- Nagai, H., and Runggaldier, W. PDE Approach to Utility Maximization for Market Models with Hidden Markov Factors. In *Seminars on Stochastics Analysis, Random Fields, and Applications V*, Progress in Probability 59, 493-506, 2008.
- Pham, H. Stochastic control under progressive enlargement of filtrations and applications to multiple defaults risk management. *Stochastic Processes and their applications* 120, 1795-1820, 2010.
- Protter, P., *Stochastic Integrations and Differential Equations*, 2nd Edition. Springer-Verlag, New York, 2004.
- Protter, P., and Shimbo, K. No Arbitrage and General Semimartingales. In *Markov Processes and Related Topics: A Festschrift for Thomas G. Kurtz*. IMS Collections, 4, 267-283, 2008.
- Rogers, C., and Williams, D. *Diffusions, Markov Processes, and Martingales: Itô calculus*, Wiley, 1987.
- Sass, J., and Haussmann, U. Optimizing the terminal wealth under partial information: The drift process as a continuous time Markov chain. *Finance and Stochastics* 8(4), 553-577, 2004.
- Siu, T.K. A BSDE approach to optimal investment of an insurer with hidden regime switching. *Stochastic Analysis and Applications* 31(1), 1-18, 2013.
- Sotomayor, L., and Cadenillas, A. Explicit Solutions of consumption investment problems in financial markets with regime switching. *Mathematical Finance* 19(2), 251–279, 2009.
- Tamura, T., and Watanabe, Y. Risk-Sensitive portfolio optimization problems for hidden Markov factors on infinite time horizon. *Asymptotic Analysis* 75, 169–209, 2011.
- Wonham, W. M. Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering, *J. Soc. Indust. Appl. Math. Ser. A Control* 2, 347-369 1965.
- Wong, E., and Hajek, B. *Stochastic Processes in Engineering Systems*, Springer Verlag, 1985.
- Zariphopoulou, T. Investment-Consumption Models with Transaction Fees and Markov-Chain Parameters. *Siam Journal on Control and Optimization* 30(3), 613-636.