Optimal Liquidation under Partial Information with Price Impact

Katia Colaneri ‡ Zehra Eksi ‡ Rüdiger Frey ‡ Michaela Szölgyenyi ¶

Abstract

We study the problem of a trader who wants to maximize the expected revenue from liquidating a given stock position. We model the stock price dynamics as a geometric pure jump process with local characteristics driven by an unobservable finite-state Markov chain and by the liquidation rate. This reflects uncertainty about activity of other traders and feedback effects from trading. We use stochastic filtering to reduce the optimization problem under partial information to an equivalent one under complete information. This leads to a stochastic control problem for piecewise deterministic Markov processes (PDMPs). We carry out a detailed mathematical analysis of this problem via In particular, we derive the optimality equation for the value function, we characterize the value function as continuous viscosity solution of the associated dynamic programming equation, and we prove a novel comparison result. The paper concludes with numerical results illustrating the impact of partial information and feedback effects on the value function and on the optimal liquidation rate.

Keywords: Optimal liquidation, Stochastic filtering, Piecewise deterministic Markov process, Viscosity solutions and comparison principle.

1 Introduction

Traders on financial markets frequently face the task of selling a large amount of a given asset over a short time period. This has sparked a lot of research on how to carry out such transactions in an optimal way. The existing literature on optimal portfolio liquidation can essentially be divided into two classes: market impact models and order book models. In a market impact model one directly specifies the impact of a given trading strategy on the bid price of the asset. The fundamental price (the price if the trader is inactive) is usually modelled as a diffusion process such as geometric Brownian motion. In an order book model one specifies the dynamics of the limit order book which is more complex but results in an endogenous price impact.
It is well known that on fine time scales the bid price of an asset is best described by a pure jump process, since in reality the bid price is constant between events and jumps only when a market sell order is executed or when a new limit buy order above the current bid price arrives. In this paper we therefore study the optimal liquidation problem in a novel market impact model where the bid price follows a marked point process. Following the literature we assume that the trader uses absolutely continuous liquidation strategies with liquidation rate \( \nu_t \) for \( 0 \leq t \leq T \). In our setup the local characteristics (intensity and jump size distribution) of the bid price process depend on \( \nu_t \), so that there is permanent price impact; in particular, a high liquidation rate will push prices down. Moreover, we assume that the local characteristics depend on an unobservable Markov chain \( Y \), so that the liquidity and the trend of the market are random and not directly observable. The chain \( Y \) can among others be used to model feedback effect from the trading activity of the rest of the market. It is well known that in reality the order flow is clustered in time: there are periods with a lot of buy orders and periods with a lot of sell orders. Cont [25] attributes this to the fact that many observed orders are components of a larger parent order that is executed in small blocks. This clustering of the order flow can be mimicked by having a high intensity of downward jumps in one state of \( Y \) and a high intensity of upward jumps in another state of \( Y \). Note that with this interpretation of \( Y \) in mind, it is natural to assume that the chain is not directly observable, since the trading activity of other investors can be observed only indirectly via its impact on the price.

In mathematical terms the trader faces a control problem for a marked point processes with partial information (as \( Y \) is unobservable). The first step in the analysis of this problem is to derive an equivalent control problem under full information via stochastic filtering. In order to circumvent the issue that the information available to the investor depends on her liquidation strategy, we resort to the reference probability approach. Hence we cannot use the existing filtering results for marked point process observation which are mostly derived using the innovations approach. We end up with a control problem whose state process \( X \) consists of the stock price, the inventory level and the filter process. We provide a detailed mathematical analysis of this problem. The form of the asset return dynamics implies that \( X \) is a piecewise deterministic Markov process (PDMP) so that we rely on control theory for PDMPs; a general introduction to this theory is given in Davis [30] or in Bäuerle and Rieder [13]. We establish the dynamic programming equation for the value function and we derive conditions on the data of the problem that guarantee the continuity of the value function and the existence of optimal relaxed controls.

For this we need a careful analysis of the behavior of the value function close to the boundary. As a further step, using ideas from Davis and Farid [31] we characterize the value function as continuous viscosity solution of the Hamilton Jacobi Bellman (HJB) partial integro-differential equation associated with the problem. Moreover, we prove a novel comparison theorem for the HJB equation which is valid in more general setups. A comparison principle is important to ensure the convergence of numerical schemes for the value function, see Barles and Souganidis [9]. In order to gain intuition on the form of the resulting optimal liquidation rate we carry out a numerical study, based on discretization of the HJB equation. We calculate the additional liquidation profit from the use of stochastic filtering and we study the influence of the temporary and permanent price impact parameters on the form of the liquidation strategy. We find that for certain parameter constellations the optimal strategy displays a surprising yet economically plausible gambling behaviour of the trader that cannot be guessed upfront.

We continue with a brief discussion of the existing literature. Starting with market impact models, the first contribution is Bertsimas and Lo [16] who analyze the optimal portfolio execution problem for a risk-neutral agent in a model with linear and purely permanent price impact where
the fundamental price follows an arithmetic random walk. This model has been generalized in continuous time by Almgren and Chriss [3] allowing for risk aversion and temporary price impact. Since then, market impact models have been extensively studied. Important contributions include He and Mamaysky [12], Schied and Schöneborn [50], Bian et al. [17], Schied [19], Ankirchner et al. [9], Guo and Zervos [11]. Recently, Cayé and Muhle-Karbe [21] considered an extension of the Almgren Chriss model with a self-exciting temporary price impact. Bian et al. [17] use dynamic programming equations and viscosity solutions to study a model with market impact and regime switching. All these models work in a (discretized) diffusion framework.

In the order book literature on the other hand a few contributions based on point-process models exist. Bayraktar and Ludkovski [14] study the optimal portfolio execution problem in a model with discrete order flow represented by a Poisson process with observable intensity. The price impact is purely temporary and is represented in terms of a cost function. Bäuerle and Rieder [12] consider the same setting with a standard Poisson process and solve the cost minimization problem by using tools from the control theory of PDMPs. A further order book model based on point process methodology is Bayraktar and Ludkovski [13]. There it is assumed that the trader uses limit orders and that she can control the intensity of the order flow by choosing the spread at which she is willing to trade. Additional contributions based on diffusion models are Alfonsi et al. [1, 2], Obizhaeva and Wang [17], Cartea and Jaimungal [19]. For a detailed overview we refer to the surveys Gökay et al. [40], Gatheral and Schied [39], Cartea et al. [20]. From a methodological point of view our analysis is also related to the literature on expected utility maximization or hedging for pure jump process such as Bäuerle and Rieder [11] or Kirch and Runggaldier [44]. Important contributions to the control theory of PDMPs include among others Davis [30], Lenhart and Liao [15], Costa and Davis [26], Dempster and Ye [32], Almudevar [3], Forwick et al. [36], Bäuerle and Rieder [12], Costa and Dufour [27].

The outline of the paper is the following. In Section 2, we introduce our model, the main assumptions and the optimization problem. In Section 3, we derive the filtering equations for our model. Section 4 contains the mathematical analysis of the optimization problem via PDMP techniques. In Section 5, we provide a viscosity solution characterization of the value function. Finally, in Section 6, we present the results of our numerical experiments. The appendix contains additional proofs and some background on Markov decision models.

2 The Model

Throughout we work on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where the filtration $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ satisfies the usual conditions. Here $\mathbb{F}$ is the global filtration, i.e. all considered processes are $\mathbb{F}$-adapted, and $\mathbb{P}$ is the historical probability measure. We consider a trader who wants to liquidate $w_0 > 0$ units of a given security (referred to as the stock in the sequel) over the period $[0, T]$ for a given time horizon $T$. We denote the bid price process by $S = (S_t)_{t \geq 0}$ and $\mathbb{F}^S$ is the filtration generated by $S$. In what follows we assume that $\mathbb{F}^S$ satisfies the usual conditions. We assume that the trader sells the shares at a nonnegative $\mathbb{F}^S$-adapted rate $\nu = (\nu_t)_{0 \leq t \leq T}$. Hence her inventory, i.e. the amount of shares she holds at time $t \in [0, T]$, is given by

$$W_t = w_0 - \int_0^t \nu_u \, du, \quad t \in [0, T].$$

Modelling the inventory as an absolutely continuous process corresponds to the situation where the trader is frequently submitting small sell orders. By taking $\nu$ to be nonnegative, we confine the trader to pure selling strategies.
The goal of the trader is to maximize the expected revenue from her trading strategy. We assume that the implementation of the liquidation strategy generates temporary and permanent price impact, where permanent price impact is the impact of trading on the dynamics of $S$ and temporary price impact is the impact of trading on the execution price of the current trade.

2.1 Dynamics of the bid price. We model the bid price as a Markov-modulated geometric finite activity pure jump process. Let $Y = (Y_t)_{t \geq 0}$ be a continuous-time finite-state Markov chain on $(\Omega, F, \mathbb{F}, P)$ with state space $\mathcal{E} = \{e_1, e_2, \ldots, e_K\}$ ($e_k$ is $k$-th unit vector in $\mathbb{R}^K$), generator matrix $Q = (q_{ij})_{i,j=1,\ldots,K}$ and initial distribution $\pi_0 = (\pi_0^1, \cdots, \pi_0^K)$. We assume that the bid price has the dynamics

$$dS_t = S_{t-}dR_t,$$

where the return process $R = (R_t)_{t \geq 0}$ is a finite activity pure jump process. We assume that $\Delta R_t > -1$ so that $S$ is strictly positive. Denote the random measure associated with $R$ by

$$\mu^R(dt, dz) := \sum_{u \geq 0, \Delta R_u \neq 0} \delta_{\{u, \Delta R_u\}}(dt, dz),$$

and by $\eta^P(dt, dz)$ the $(\mathbb{F}, P)$-dual predictable projection (or compensating random measure) of $\mu^R$. $\eta^P$ is absolutely continuous and of the form

$$\eta^P(dt, dz) = \eta^P(t, Y_{t-}, \nu_{t-}, dz)dt,$$

for a finite measure $\eta^P(t, e, \nu, dz)$ on $\mathbb{R}$. The fact that $\eta^P(t, e, \nu)$ depends explicitly on time can be used to model the strong intraday seasonality patterns observed for high frequency data.

In order to specify the joint law of $Y$ and $R$, let $G = \{e_k - e_l: 1 \leq k, l \leq K, k \neq l\}$ be the set of possible jumps of the chain $Y$ and denote the elements of $G$ by $\gamma$. Let $\mu^{(Y,R)}(dt, d\gamma, dz)$ be the jump measure of the pair $(Y, R)$ on $G \times \mathbb{R}$. We assume that the compensator $\tilde{\eta}^P$ of $\mu^{(Y,R)}$ is absolutely continuous, $\tilde{\eta}^P(dt, d\gamma, dz) = \tilde{\eta}^P(t, Y_{t-}, \nu_{t-}, d\gamma, dz)dt$, where

$$\tilde{\eta}^P(t, e, \nu, d\gamma, dz) = \left(\sum_{i=1}^K 1_{\{Y_{t-} = e_i\}} (\eta^P(t, e, \nu, dz) + \sum_{j \neq i} q_{ij} \delta_{\{e_j - e_i\}}(d\gamma))\right)dt. \quad (2.3)$$

Note that $(2.3)$ implies that $Y$ and $R$ have no common jumps.

Now we turn to the semimartingale decomposition of the bid price with respect to the full information filtration $\mathbb{F}$. Denote by

$$\tilde{\eta}^P(t, e, \nu, dz) := \int_{\mathbb{R}} z \eta^P(t, e, \nu, dz) \quad (2.4)$$

the mean of $\eta^P(t, e, \nu, dz)$ ($\eta^P(t, e, \nu)$ exists under Assumption 2.1 below). Fix some liquidation strategy $\nu$. Then the martingale part $M^R$ of the return process is given by $M^R_t = R_t - \int_0^t \tilde{\eta}^P(s, Y_{s-}, \nu_{s-})ds$, $t \in [0, T]$, and the $\mathbb{F}$-semimartingale decomposition of $S$ equals

$$S_t = S_0 + \int_0^t S_{s-} dM^R_s + \int_0^t S_{s-} \tilde{\eta}^P(s, Y_{s-}, \nu_{s-})ds, \quad t \in [0, T]. \quad (2.5)$$

It is well-known that the semimartingale decomposition of $S$ with respect to the trader filtration $\mathbb{F}^S$ is obtained by projecting the process $\tilde{\eta}^P(t, Y_{t-}, \nu_{t-})$ onto $\mathbb{F}^S$.

In the sequel we assume that for all $t \in [0, T], e \in \mathcal{E}$, the mapping $\nu \mapsto \tilde{\eta}^P(t, e, \nu)$ is decreasing on $[0, \infty)$, that is selling pushes the price down on average. Furthermore, we make the following regularity assumption.
**Assumption 2.1 (Properties of $\eta^P$).** There is a deterministic finite measure $\eta^Q(dz)$ on $\mathbb{R}$ whose support, denoted $\text{supp}(\eta)$, is a compact subset of $(-1, \infty)$, such that for all $(t,e,\nu) \in [0,T] \times \mathcal{E} \times [0,\infty)$ the measure $\eta^P(t,e,\nu, dz)$ is equivalent to $\eta^Q(dz)$. Furthermore, for every $\nu^{\max} < \infty$ there is some constant $M > 0$ such that

$$M^{-1} < \frac{d\eta^P(t,e,\nu)}{d\eta^Q}(z) < M \quad \text{for all} \quad (t,e,\nu) \in [0,T] \times \mathcal{E} \times [0,\nu^{\max}]. \tag{2.6}$$

The assumption implies that for every $\nu^{\max}$ there is a $\lambda^{\max} < \infty$ such that

$$\sup\{\eta^P(t,e,\nu, \mathbb{R}) : t \in [0,T], e \in \mathcal{E}, \nu \in [0,\nu^{\max}]\} \leq \lambda^{\max}, \tag{2.7}$$

in particular the counting process associated to the jumps of $S$ is $P$-nonexplosive. Moreover, it provides a sufficient condition for the existence of a *reference probability measure*, i.e. a probability measure $Q$ equivalent to $P$ on $(\Omega, \mathcal{F}_T)$, such that under $Q$ the compensator of $\mu^R$ is a Poisson random measure with intensity measure $\eta^Q(dz)$ independent of $Y$ and $\nu$. This measure is needed in the analysis of the filtering problem of the trader in Section 3. Note that the equivalence of $\eta^P$ and $\eta^Q$ implies that for all $(t,e,\nu) \in [0,T] \times \mathcal{E} \times [0,\infty)$ the support of $\eta^P(t,e,\nu, dz)$ is equal to $\text{supp}(\eta)$. The assumption that $\text{supp}(\eta)$ is compact is not restrictive, since in reality the bid price moves only by a few ticks at a time.

The following examples serve to illustrate our framework; they will be taken up in our numerical experiments in Section 3.

**Example 2.2.** Consider the case where the return process $R$ follows a bivariate point process, i.e. there are two possible jump sizes, $\Delta R \in \{-\theta, \theta\}$ for some $\theta > 0$. In this example we assume that the dynamics of $S$ is independent of $Y$ and $t$. Moreover, the intensity $\lambda^+$ of an upward jump is constant and equal to $c_{\text{up}} > 0$, and the intensity $\lambda^-$ of a downward jump depends on the rate of trading and is given by $\lambda^-(\nu) = c_{\text{down}}(1 + a\nu)$ for constants $c_{\text{down}}, a > 0$. Note that, with this choice of $\lambda^-$, the intensity of a downward jump in $S$ is linearly increasing in the liquidation rate $\nu$. The function $\eta^P$ from (2.3) is independent of $t$ and $e$ and linearly decreasing in $\nu$; it is given by $\eta^P(\nu) = \theta(c_{\text{up}} - c_{\text{down}}(1 + a\nu))$. Linear models for the permanent price impact are frequently considered in the literature as they have theoretical and empirical advantages; see for instance Almgren et al. [4] or Gatheral and Schied [39].

**Example 2.3.** Now we generalize Example 2.2 and allow $\eta^P$ to depend on the state process $Y$. We consider a two-state Markov chain $Y$ with the state space $\mathcal{E} = \{e_1, e_2\}$ and we assume that $e_1$ is a ‘good’ state and $e_2$ a ‘bad’ state in the following sense: in state $e_1$ the intensity of an upward move of the stock is larger than in state $e_2$; the intensity of a downward move on the other hand is larger in state $e_2$ than in $e_1$. State $e_1$ could thus represent a scenario where another trader is executing a buy program that pushes prices up; state $e_2$ might correspond to the situation where another trader is executing a large sell program. Given constants $c_{1\text{up}} > c_{2\text{up}} > 0$, $c_{1\text{down}} > c_{2\text{down}} > 0$ and $a > 0$ we set

$$\lambda^+(e_i, \nu) = (c_{1\text{up}}, c_{2\text{up}})e_i \quad \text{and} \quad \lambda^-(e_i, \nu) = (1 + a\nu)(c_{1\text{down}}, c_{2\text{down}})e_i,$$

for $i = 1, 2$. Then, $\eta^P(e_i, \nu, dz) = \lambda^+(e_i, \nu)\delta_{\{\theta\}}(dz) + \lambda^-(e_i, \nu)\delta_{\{-\theta\}}(dz)$. Note that in this model we again have a linear permanent price impact.

### 2.2 The optimization problem
In this section we specify the ingredients of the traders optimization problem in detail.
**Liquidation strategies.** We assume that the state process $Y$ is not directly observable by the trader. Instead, she observes the price process $S$ and knows the model parameters, so that the information available to her is carried by the filtration $\mathbb{F}^S$ or, equivalently, by the filtration generated by the return process $R$. Hence we assume that the trader uses only liquidation strategies that are $\mathbb{F}^S$-adapted. Moreover we impose a bound on the maximal speed of trading: we fix some constant $\nu^\max > w_0/T$ and we call a liquidation strategy $\nu$ admissible if $\nu$ is $\mathbb{F}^S$-adapted and if $\nu_t \in [0, \nu^\max]$ for all $t \in [0, T]$ $\mathbb{P}$ a.s. Note that the condition $\nu^\max > w_0/T$ ensures that it is feasible for the trader to liquidate the whole inventory over the period $[0, T]$.

The assumption of a bounded liquidation rate merits a discussion. From a mathematical point of view a bound on the liquidation rate facilitates the application of results for the control of piecewise deterministic Markov processes, since in this theory it is typically assumed that the strategies take values in a compact control space. Moreover, without this assumption the viscosity solution characterization of the value function (Theorem 5.3 below) does not hold; a counterexample is given in Section 5.2 There we show that for unbounded liquidation rate the value function turns out to be a strict supersolution of the corresponding dynamic programming equation. Finally, the upper bound on $\nu_t$ ensures that under Assumption 2.1 for every admissible strategy $\nu$ a return process $R$ with compensating measure $\eta^R(t, Y_t-, \nu_{t-}, dz)dt$ (and hence the bid price process (2.2)) exists.

From a financial point of view an upper bound on $\nu_t$ is reasonable, as trading at infinite speed would correspond to large block transactions; allowing such transactions at some time point $t < T$ would require an explicit model for market resiliency. It is however not clear how to determine $\nu^\max$ empirically. In Proposition 2.3 we therefore show that $J^{*, m}$, the optimal proceeds from liquidation if the trader uses $\mathbb{F}^S$ adapted strategies with $\nu_t \leq m$ for all $t$, is bounded independently of $m$. Now the sequence $\{J^{*, m}\}$ is obviously increasing, since a higher $m$ means that the trader can optimize over a larger set of strategies. Hence $\{J^{*, m}\}$ is Cauchy, and the optimal proceeds from liquidation are essentially independent of the precise numerical value chosen for $\nu^\max$, provided that the upper bound is sufficiently large. This is confirmed by our numerical experiments, see Table 2.

**Objective of the trader.** To account for the case where not all shares have been sold prior to time $T$ we specify the liquidation value of the remaining share position $W_T$. This liquidation value is of the form $h(W_T)S_T$ for some increasing, continuous and concave function $h$ with $h(w) \leq w$ and $h(0) = 0$. We give some examples for $h$. The choice $h(w) \equiv 0$ corresponds to the case where liquidation at the terminal date is not possible; the choice $h(w) = \frac{w}{1-\vartheta w}$ for some $\vartheta > 0$ models a situation where the liquidation value of the remaining shares is positive but strictly smaller than the book value, reflecting the limited liquidity of the market for the stock.

We model the temporary price impact by a nonnegative and increasing function $f$ so that the proceeds from liquidation are given by $\int_0^T \nu_s S_s(1 - f(\nu_s))ds$. For instance, Almgren et al. propose a power function of the form $f(\nu) = c_f \nu^{\varsigma}$ and they estimate the coefficient $\varsigma \approx 0.6$.

Now we define the time $\tau$ as the minimum of the first time the inventory is completely liquidated and the horizon $T$:

$$\tau := \inf \{ t \geq 0 : W_t \leq 0 \} \wedge T. \quad (2.8)$$

\[1\] Note that when considering a block transaction at the horizon date $T$ we do not need to model market resiliency or permanent price impact as the model ‘ends’ at $T$. 

6
Denote by $\rho$ the (subjective) discount rate of the trader. Consider an admissible strategy $\nu$ and denote the corresponding bid price by $S^\nu$. The expected discounted value of the proceeds generated by the liquidation strategy $\nu$ is equal to

$$J(\nu) = \mathbb{E} \left( \int_0^T e^{-\rho u} \nu_u S^\nu_u (1 - f(\nu_u)) du + e^{-\rho \tau} S^\nu_\tau h(W_\tau) \right).$$

(2.9)

The trader wants to maximize (2.9) over all admissible strategies; the corresponding optimal value is denoted by $J^*$, or, if we want to emphasize the dependence on the upper bound on the liquidation rate, by $J^{*,\text{max}}$. Note that the form of the objective function in (2.9) implies that the trader is risk neutral. Risk neutrality is frequently assumed in the literature on optimal order execution, see for instance Bertsimas and Lo [16]. Moreover, in our setup the risk the trader may take is limited as we consider pure selling strategies and as the time period $[0, T]$ is fairly short.

The next proposition establishes an upper bound on $J^*$.

**Proposition 2.4.** Suppose that Assumption 2.1 holds and that the function $\eta^P(t, e, \nu)$ from (2.4) is decreasing in $\nu$, and set

$$\eta = 0 \vee \sup \{\eta^P(t, e, 0) - \rho: t \in [0, T], e \in \mathcal{E}\}.$$

Then $\sup_{m > 0} J^{*,m} \leq w_0 S_0 e^{\rho T}$.

Proof. Fix some $\mathbb{F}^S$-adapted strategy $\nu$ with values in $[0, m]$ and let $\tilde{S}_\tau^\nu = e^{-\rho \tau} S^\nu_\tau$. Since $W_t = w_0 - \int_0^t \nu_s ds$ we get by partial integration that

$$\int_0^T \nu_s \tilde{S}_s^\nu ds = - \int_0^T \tilde{S}_s^\nu dW_s = S_0 w_0 - \tilde{S}_\tau^\nu W_\tau + \int_0^T W_s d\tilde{S}_s^\nu.$$

Since $h(w) \leq w$ and $f(\nu) \geq 0$ we thus get that

$$\int_0^T \nu_s \tilde{S}_s^\nu (1 - f(\nu_s)) du + \tilde{S}_\tau^\nu h(W_\tau) \leq \int_0^T \nu_s \tilde{S}_s^\nu du + \tilde{S}_\tau^\nu W_\tau = S_0 w_0 + \int_0^T W_s d\tilde{S}_s^\nu.$$

Now $\int_0^T W_s d\tilde{S}_s^\nu = \int_0^T W_s \tilde{S}_s^\nu dM_s^R + \int_0^T W_u \tilde{S}_u^\nu (\eta^P(u, Y_u, \nu_u) - \rho) du$. Moreover, $\int_0^{\tau \wedge T} W_u \tilde{S}_u^\nu dM_u^R$ is a true martingale: as $0 \leq W_u \leq w_0$, a similar argument as in the proof of Lemma A.1 shows that this process is of integrable quadratic variation. Since $\eta^P(u, Y_u, \nu_u) - \rho \leq \eta$, $\tau \leq T$ and $W_u \leq w_0$, we get

$$J(\nu) \leq S_0 w_0 + \mathbb{E} \left( \int_0^T W_u \tilde{S}_u^\nu (\eta^P(u, Y_u, \nu_u) - \rho) du \right) \leq S_0 w_0 + \mathbb{E} \left( \int_0^T w_0 \eta^P du \right).$$

(2.10)

Next we show that $\mathbb{E}(\tilde{S}_\tau^\nu) \leq S_0 e^{\rho T}$. To this, note that by Lemma A.1 $\int_0^T S_s^\nu dM_s^R$ is a true martingale so that

$$\mathbb{E}(\tilde{S}_\tau^\nu) = S_0 + \mathbb{E}(\int_0^\tau \tilde{S}_s^\nu (\eta^P(u, Y_u, \nu_u) - \rho) du) \leq S_0 + \eta \int_0^\tau \mathbb{E}(\tilde{S}_u^\nu) du,$$

and the claim follows from the Gronwall inequality. Using (2.10) we finally get that $J(\nu) \leq S_0 w_0 (1 + \int_0^\tau \eta^P du) = S_0 w_0 e^{\rho T}$, and hence the result. $\square$

7
3 Partial Information and Filtering

Considering $\mathbb{F}^S$-adapted investment strategies results in an optimal control problem under partial information. The standard approach to dealing with such problems is to introduce the filter for the Markov chain as additional state variable of the control problem. In this section we therefore derive the filtering equations for our model. Filtering for point processes observation is for instance considered in Ceci and Gerardi [24], Elliott and Malcolm [34], Frey and Schmidt [38], Ceci and Colaneri [22, 23]. This literature is mostly based on the innovations approach. In this paper, instead, we address the filtering problem via the reference probability approach. This methodology relies on the existence of an equivalent probability measure such that the observation process is driven by a random measure with dual predictable projection independent of the Markov chain, see for instance Brémaud [18, Chapter 6]. The reference probability approach permits us to overcome the difficulties caused by the fact that the observation process $S$ is affected by the liquidation strategy chosen by the trader.

3.1 Reference probability. We start from a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{Q})$ that supports a Markov chain $Y$ with state space $\mathcal{E}$ and generator matrix $Q$, and an independent Poisson random measure $\mu^R$ with compensator $\eta^Q(dz)dt$ as in Assumption 2.1.2; $\mathbb{Q}$ is known as the reference probability. Note that under $\mathbb{Q}$, the dynamics of $S$ and $R$ are independent of the liquidation strategy $\nu$ so that $\mathbb{F}^S$ is exogenously given. Moreover, due to the independence of $R$ and $Y$, the $(\mathbb{Q}, \mathbb{F})$-dual predictable projection of the jump measure $\mu^{(Y,R)}$ of the pair $(Y, R)$ has the form

$$\tilde{\eta}^Q(d\gamma, dz)dt := \left\{ \sum_{i=1}^{K} \mathbf{1}_{\{Y_{t-} = e_i\}} \left( \eta^Q(dz) + \sum_{j \neq i} q^{ij} \delta_{\{e_j - e_i\}}(d\gamma) \right) \right\} dt .$$

Fix an admissible liquidation strategy $\nu$ and define the function $\beta$ by

$$\beta(t, e, \nu, z) := \frac{d\eta^P(t, e, \nu)}{d\eta^Q}(z) - 1 , \quad (3.1)$$

i.e. $\beta(t, e, \nu, z) + 1$ is the Radon-Nikodym derivative of the measure $\eta^P(t, e, \nu, dz)$ with respect to $\eta^Q(dz)$. Define for $t \leq T$ the stochastic exponential $\tilde{Z}$ by

$$\tilde{Z}_t = 1 + \int_0^t \int_{\mathbb{R}} \tilde{Z}_{s-} \beta(s, Y_{s-}, \nu_{s-}, z) \left( \mu^R(ds, dz) - \eta^Q(dz)ds \right) .$$

Then we have the following result.

**Lemma 3.1.** Let Assumption 2.1 prevail. Then the process $\tilde{Z}$ is a strictly positive martingale with $\mathbb{E}^Q(\tilde{Z}_T) = 1$. Define a measure $\mathbb{P}$ on $\mathcal{F}_T$ by setting $\frac{d\mathbb{P}}{d\mathbb{Q}}|_{\mathcal{F}_T} = \tilde{Z}_T$. Then $\mathbb{P}$ and $\mathbb{Q}$ are equivalent and, under $\mathbb{P}$, the random measure $\mu^{(Y,R)}$ has the compensator $\tilde{\eta}^P(dt, d\gamma, dz)$ given in (2.3). Conversely, given a pair $(Y, R)$ with compensating random measure $\tilde{\eta}^P$ as in (2.3), the reference probability $\mathbb{Q}$ can be constructed by the inverse change of measure.

The proof of the lemma is postponed to Appendix A. Note that Lemma 3.1 ensures, for every admissible $\nu$, the existence of a pair $(Y, R)$ with compensating measure $\eta^P(t, Y_{t-}, \nu_{t-}, d\gamma, dz)dt$, that is unique in law.
3.2 Filtering equations. For a function $f: \mathcal{E} \to \mathbb{R}$, we introduce the filter $\pi(f)$ as the optimal projection of the process $f(Y)$ on the filtration $\mathbb{F}^S$, i.e. $\pi(f)$ is a càdlàg process such that for all $t \in [0, T]$, it holds that $\pi_t(f) = \mathbb{E}\left(f(Y_t) \mid \mathcal{F}^S_t\right)$. Note that $f(Y_t) = \langle f, Y_t \rangle$, where $\langle \cdot, \cdot \rangle$ denotes the scalar product on $\mathbb{R}^K$ and $f_k = f(e_k)$, $1 \leq k \leq K$, so that functions of the Markov chain can be identified with $K$-vectors. Let $\pi^f_t := \mathbb{E}\left(1_{\{Y_t = e_i\}} \mid \mathcal{F}^S_t\right)$. Then, we can represent the filter as

$$
\pi_t(f) = \sum_{j=1}^K f_j \pi_t^j = \langle f, \pi_t \rangle, \quad t \in [0, T].
$$

In this section we derive the dynamics of the process $\pi = (\pi^1_t, \ldots, \pi^K_t)_{0 \leq t \leq T}$.

By the Kallianpur-Striebel formula we have $\pi_t(f) := \frac{p_t(f)}{p_t(1)}$, for every $t \in [0, T]$, where $p(f)$ denotes the unnormalized version of the filter defined by

$$
p_t(f) := \mathbb{E}^Q \left( \tilde{Z}_t(f, Y_t) \mid \mathcal{F}^S_t \right), \quad t \in [0, T].
$$

The next theorem gives the dynamics of $p(f)$.

**Theorem 3.2 (The Zakai equation).** Suppose Assumption 2.1 holds and let $f: \mathcal{E} \to \mathbb{R}$. Then, for every $t \in [0, T]$, the unnormalized filter (3.2) solves the equation:

$$
p_t(f) = \pi_0(f) + \int_0^t p_s(Qf)ds + \int_0^t \int_\mathbb{R} p_s^- (\beta(z)f) (\mu^R(ds, dz) - \eta^Q_s(dz)ds),
$$

where $p_t^- (\beta(z)f) := \mathbb{E}^Q \left( f(Y_{t^-})\tilde{Z}_{t^-} \beta(t, Y_{t^-}, \nu_{t^-}, z) \mid \mathcal{F}^S_t \right)$ and $p_t(Qf) := \mathbb{E}^Q \left( \tilde{Z}_t(Qf, Y_t) \mid \mathcal{F}^S_t \right)$.

**Proof.** We extend the proof for the classical case of diffusion information (see for instance Bain and Crisan [4], Theorem 3.24) to the situation where the observation process is a marked point process. Recall that for a function $f: \mathcal{E} \to \mathbb{R}$ the semimartingale decomposition of $f(Y_t)$ is given by $f(Y_t) = f(Y_0) + \int_0^t (Qf, Y_s)ds + M^f_t$, where $M^f$ is a true $(\mathcal{F}, Q)$-martingale. Define the process $\tilde{Z}^f = (\tilde{Z}^f_t)_{t \in [0, T]}$ by

$$
\tilde{Z}^f_t := \frac{\tilde{Z}_t}{1 + \epsilon \tilde{Z}_t},
$$

and note that $\tilde{Z}^f_t < 1/\epsilon$ for every $t \in [0, T]$, i.e. $\tilde{Z}^f_t$ is a bounded process. Itô’s formula gives that

$$
d\tilde{Z}^f_t = \tilde{Z}^f_t \int_\mathbb{R} \frac{\beta(t, Y_{t^-}, \nu_{t^-}, z)}{1 + \epsilon \tilde{Z}_{t^-}} \mu^R(dt, dz) - \tilde{Z}^f_t \int_\mathbb{R} \frac{\beta(t, Y_{t^-}, \nu_{t^-}, z)}{1 + \epsilon \tilde{Z}_{t^-}} \eta^Q(dz)dt.
$$

Now we compute $\tilde{Z}^f f(Y)$. Notice that $[\tilde{Z}^f, Y]_t = 0$ for every $t \in [0, T]$, as $R$ and $Y$ have no common jumps. Hence, from Itô’s product rule we get

$$
d(\tilde{Z}^f_t f(Y_t)) = \tilde{Z}^f_t (Qf, Y_t)dt + \tilde{Z}^f_t dM^f_t - f(Y_t)\tilde{Z}^f_t \int_\mathbb{R} \frac{\beta(t, Y_{t^-}, \nu_{t^-}, z)}{1 + \epsilon \tilde{Z}_{t^-}} \eta^Q(dz)dt + f(Y_t)\tilde{Z}^f_t \int_\mathbb{R} \frac{\beta(t, Y_{t^-}, \nu_{t^-}, z)}{1 + \epsilon \tilde{Z}_{t^-}} \mu^R(dt, dz).
$$

Next we show that $\mathbb{E}^Q \left( \int_0^t \tilde{Z}^f_s dM^f_s \mid \mathcal{F}^S_t \right) = 0$. By the definition of conditional expectation, this is equivalent to $\mathbb{E}^Q \left( H \int_0^t \tilde{Z}^f_s dM^f_s \right) = 0$ for every bounded, $\mathcal{F}^S_t$-measurable random variable $H$. Define an $(\mathcal{F}^S, Q)$-martingale by $H_u := \mathbb{E}^Q \left( H \mid \mathcal{F}^S_u \right)$, $0 \leq u \leq t \leq T$, and note that $H = H_t$. By
the martingale representation theorem for random measures, see, e.g., Jacod and Shiryaev [43] Ch. III, Theorem 4.37 or Brémaud [18] Ch. VIII, Theorem T8, we get that there is a bounded \(F^S\)-predictable random function \(\phi\) such that

\[
H_t = H_0 + \int_0^t \int_{\mathbb{R}} \phi(s, z)(\mu^R(ds, dz) - \eta^Q(dz)ds), \quad t \in [0, T].
\]

Now, applying the Itô product rule and using that \([M^f, H]_t = [Y, R]_t = 0\) for every \(t \in [0, T]\), we obtain

\[
H_t \int_0^t \tilde{Z}^c_s \ dM^f_s = \int_0^t H_s \tilde{Z}^c_s \ dM^f_s + \int_0^t \left( \int_{0}^{s} \tilde{Z}^c_u \ dM^f_u \right) \phi(s, z)(\mu^R(ds, dz) - \eta^Q(dz)ds).
\]

Both integrals on the right hand side of the above representation are martingales. This follows from the finite-state property of the Markov chain \(Y\) and the boundedness of \(\tilde{Z}^c\) and \(H\). Hence, taking the expectation we get that \(\mathbb{E}^Q \left( H_0 \int_0^t \tilde{Z}^c_s \ dM^f_s \right) = 0\) as claimed.

Taking the conditional expectation from (3.3) and applying Lemma A.2 and the Fubini theorem we get for every \(t \in [0, T]\),

\[
\begin{align*}
\mathbb{E}^Q \left( \tilde{Z}^c_t f(Y_t) \mid F^S_t \right) &= \frac{\pi_0(f)}{1 + \epsilon} + \int_0^t \mathbb{E}^Q \left( \tilde{Z}^c_s - \langle Qf, Y_s \rangle \mid F^S \right) ds \\
&+ \int_0^t \int_{\mathbb{R}} \mathbb{E}^Q \left( f(Y_s) \tilde{Z}^c_s \frac{\beta(s, Y_{s-}, \nu_{s-}, z)}{1 + \epsilon \tilde{Z}^c_s (1 + \beta(s, Y_{s-}, \nu_s))} \mid F^S \right) \mu^R(ds, dz) \\
&- \int_0^t \int_{\mathbb{R}} \mathbb{E}^Q \left( f(Y_s) \tilde{Z}^c_s \frac{\beta(s, Y_{s-}, \nu_{s-}, z)}{1 + \epsilon \tilde{Z}^c_s} \mid F^S \right) \eta^Q(dz)ds.
\end{align*}
\]

Note that, for every \(t \in [0, T]\), \(\tilde{Z}^c_t < \tilde{Z}_t\) and that \(\tilde{Z}_t\) is integrable. Since \(\beta\) is bounded by assumption, by dominated convergence we get the following three limits

\[
\begin{align*}
\lim_{\epsilon \to 0} \mathbb{E}^Q \left( \tilde{Z}^c_t f(Y_t) \mid F^S_t \right) &= \mathbb{E}^Q \left( \tilde{Z}_t f(Y_t) \mid F^S_t \right), \\
\lim_{\epsilon \to 0} \int_0^t \mathbb{E}^Q \left( \tilde{Z}^c_s - \langle Qf, Y_s \rangle \mid F^S \right) ds &= \int_0^t \mathbb{E}^Q \left( \tilde{Z}^c_s - \langle Qf, Y_s \rangle \mid F^S \right) ds, \\
\lim_{\epsilon \to 0} \int_0^t \int_{\mathbb{R}} \mathbb{E}^Q \left( f(Y_s) \tilde{Z}^c_s \frac{\beta(s, Y_{s-}, \nu_{s-}, z)}{1 + \epsilon \tilde{Z}^c_s (1 + \beta(s, Y_{s-}, \nu_s))} \mid F^S \right) \eta^Q(dz)ds \\
&= \int_{\mathbb{R}} \mathbb{E}^Q \left( f(Y_s) \tilde{Z}^c_s \frac{\beta(s, Y_{s-}, \nu_{s-}, z)}{1 + \epsilon \tilde{Z}^c_s (1 + \beta(s, Y_{s-}, \nu_s))} \mid F^S \right) \eta^Q(dz)ds.
\end{align*}
\]

Finally we consider the integral with respect to \(\mu^R(ds, dz)\) in (3.4). Let \(\{T_n, Z_n\}\) be the sequence of jump times and the corresponding jump sizes of the process \(R\). Denote by \(n(t)\) the number of jumps up to time \(t\), so that \(T_{n(t)}\) is the last jump time before \(t\). Then

\[
\lim_{\epsilon \to 0} \int_0^t \int_{\mathbb{R}} \mathbb{E}^Q \left( f(Y_s) \tilde{Z}^c_s \frac{\beta(s, Y_{s-}, \nu, z)}{1 + \epsilon \tilde{Z}^c_s (1 + \beta(s, Y_{s-}, \nu, z))} \mid F^S \right) \mu^R(ds, dz)
\]
for every $t$

The process $Q$ solves the following SDE system

$$
\pi_t = \pi_0 + \int_0^t \sum_{j=1}^K q^j \pi_s^j ds + \int_0^t \int_\mathbb{R} \pi_s u(s, \nu_s, \pi_s, z) (\mu^R(ds, dz) - \pi_s(\eta^P(z)dz)) ds,
$$

(3.5)

for every $t \in [0, T]$, where $u^i(t, \nu, \pi, z) := \frac{d\eta^P(t, e_i, \nu)}{d\eta^Q(z)} - 1$.

Proof. We set $f(Y_t) = 1_{\{Y_t = e_i\}}$ in Proposition 3.3 and obtain the result. 

We introduce the notation

$$
\pi_{t-}(\eta^P(dz)) := \sum_{i=1}^K \pi_{t-}^i \eta^P(t, e_i, \nu_t, dz), \quad t \in [0, T].
$$

By applying Brémaud [18, Ch. II, Theorem T14] it is easy to see that $\pi_{t-}(\eta^P(dz))dt$ provides the $(\mathbb{F}^S, \mathbb{P})$-dual predictable projection of the measure $\mu^R$.

The next proposition provides the filter dynamics.

**Proposition 3.3** (The Kushner-Stratonovich equation). The $(\mathbb{F}^S, \mathbb{P})$-semimartingale decomposition of the normalized filter $\pi(f)$ is given by

$$
\begin{align*}
&d\pi_t(f) = \pi_t(Qf) dt \\
&+ \int_\mathbb{R} \left( \frac{\pi_{t-}(f d\eta^P/d\eta^Q(z))}{\pi_{t-}(d\eta^P/d\eta^Q(z))} - \pi_{t-}(f) \right) (\mu^R(dt, dz) - \pi_{t-}(\eta^P(dz))) dt,
\end{align*}
$$

where $\pi_t(Qf) = \mathbb{E}(\langle Qf, Y_t \rangle | \mathbb{F}^S)$.

Note that the filtering equation does not depend on the particular choice of the reference probability $Q$. The proof is provided in Appendix A.
Example 2.2 revisited  In the following we now give the dynamics of $\pi$ for Example 2.3. In this case it is sufficient to specify the dynamics of $\pi^1$, since $\pi^2 = 1 - \pi^1$. Define two point processes by $N^\text{up}_t = \sum_{T_n \leq t} \mathbb{1}_{\{\Delta R_{T_n} = \theta\}}$ and $N^\text{down}_t = \sum_{T_n \leq t} \mathbb{1}_{\{\Delta R_{T_n} = -\theta\}}$, $t \in [0, T]$, that count the upward and the downward jumps of the return process, respectively. Note that the function $u^1$ is given by

\[
 u^1(\nu, \pi^1, z) = \frac{\pi^1 \lambda^+(e_1, \nu)}{\pi^1 \lambda^+(e_1, \nu) + (1 - \pi^1) \lambda^+(e_2, \nu)} \mathbb{1}_{\{z = \theta\}} + \frac{\pi^1 \lambda^-(e_1, \nu)}{\pi^1 \lambda^-(e_1, \nu) + (1 - \pi^1) \lambda^-(e_2, \nu)} \mathbb{1}_{\{z = -\theta\}}.
\]

By Corollary 3.4 we then get the following equation for $\pi^1_t$

\[
d\pi^1_t = \left(q^1 \pi^1_t + q^2 (1 - \pi^1_t)\right) dt + \pi^1_t (1 - \pi^1_t) \left((\lambda^+(e_1, \nu_1) + \lambda^-(e_1, \nu_1)) - (\lambda^+(e_2, \nu_1) + \lambda^-(e_2, \nu_1))\right) dt + \left(\frac{\pi^1_t \lambda^+(e_1, \nu_1)}{\pi^1_t \lambda^+(e_1, \nu_1) + (1 - \pi^1_t) \lambda^+(e_2, \nu_1) - 1}\right) dN^\text{up}_t + \left(\frac{\pi^1_t \lambda^-(e_1, \nu_1)}{\pi^1_t \lambda^-(e_1, \nu_1) + (1 - \pi^1_t) \lambda^-(e_2, \nu_1) - 1}\right) dN^\text{down}_t.
\]

4 Control Problem I: Analysis via PDMPs

We begin with a brief overview of our analysis of the control problem (2.9). In Proposition 4.2 below we show that the Kushner Stratonovich equation (3.5) has a unique solution. Then, standard arguments (see for instance Bäuerle and Rieder [10]) ensure that the original control problem under incomplete information is equivalent to a control problem under complete information with state process equal to the $(K + 2)$-dimensional process $X := (W, S, \pi)$. This process is a PDMP in the sense of Davis [30]. A trajectory of a PDMP consists of a deterministic part, which is a solution of an ordinary differential equation (ODE), interspersed by random jumps. We apply control theory for PDMPs to the optimal liquidation problem. This theory is based on the observation that a control problem for a PDMP is discrete in time: loosely speaking, at every jump-time $T_n$ of the process one chooses a control policy to be followed up to the next jump time or until maturity. Therefore, one can identify the control problem for the PDMP with a control problem for a discrete-time, infinite-horizon Markov decision model (MDM). Using this connection we show that the value function of the optimal liquidation problem is continuous and the unique solution of the dynamic programming or optimality equation for the MDM, and we establish the existence of an optimal strategy in the set of all relaxed controls. These results are the basis for the viscosity-solution characterization of the value function in Section 5.

There are several contributions where optimal control for PDMPs is studied via MDMs, see for instance Davis [30], Lenhart and Liao [46], Costa and Davis [26], Dempster and Ye [32], Almudevar [3], Forwick et al. [36], Bäuerle and Rieder [12], Costa and Dufour [27].

4.1 Optimal liquidation as a control problem for a PDMP  From the viewpoint of the trader endowed with the filtration $\mathbb{F}^S$, the state of the economic system at time $t \in [0, T]$ is given by $X_t = (W_t, S_t, \pi_t)$. To make the piecewise deterministic process Markovian it is convenient to include time into the state; we set $\bar{X}_t = (t, X_t)$. In this section we show that this process is
a PDMP in the sense of Davis [30] and we rewrite the original control problem (2.9) as control problem for a PDMP.

The corresponding state space is \( \tilde{X} = [0, T] \times X \) where \( X = [0, u_0] \times \mathbb{R}^+ \times S^K \) and \( S^K \) is the \( K \)-dimensional simplex. Let \( \nu \) be the liquidation strategy followed by the trader. It follows from (2.1), (3.5), and from the fact that the bid price is a pure jump process that between jump times the state process follows the ODE \( d\tilde{X}_t = g(\tilde{X}_t, \nu_t)dt \), where the vector field \( g(\tilde{x}, \nu) \in \mathbb{R}^{K+3} \) is given by \( g^1(\tilde{x}, \nu) = 1 \), \( g^2(\tilde{x}, \nu) = -\nu \), \( g^3(\tilde{x}, \nu) = 0 \), and for \( k = 1, \ldots, K \),

\[
g^{k+3}(\tilde{x}, \nu) = \sum_{j=1}^{K} g^{k} \pi^{j} - \pi^{k} \sum_{j=1}^{K} \int_{\mathbb{R}} u^{k}(t, \nu, \pi, z) \eta^{P}(t, e_j, \nu, dz). \quad (4.1)
\]

For our analysis we need the following regularity property of \( g \).

**Lemma 4.1.** Under Assumption 2.1 the function \( g \) is Lipschitz continuous in \( \tilde{x} \) uniformly in \( t \) and \( \nu \).

The proof is postponed to Appendix B.2.

The jump rate of the state process \( \tilde{X} \) is given by \( \lambda(\tilde{X}_{t^-}, \nu_{t^-}) \), where

\[
\lambda(\tilde{x}, \nu) = \lambda(t, w, s, \pi, \nu) := \sum_{j=1}^{K} \pi^{j} \eta^{P}(t, e_j, \nu, \mathbb{R}).
\]

Note that \( \lambda(\tilde{x}, \nu) \) is independent of \( w \) and \( s \). Next, we identify the transition kernel \( Q_{\tilde{X}} \) that governs the jumps of the state process. Denote by \( Q_{\tilde{X}}(B \mid \tilde{x}, \nu) \) the probability of jumping to a point in the set \( B \subseteq \tilde{X} \) for a given state \( \tilde{x} \in \tilde{X} \) and a given control \( \nu \in [0, \nu_{\text{max}}] \). It follows that for any bounded and measurable function \( f : \tilde{X} \to \mathbb{R}^+ \),

\[
Q_{\tilde{X}} f(\tilde{x}, \nu) := \int_{\tilde{X}} f(\tilde{y}) Q_{\tilde{X}}(d\tilde{y} \mid \tilde{x}, \nu) = \frac{1}{\lambda(\tilde{x}, \nu)} Q_{\tilde{X}} f(\tilde{x}, \nu), \quad (4.2)
\]

where the unnormalized kernel \( Q_{\tilde{X}} \) is given by

\[
Q_{\tilde{X}} f(\tilde{x}, \nu) = \sum_{j=1}^{K} \pi^{j} \int_{\mathbb{R}} f(t, w, s(1+z), \pi^{j}(1 + u^{j}(t, \nu, \pi, z)), \ldots, \\
\pi^{K}(1 + u^{K}(t, \nu, \pi, z))) \eta^{P}(t, e_j, \nu, dz).
\]

The state process of the optimal liquidation problem is a PDMP with characteristics given by the vector field \( g \), the jump rate \( \lambda \) and the transition kernel \( Q_{\tilde{X}} \). It is standard in control theory for PDMPs to work with so-called open-loop controls. In the current context this means that the trader chooses at each jump time \( T_n < \tau \) a liquidation policy \( \nu^n \) to be followed up to \( T_{n+1} \land \tau \). This policy may depend on the state \( \tilde{X}_{T_n} = (T_n, X_{T_n}) \) and is therefore denoted by \( \nu^n(\tilde{x}) \).

Denote by \( \mathcal{A} \) the set of measurable mappings \( \alpha : [0, T] \to [0, \nu_{\text{max}}] \). We define an admissible open loop liquidation strategy as a sequence of mappings \( \{\nu^n\} : \tilde{X} \to \mathcal{A} \); the liquidation rate at time \( t \) is given by

\[
\nu_t = \sum_{n=0}^{\infty} 1_{(T_n \land \tau, T_{n+1} \land \tau]}(t) \nu^n(t - T_n, \tilde{X}_{T_n}). \quad (4.3)
\]
It follows from Brémaud [13, Theorem T34, Appendix A2] that an admissible strategy as defined
in Section 2.2 has the form (4.3) for \( F^\infty_T \) measurable mappings \( \nu^n : \Omega \to A \), where \( \nu^n(\cdot) \) may
depend on the entire history of the system. General results for Markov decision models (see
Bäuerle and Rieder [13, Theorem 2.2.3]) show that the expected profit of the trader stays the
same if instead we consider the smaller class of admissible open-loop strategies, so that from now
on it is implicitly understood that an admissible strategy is automatically an open-loop
strategy of the form (4.3).

**Proposition 4.2.** Let Assumption 2.1 hold. For every admissible liquidation strategy \( \{\nu^n\} \) and
every initial value \( \tilde{x} \), a unique PDMP with characteristics \( g, \lambda, \) and \( Q_\tilde{x} \) as above exists. In
particular the Kushner-Stratonovic equation (4.5) has a unique solution.

**Proof.** Lemma 4.1 implies that for \( \alpha \in A \) the ODE \( d\tilde{X}_t = g(\tilde{X}_t, \alpha_t)dt \) has a unique solution
so that between jumps the state process is well-defined. At any jump time \( T_n, \tilde{X}_{T_n} \) is uniquely
defined in terms of observable data \( (T_n, \Delta R_{T_n}) \). Moreover, since the jump intensity is bounded
by \( \lambda^{\text{max}} \), jump times cannot accumulate. \( \square \)

Denote by \( P_{\{\nu^n\}} \) (equiv. \( P_{\tilde{x}} \)) the law of the state process provided that \( X_t = x \in X \) and that
the trader uses the open-loop strategy \( \{\nu^n\}_{n \in \mathbb{N}} \). The reward function associated to an admissible
liquidation strategy \( \{\nu^n\} \) is

\[
V(t, x, \{\nu^n\}) = E_{\{\nu^n\}}^{\{\tilde{x}, t, x\}}(\int^\tau \rho(\tilde{u}_t)\nu_u S_u(1 - f(\nu_u))du + e^{\rho(\tau - t)}h(W_\tau)S_\tau),
\]

and the value function of the liquidation problem under partial information is

\[
V(t, x) = \sup \{V(t, x, \{\nu^n\}) : \{\nu^n\} \text{ admissible liquidation strategy}\}. \quad (4.4)
\]

**Remark 4.3.** Note that the compensator \( \eta^P \) and the dynamics of the filter \( \pi \) are independent of
the current bid price \( s \), and that the payoff of a liquidation strategy \( \{\nu^n\} \) is positively homogeneous in \( s \). This implies that the reward and the value function of the liquidation problem are positively homogeneous in \( s \) and in particular \( V(t, w, s, \pi) = sV(t, w, 1, \pi) \).

### 4.2 Associated Markov Decision Model

The optimization problem in (4.4) is essentially discrete: the control policy is chosen at the discrete time points \( T_n \wedge \tau, n \geq 0 \), and, given an
admissible control policy, the sequence \( \{\tilde{X}_{T_n \wedge \tau}\} \) is Markov. Hence (4.4) can be rewritten as
a control problem in an infinite horizon MDM. In this section we identify this MDM. A short
introduction to MDMs is given in Appendix B.1.

**Basic definitions.** Let \( \Delta \) be a cemetery state. Consider the sequence \( \{L_n\} \) of random variables
defined by

\[
L_n = \tilde{X}_{T_n} \quad \text{for} \quad T_n < \tau \quad \text{and} \quad L_n = \Delta \quad \text{for} \quad T_n \geq \tau, \quad n \in \mathbb{N}.
\]

In order to derive the transition kernel of the sequence \( \{L_n\} \) and the corresponding reward function of the MDM, we introduce some notation. For a function \( \alpha \in \mathcal{A} \) we denote by \( \phi^\alpha(\tilde{x}) \) or by \( \phi(\alpha, \tilde{x}) \) the flow of the initial value problem \( \frac{d}{ds} \tilde{X}(s) = g(\tilde{X}(s), \alpha_s) \) with initial condition
\( \tilde{X}(0) = \tilde{x} \). Whenever we want to make the dependence on time explicit we write \( \tilde{\varphi}^\alpha \) in the form \((t, \varphi^\alpha)\). Next we define the functions

\[
\lambda^\alpha_s(\tilde{x}) = \lambda(\tilde{\varphi}^\alpha_s(\tilde{x}), \alpha_s) = \lambda((t + s, \varphi^\alpha_s), \alpha_s),
\]

\[
\Lambda^\alpha_s(\tilde{x}) = \Lambda^\alpha(s; \tilde{x}) := \int_0^s \lambda^\alpha_u(\tilde{x})du.
\] (4.5)

Finally we consider the boundary of the state space \( \tilde{X} \). Note that \( \pi_t \) belongs to the hyperplane \( \mathcal{H}^K = \{x \in \mathbb{R}^K : \sum_{i=1}^K x_i = 1\} \), so that \( \tilde{X} \) is contained in the set \( \mathcal{H} = \mathbb{R}^3 \times \mathcal{H}^K \) which is a hyperplane of \( \mathbb{R}^{K+3} \). When considering the boundary or the interior of \( \tilde{X} \) we always refer to the relative boundary or the relative interior with respect to \( \mathcal{H} \). Of particular interest to us is the active boundary \( \Gamma \) of the state space, that is the part of the boundary of \( \tilde{X} \) which can be reached by the flow \( \tilde{\varphi}^\alpha(t) \) starting in an interior point \( \tilde{x} \in \tilde{X} \). The boundary of the state space \( \tilde{X} \) can only be reached if \( W = 0 \), if \( t = T \), or if the filter process reaches the boundary of the \( K \)-dimensional simplex. The latter is not possible: indeed, if \( \tau_i^0 > 0 \), then \( \tau_i^0 > 0 \) for all \( t \in [0, T] \), since there is a positive probability that the Markov chain has not changed its state. Hence the active boundary equals \( \Gamma = \Gamma_1 \cup \Gamma_2 \), where

\[
\Gamma_1 = [0, T] \times \{0\} \times (0, \infty) \times S^K_0 \quad \text{and} \quad \Gamma_2 = \{T\} \times [0, w_0] \times (0, \infty) \times S^K_0,
\] (4.6)

and where \( S^K_0 \) is the interior of \( S^K \), i.e. \( S^K_0 := \{x \in \mathbb{R}^K_+ : \sum_{i=1}^K x_i = 1\} \). In \( \{4.6\} \), \( \Gamma_1 \) is the lateral part of the active boundary that corresponds to an inventory level equal to zero and \( \Gamma_2 \) is the terminal boundary related the exit from the state space at maturity \( T \).

In the sequel we denote the first exit time of the flow \( \tilde{\varphi}^\alpha(t) \) from \( \tilde{X} \) by

\[
\tau^\varphi = \tau^\varphi(\tilde{x}, \alpha) = \inf\{s \geq 0 : \tilde{\varphi}^\alpha_s(\tilde{x}) \in \Gamma\}.
\]

Notice that the stopping time \( \tau \) defined in (2.8) corresponds to the first time the state process \( \tilde{X} \) reaches the active boundary \( \Gamma \) and also that \( \tau^\varphi((t, x), \alpha) + t \) is equal to \( \tau \), if there is no jump in the time interval \([t, t + \tau^\varphi] \).

**Transition kernel.** To identify the transition kernel of the sequence \( \{L_n\} \), we denote by \( U_{n+1} = T_{n+1} - T_n \) the inter-arrival times of the jumps of \( R \). For a bounded and measurable function \( f : \tilde{X} \cup \{\tilde{x}\} \rightarrow \mathbb{R} \), it holds that

\[
\mathbb{E}(f(L_{n+1}) \mid L_n = (t, x), \nu^n = \alpha) = f(\tilde{\Delta})P(U_{n+1} + t > \tau \mid L_n = (t, x), \nu^n = \alpha) + \mathbb{E}(f(U_{n+1} + t, X_{U_{n+1} + t})1_{\{U_{n+1} + t \leq \tau\}} \mid L_n = (t, x), \nu^n = \alpha).
\]

Now note that \( \{U_{n+1} + t > \tau\} = \{U_{n+1} > \tau^\varphi\} \). Moreover, the distribution of \( U_{n+1} \) given \( L_n = (t, x) \) and \( \nu^n = \alpha \) is equal to \( \lambda^\alpha_t(\tilde{x})e^{-\Lambda^\alpha_{\nu^n}(\tilde{x})}d\mu \). Hence, we get

\[ P(U_{n+1} + t > \tau \mid L_n = (t, x), \nu^n = \alpha) = e^{-\lambda^\alpha_t(\tilde{x})e^{-\Lambda^\alpha_{\nu^n}(\tilde{x})}d\mu}. \]

Moreover, using the transition kernel of the state process given in (4.2), for \( u \leq \tau^\varphi \) we have

\[
\mathbb{E}(f(U_{n+1} + t, X_{U_{n+1} + t})1_{\{U_{n+1} \leq \tau^\varphi\}} \mid U_{n+1} = u, L_n = (t, x), \nu^n = \alpha) = Q_Xf(u + t, \varphi_u(\tilde{x}), \alpha_u).
\]

By integrating with respect to the conditional distribution of \( U_{n+1} \), we get that the transition kernel \( Q_L \) of \( \{L_n\} \) is given by

\[
Q_Lf((t, x), \alpha) = \int_0^{\tau^\varphi(\tilde{x})} e^{-\Lambda^\alpha(u)}Q_Xf(u + t, \varphi_u(\tilde{x}), \alpha_u)du + e^{-\Lambda^\alpha_{\nu^n}(\tilde{x})}f(\tilde{\Delta}).
\]

Moreover, since the cemetery state is absorbing, \( Q_L1_{\{\tilde{\Delta}\}}(\tilde{\Delta}, \alpha) = 1 \).
Reward function of the associated MDM. We now identify the optimal liquidation problem \([4.4]\) with a MDM for the sequence \(\{L_n\}\). For this, we define for \(\tilde{x}\) the one-period reward function \(r: \tilde{X} \times A \to \mathbb{R}^+\) by
\[
    r(\tilde{x}, \alpha) = \int_0^{\tau^*} e^{-\rho \tau} e^{-\gamma(\tilde{x})} \alpha_u s(1 - f(\alpha_u)) d\tau + e^{-\rho \tau} e^{-\gamma(\tilde{x})} h(w_{\tau^*}) s,
\]
where \(w_{\tau^*}\) represents the inventory-component of \(\tilde{\varphi}\), and we set \(r(\tilde{x}) = 0\). For an admissible strategy \(\nu^n\) we let \(J^{(\nu^n)}_\infty(\tilde{x}) = \mathbb{E}_{\tilde{x}}(\nu^n) \left( \sum_{n=0}^{\infty} r(L_n, \nu^n(L_n)) \right)\), and
\[
    J^{(\nu^n)}_\infty(\tilde{x}) := \sup \left\{ J^{(\nu)}_\infty(\tilde{x}) : \{\nu^n\} \text{ admissible liquidation strategy} \right\}.
\]
(4.7)

The next lemma shows that the MDM with transition kernel \(Q_L\) and one-period reward \(r(L, \alpha)\) is equivalent to the optimization problem \([4.4]\).

**Lemma 4.4.** For every admissible strategy \(\{\nu^n\}\) it holds that that \(V(\nu^n) = J^{(\nu^n)}_\infty\). Hence \(V = J_\infty\), and the control problems \([4.4]\) and \([4.7]\) are equivalent.

The proof is in Appendix B.2.

**4.3 The optimality equation.** Next we use Theorem [3.4] to derive the optimality equation for the value function \(J_\infty\) and hence for \(V\). For this we face a number of tasks. First, we need to introduce a topology on the control set \(A\) so that \(A\) is compact; second, we need to find a suitable bounding function; finally, we need to ensure that \(r\) and \(Q_L\) are continuous with respect to the topology on \(A\). We introduce the following continuity conditions on the data of our model:

**Assumption 4.5.** For all \(j = 1, \ldots, K\),

1. the measure \(\eta^j(t, \nu, dz)\) is continuous in the weak topology, i.e. for all bounded and continuous functions \(f\), the mapping \((t, \nu) \mapsto \int_{\mathbb{R}} f(z) \eta^j(dz)\) is continuous on \([0, T] \times [0, \nu^{\max}]\);
2. for the functions \(u^j\) introduced in Corollary [3.4] the following holds: for any sequence \(\{(t^n, \nu^n, \pi^n)\} \in [0, T] \times [0, \nu^{\max}] \times S^K\), such that \((t^n, \nu^n, \pi^n) \xrightarrow{n \to \infty} (t, \nu, \pi)\), one has
\[
    \lim_{n \to \infty} \sup_{z \in \text{supp}(\eta)} |u^j(t^n, \nu^n, \pi^n, z) - u^j(t, \nu, \pi, z)| = 0.
\]

**Relaxed controls and the Young topology.** The standard approach to ensure that \(A\) is compact is to extend the control set and to allow for randomized or relaxed controls. We define the set of relaxed controls by
\[
    \tilde{A} := \{ \alpha : [0, T] \to \mathcal{M}^1([0, \nu^{\max}]) \text{ measurable} \},
\]
where \(\mathcal{M}^1([0, \nu^{\max}])\) is the set of probability measures on \([0, \nu^{\max}]\). In line with the definition of an admissible liquidation strategy in the non-relaxed framework, an admissible relaxed liquidation strategy is a sequence of mappings \(\{\nu^n\} : \tilde{X} \to \tilde{A}\). A suitable topology on \(\tilde{A}\) is the so-called Young topology.

**Definition 4.6 (Young topology).** The Young topology on \(\tilde{A}\) is the coarsest topology such that all mappings of the form
\[
    \tilde{A} \ni \alpha \to \int_0^T \int_0^{\nu^{\max}} f(t, u) \alpha_u d\mu dt
\]
are continuous for all functions \(f : [0, T] \times [0, \nu^{\max}] \to \mathbb{R}\) that are continuous in the second argument, measurable in the first one, and satisfy \(\int_0^T \max \{|f(t, u)| : u \in [0, \nu^{\max}]\} dt < \infty\).
Notice that $\mathcal{A}$ is a subset of $\tilde{\mathcal{A}}$, since standard (non-relaxed) controls can be identified with Dirac measures on $[0,\nu^{\max}]$. Furthermore, applying the Chattering Theorem, see Kushner [15, Theorem 2.2], it is possible to show that $\mathcal{A}$ is dense in $\tilde{\mathcal{A}}$ with respect to the Young topology. These facts imply that relaxed controls can be approximated by deterministic controls. For further details on the Young topology we refer to Davis [30] or Bäuerle and Rieder [13, Ch. 8].

Next, we extend the definition of the reward function and the transition kernel of the MDM $\{L_n\}$ to a setup with relaxed controls. For a measurable function $A$ of the PDMP is defined as

$$g(\tilde{x}, \alpha) = \langle \alpha, g(\tilde{x}, \cdot) \rangle = \int_0^{\nu^{\max}} g(\tilde{x}, \nu)\alpha_s(d\nu).$$

As before, the flow associated to this vector field is denoted by $\tilde{\varphi}$ into the active boundary $\Gamma$ of $\tilde{\mathcal{X}}$. For a relaxed control $\alpha \in \tilde{\mathcal{A}}$ the jump intensity is given by $\lambda^\alpha_s(\tilde{x}) = \langle \alpha_s(d\nu), \lambda(t + s, \varphi^\alpha_s, \nu) \rangle$, and we set $\Lambda^\alpha_s = \Lambda^\alpha(\tilde{x}) = \int_0^s \lambda^\alpha_s(\tilde{x})du$.

Recall that the temporary price impact is modelled by the function $f$. Then the reward function is extended for controls from $\tilde{\mathcal{A}}$ as follows:

$$r(\tilde{x}, \alpha) = \int_0^{\tau^\tilde{x}} e^{s} e^{-\Lambda^\alpha_s} s(\alpha_s(d\nu), \nu - \nu f(\nu))du + e^{-\rho_{\tau^\tilde{x}}} e^{-\Lambda^\alpha_s} h(w_{\tau^\tilde{x}})s. \quad (4.8)$$

Finally we extend the transition kernel $Q_L$ to $\tilde{\mathcal{A}}$ by setting

$$Q_L\phi(\tilde{x}, \alpha) = \int_0^{\tau^\tilde{x}} e^{-\Lambda^\alpha_s} \langle \alpha_s(d\nu), \phi(\tilde{\varphi}(\tilde{x}), \nu) \rangle du + e^{-\Lambda^\alpha_s} \phi(\Delta).$$

**Bounding function** The following lemma gives the bounding function $b$ (see Definition B.3) for the current problem. Our choice of $b$ is motivated by Bäuerle and Rieder [13].

**Lemma 4.7.** For $\gamma > 0$, we define $b: \tilde{\mathcal{X}} \cup \{\Delta\} \to \mathbb{R}^+$ by $b(\tilde{x}) = b(t, x) := s\exp(T-t)$ for every $\tilde{x} \in \tilde{\mathcal{X}}$, and $b(\Delta) = 0$. Then under Assumptions 2.1 and 4.5, $b$ is a bounding function for the MDM with transition kernel $Q_L$ and reward function $r$. Moreover, the MDM is contracting for $\gamma$ sufficiently large.

The proof is postponed to Appendix B.2. In the sequel we denote by $B_b$ the set of all functions $v: \tilde{\mathcal{X}} \cup \{\Delta\} \to \mathbb{R}^+$ such that $v(\tilde{x}) \leq C_v b(\tilde{x})$ for all $\tilde{x} \in \tilde{\mathcal{X}} \cup \{\Delta\}$ and some $C_v \geq 0$. Note that for $v \in B_b$, $v(\Delta) = 0$.

**Continuity of $r$ and $Q_L$.** Finally we show that under Assumptions 2.1 and 4.5, $r$ and $Q_L$ satisfy the continuity conditions in Theorem B.4.

**Proposition 4.8.** Suppose that Assumptions 2.1 and 4.5 hold. Consider a continuous function $v \in B_b$. Then the mappings $(\tilde{x}, \alpha) \mapsto r(\tilde{x}, \alpha)$ and $(\tilde{x}, \alpha) \mapsto Q_L v(\tilde{x}, \alpha)$ are continuous on $\tilde{\mathcal{X}} \times \tilde{\mathcal{A}}$ with respect to the Young topology on $\tilde{\mathcal{A}}$.

**Proof.** Consider some sequence $(\tilde{x}_n, \alpha^n) \to (\tilde{x}, \alpha)$, where $\tilde{x} = (t, w, s, \pi)$. It follows from Davis [30, Theorem 43.5] that

$$\sup_{0 \leq u \leq T} |\tilde{\varphi}^n_u(\tilde{x}_n) - \tilde{\varphi}^n_u(\tilde{x})| \to 0 \text{ as } n \to \infty. \quad (4.9)$$
This does however not imply that \( \tau^{n} \), the entrance time of \( \tilde{\varphi}^{n}(\tilde{x}_{n}) \) into the active boundary of the state space, converges to \( \tau^{n} \); for this we would need a strictly positive lower bound on the liquidation rate, see also Davis \cite{Davis2015} Example 44.15. To deal with this issue we distinguish two cases:

**Case 1.** The flow \( \tilde{\varphi}^{n}(\tilde{x}) \) exits the state space \( \tilde{\Gamma} \) at the terminal boundary \( \Gamma_{2} \) (see (4.6)). This implies that \( \tau^{n} = T - t \) and that the inventory level \( w_{u} = w_{u}^{n}(\tilde{x}) \) is strictly positive for \( u < T - t \). We therefore conclude from (4.9) that \( \tau^{n} \) converges to \( T - t \) and that \( \varphi_{\tau^{n}}^{n}(\tilde{x}_{n}) \) converges to \( \varphi_{\tau^{n}}(\tilde{x}) \). The continuity of the one-period reward \( r \) and of \( Q_{\tilde{\nu}} \) follows immediately using the definition of \( r \) and the continuity of the mapping \( (\tilde{x}, \nu) \mapsto \tilde{Q}v(\tilde{x}, \nu) \) established in Lemma B.5 see Appendix B.2.

**Case 2.** The flow \( \tilde{\varphi}^{n}(\tilde{x}) \) exits \( \tilde{\Gamma} \) at the lateral boundary \( \Gamma_{1} \) so that \( w_{r^{n}} = 0 \). In that case it may happen that \( \limsup_{n \to \infty} \tau^{n} > \tau^{n} \); (4.9) only implies that \( \liminf_{n \to \infty} \tau^{n} \geq \tau^{n} \). In order to prove the continuity of \( r \) we decompose \( r(\tilde{x}_{n}, \alpha_{n}) \) (setting \( p = 0 \) for simplicity) as follows:

\[
\begin{align*}
  r(\tilde{x}_{n}, \alpha_{n}) &= s \int_{0}^{\tau_{\nu}} e^{-\Lambda_{u}^{n}(\tilde{x}_{n})(\alpha_{u,n}(d\nu), \nu - \nu f(\nu))} du \\
  &\quad + s \int_{\tau_{\nu} \wedge \tau_{r^{n}}} e^{-\Lambda_{u}^{n}(\tilde{x}_{n})(\alpha_{u,n}(d\nu), \nu - \nu f(\nu))} du + se^{-\Lambda_{r}^{n}(\tilde{x})} h(w_{r^{n}}). \tag{4.10}
\end{align*}
\]

Now it follows from (4.9) that the integral in (4.10) converges to \( r(\tilde{x}, \alpha) \). The terms in (4.11) are bounded from above by \( su_{\tau_{\nu} \wedge \tau_{r^{n}}} \); this can be shown via a similar partial integration argument as in the proof of Lemma 4.7 see Appendix B.2. Moreover, \( u_{\tau_{\nu} \wedge \tau_{r^{n}}}^{n} \) converges to \( w_{r^{n}} = 0 \), which gives the convergence \( r(\tilde{x}_{n}, \alpha_{n}) \to r(\tilde{x}, \alpha) \).

Finally we turn to the continuity of \( Q_{\tilde{\nu}} \). We decompose \( Q_{\tilde{\nu}} \) as follows:

\[
\begin{align*}
  Q_{\tilde{\nu}}v(\tilde{x}_{n}, \alpha_{n}) &= \int_{0}^{\tau_{\nu} \wedge \tau^{n}} e^{-\Lambda_{u}^{n}(\tilde{x}_{n})(\alpha_{u,n}(d\nu), \tilde{Q}v(\tilde{x}_{n}, \nu))} du \\
  &\quad + \int_{\tau_{\nu} \wedge \tau^{n}} e^{-\Lambda_{u}^{n}(\tilde{x}_{n})(\alpha_{u,n}(d\nu), \tilde{Q}v(\tilde{x}_{n}, \nu))} du. \tag{4.13}
\end{align*}
\]

The integral in (4.12) converges to \( Q_{\tilde{\nu}}v(\tilde{x}, \alpha) \) by (4.9) and the continuity of the mapping \( (\tilde{x}, \nu) \mapsto \tilde{Q}v(\tilde{x}, \nu) \) (Lemma B.5). To deal with the integral in (4.13) note that \( \tilde{Q}v(\tilde{x}, \nu) \leq csw_{\tilde{\lambda}}(\tilde{x}, \nu) \) (as \( \frac{1}{\tilde{\lambda}}\tilde{Q} \) is a probability transition kernel), so that (4.13) is smaller than

\[
sw_{\tau_{\nu} \wedge \tau^{n}}^{n} \int_{\tau_{\nu} \wedge \tau^{n}} \lambda_{u}^{n} e^{-\Lambda_{u}^{n}(\tilde{x})} du \leq sw_{\tau_{\nu} \wedge \tau^{n}}^{n} \to 0 \text{ for } n \to \infty.
\]

\(\square\)

**Optimality equation** In this section we use Theorem B.4 to derive the optimality equation for the value function \( V \). To formulate our result, for \( v : \tilde{\mathcal{X}} \cup \{\tilde{\Delta}\} \to \mathbb{R} \) we define the mapping \( \mathcal{L}v : (\tilde{\mathcal{X}} \cup \{\tilde{\Delta}\}) \times \tilde{\mathcal{A}} \to \mathbb{R} \) by \( \mathcal{L}v(\tilde{x}, \alpha) = r(\tilde{x}, \alpha) + Q_{\tilde{\nu}}v(\tilde{x}, \alpha) \) and the maximal reward operator by \( \tilde{T}v(\tilde{x}) = \sup_{\alpha \in \tilde{\mathcal{A}}} \mathcal{L}v(\tilde{x}, \alpha). \)

**Theorem 4.9.** Suppose that Assumptions 2.1 and 4.5 hold. Then the following assertion hold:
1. The value function $V$ is continuous on $\tilde{X}$ and satisfies the boundary conditions $V(\tilde{x}) = 0$ for $\tilde{x}$ in the lateral boundary $\Gamma_1$ and $V(T, x) = h(w)$. Moreover, $V$ is the unique fixed point of the operator $\tilde{T}$ in $B_b$.

2. There exists an optimal relaxed Markov control. More precisely, the stationary strategy \( \{\nu^{*,n}\} \) with $\nu^{*,n} = \nu^*$, for every $n \in \mathbb{N}$ and $\nu^*(\tilde{x}) \in \arg\max\{Lv(\tilde{x}, \alpha): \alpha \in \tilde{A}\}$, is optimal.

**Proof.** The claim follows immediately from Theorem 4.4 in Appendix B.1. Lemma 4.7 ensures the existence of a legitimate bounding function for the MDM, and Proposition 4.8 ensures that the continuity conditions are satisfied.

The equation $V = \tilde{T}V$ is the optimality equation for $V$. In explicit form it reads as follows:

$$V(\tilde{x}) = \sup_{\alpha \in \tilde{A}} \left\{ \int_0^{\tau^\varphi} e^{-\rho u} e^{-\Lambda^a_\nu(\tilde{x})} (\alpha_u(\nu), \nu s(1 - f(\nu)) + \bar{Q}V(\nu^a_u(\tilde{x}), \nu)) \, du 
\right. 
\left. + e^{-\rho \tau^\varphi} e^{-\Lambda^a_{\nu^*}(\tilde{x})} h(w(\tau^\varphi)) \right\}.$$ 

Since relaxed controls are difficult to interpret, we next provide a characterization of $V$ that involves only non-relaxed controls. Define the operator $T$ with $Tv(\tilde{x}) = \sup_{\alpha \in A} Lv(\tilde{x}, \alpha)$.

**Corollary 4.10.** Under the assumptions of Theorem 4.9, $V$ is the unique fixed point of the operator $T$ in $B_b$.

**Proof.** Clearly, $TV \leq \tilde{T}V = V$, as in the definition of $T$ the supremum is taken over a smaller set of controls than in the definition of $\tilde{T}$. Fix $\tilde{x} \in \tilde{X}$. By the Chattering Theorem Kushner [45, 3, Theorem 2.2] $A$ is dense in $\tilde{A}$ with respect to the Young topology, and that the mapping $\alpha \mapsto LV(\tilde{x}, \alpha)$ is continuous in the Young topology by Proposition 4.8. Hence we may choose a sequence $\{\alpha_n\} \in A$ with $\alpha_n \to \nu^*(\tilde{x})$ for $n \to \infty$ to obtain $\lim_{n \to \infty} LV(\tilde{x}, \alpha_n) = LV(\tilde{x}, \nu^*(\tilde{x})) = V(\tilde{x})$. This gives the equality $TV = V$. Uniqueness of the fixed point of $T$ is clear since the operator $T$ is contracting.

**5 Control Problem II: Viscosity Solutions**

In this section we follow the approach of Davis and Farid [31] and characterize the value function as the viscosity solution of the standard HJB equation associated with the controlled Markov process $(W, \pi)$. Moreover we derive a comparison principle for that equation. These results are crucial for establishing the convergence of suitable numerical schemes for the HJB equation and thus for the numerical solution of the optimal liquidation problem. Finally, we provide an example which shows that in general the HJB equation does not admit a classical solution.
5.1 Viscosity solution characterization As a first step we use the positive homogeneity of $V$ in the bid price $s$ to eliminate this variable from the set of state variables and to reduce the problem to a bounded state space. Define $\tilde{Y} = [0, T] \times [0, w_0] \times \mathcal{S}^K$ and set for $\tilde{y} \in \tilde{Y}$,

$$V'(\tilde{y}) = V'(t, w, \pi) := V(t, w, 1, \pi),$$

so that the value function satisfies the relation $V(\tilde{x}) = sV'(\tilde{y})$. For $\nu \in [0, \nu_{\text{max}}]$, $\tilde{y} \in \tilde{Y}$, and any measurable function $\Psi : \tilde{Y} \to \mathbb{R}^+$, define

$$Q^y\Psi(\tilde{y}, \nu) := \sum_{j=1}^K \pi_j \int \rho \gamma (1 + z) \Psi(t, w, (\pi_i(1 + u_i(t, \pi, \nu, z)))_{i=1,...,K}) \eta^j(t, \nu, d\nu).$$

Note that the function $g$ from (4.4) is independent of $s$, and define $g' : \tilde{Y} \times [0, \nu_{\text{max}}] \to \mathbb{R}^{K+2}$ by

$$(g')^1 = g^1, \quad (g')^2 = g^2, \quad \text{and} \quad (g')^{k+2} = g^{k+3}, \quad k = 1, \ldots, K.$$ 

Finally, denote by $\tilde{\nu}'_{\alpha}(\alpha, \tilde{y})$ the flow of $g'$ and by $\tau_{\alpha}'$ the first time this flow reaches the active boundary $I'$ of $\tilde{Y}$, where

$$I' = [0, T] \times \{0\} \times \mathcal{S}_0^K \cup \{T\} \times [0, w_0] \times \mathcal{S}_0^K.$$ 

Below, we write $\text{int} \tilde{Y}$ and $\partial \tilde{Y}$ to denote the relative interior and the relative boundary of $\tilde{Y}$ with respect to the hyperplane $\mathcal{H} = \mathbb{R}^2 \times \mathcal{H}^K$. Using Corollary 4.10, the homogeneity of $V$ and the fact that the jump intensity $\lambda$ introduced in (4.5) is independent of $s$, we get the following optimality equation for $V'$:

$$V'(\tilde{y}) = \sup_{\alpha \in \Lambda} \left\{ \int_0^{\tau_{\alpha}'(\tilde{y})} e^{-(\rho u + \Lambda^\alpha_{\text{top}}(\tilde{y}))} (\alpha_u(1 - f(\alpha_u))) + \tilde{Q}'\Psi(\tilde{\nu}_{\alpha}(\alpha, \tilde{y}), \alpha_u) \right\}$$

$$+ e^{-(\rho \tau_{\alpha}' + \Lambda^\alpha_{\text{top}}(\tilde{y}))} h(\tau_{\alpha}' \tilde{y}), \quad (5.1)$$

For $\Psi : \tilde{Y} \to \mathbb{R}^+$ bounded, define the function $\ell^\Psi : \tilde{Y} \times [0, \nu_{\text{max}}] \to \mathbb{R}^+$ by

$$\ell^\Psi(\tilde{y}, \nu) = \nu(1 - f(\nu)) + \tilde{Q}'\Psi(\tilde{y}, \nu), \quad (5.2)$$

and the operator $\mathcal{T}'$ by

$$\mathcal{T}'\Psi(\tilde{y}) = \sup_{\alpha \in \Lambda} \left\{ \int_0^{\tau_{\alpha}'(\tilde{y})} e^{-(\rho u + \Lambda^\alpha_{\text{top}}(\tilde{y}))} \ell^\Psi(\tilde{\nu}_{\alpha}(\tilde{y}), \alpha_u) du + e^{-(\rho \tau_{\alpha}' + \Lambda^\alpha_{\text{top}}(\tilde{y}))} h(\tau_{\alpha}' \tilde{y}) \right\}. \quad (5.3)$$

With this, (5.1) can be written as the fixed point equation $V' = \mathcal{T}'V'$. For fixed $\Psi$, $v^\Psi := \mathcal{T}'\Psi$ is the value function of a deterministic optimal exit-time problem. This problem is studied extensively in Barles [8], and we will use his results to obtain a viscosity solution characterization of $V'$.

For $\Psi : \tilde{Y} \to \mathbb{R}^+$, define the function $F_{\Psi} : \tilde{Y} \times \mathbb{R}^+ \times \mathbb{R}^{K+2} \to \mathbb{R}$ by

$$F_{\Psi}(\tilde{y}, v, p) := -\sup_{\nu \in [0, \nu_{\text{max}}]} \left\{ -(\rho + \lambda(\tilde{y}, \nu)) v + g(\tilde{y}, \nu)p + \ell^\Psi(\tilde{y}, \nu) \right\}.$$ 

Then the dynamic programming equation associated with the exit time optimal control problem (5.3) is

$$F_{\Psi}(\tilde{y}, v^\Psi(\tilde{y}), \nabla v^\Psi(\tilde{y})) = 0, \quad \text{for} \ \tilde{y} \in \text{int} \tilde{Y}, \quad v^\Psi(\tilde{y}) = h'(\tilde{y}) \quad \text{for} \ \tilde{y} \in \partial \tilde{Y}. \quad (5.4)$$
There are two issues with this equation: first, $v^\Psi$ is typically not sufficiently smooth to be a classical solution of (5.4); second, it is not clear how to treat the non-active part of the boundary, as for $\tilde{y} \in \partial \tilde{Y} \setminus \Gamma'$, $v^\Psi(y)$ is determined endogenously. To overcome these problems Barles [8] proposes to use the following notion of viscosity solutions.

**Definition 5.1.** (1) A bounded upper semi-continuous (u.s.c.) function $v$ on $\tilde{Y}$ is a viscosity subsolution of (5.4), if and only if for all $\phi \in C^1(\tilde{Y})$ the following holds: if $\tilde{y}_0 \in \tilde{Y}$ is a local maximum of $v - \phi$, one has

$$F^\Psi(\tilde{y}_0, v(\tilde{y}_0), \nabla \phi(\tilde{y}_0)) \leq 0 \text{ for } \tilde{y}_0 \in \text{int}\tilde{Y},$$

$$\min \{F^\Psi(\tilde{y}_0, v(\tilde{y}_0), \nabla \phi(\tilde{y}_0)), v(\tilde{y}_0) - h'(\tilde{y}_0)\} \leq 0 \text{ for } \tilde{y}_0 \in \partial \tilde{Y}.$$  \hfill (5.5)

(2) A bounded lower semi-continuous (l.s.c.) function $u$ on $\tilde{Y}$ is a viscosity supersolution of (5.4), if and only if for all $\phi \in C^1(\tilde{Y})$ the following holds: if $\tilde{y}_0 \in \tilde{Y}$ is a local minimum of $u - \phi$, one has

$$F^\Psi(\tilde{y}_0, u(\tilde{y}_0), \nabla \phi(\tilde{y}_0)) \geq 0 \text{ for } \tilde{y}_0 \in \text{int}\tilde{Y},$$

$$\max \{F^\Psi(\tilde{y}_0, u(\tilde{y}_0), \nabla \phi(\tilde{y}_0)), u(\tilde{y}_0) - h'(\tilde{y}_0)\} \geq 0 \text{ for } \tilde{y}_0 \in \partial \tilde{Y}.$$  \hfill (5.6)

(3) A viscosity solution $v^\Psi$ of (5.4) is either a continuous function on $\tilde{Y}$ that is both a sub and a supersolution of (5.4), or a bounded function the u.s.c. and l.s.c. envelopes of which are a sub and a supersolution of (5.4).

Note that Definition 5.1 allows for the case that for certain boundary points $\tilde{y}_0 \in \partial \tilde{Y}$ one has $v^\Psi(\tilde{y}_0) \neq h'(\tilde{y}_0)$. In particular, if $F^\Psi(\tilde{y}_0, v^\Psi(\tilde{y}_0), \nabla v^\Psi(\tilde{y}_0)) = 0$ in the viscosity sense, (5.5) and (5.6) hold irrespective of the value of $h'(\tilde{y}_0)$ and the boundary value $v^\Psi(y_0)$ is endogenously determined.

Now we return to the optimal liquidation problem. Since $V'$ satisfies the relation $V' = T'V'$, we expect that $V'$ solves (in the viscosity sense) the equation

$$F_{V'}(\tilde{y}, V'(\tilde{y}), \nabla V'(\tilde{y})) = 0, \text{ for } \tilde{y} \in \text{int}\tilde{Y}, \quad V'(\tilde{y}) = h'(\tilde{y}) \text{ for } \tilde{y} \in \partial \tilde{Y}.$$  \hfill (5.7)

The difference between (5.7) and (5.4) is that in the latter the function $\Psi$ is fixed.

**Definition 5.2.** A bounded u.s.c. function $v$ on $\tilde{Y}$ is a viscosity subsolution of (5.7), if and only if for all $\phi \in C^1(\tilde{Y})$ such that $\tilde{y}_0 \in \tilde{Y}$ is a local maximum of $v - \phi$ the relation (5.5) holds for $\Psi = v$.

A bounded (l.s.c.) function $u$ on $\tilde{Y}$ is a viscosity supersolution of (5.4), if and only if for all $\phi \in C^1(\tilde{Y})$ such that $\tilde{y}_0 \in \tilde{Y}$ is a local minimum of $u - \phi$, (5.6) holds with $\Psi = u$.

$V'$ is a viscosity solution of (5.7), if it is both a sub and a supersolution of that equation (with the same continuity properties as in Definition 5.1).

The next theorem is the main result of this section.

**Theorem 5.3.** Suppose that Assumptions 2.1 and 4.5 hold, and that the Markov chain $Y$ has no absorbing state ($-q_{kk} > 0$ for all $k$).

1. The value function $V'$ is a continuous viscosity solution of (5.7) in $\tilde{Y}$.  

21
2. A comparison principle holds for \((5.7)\): if \(v\) is a subsolution and \(u\) a supersolution of \((5.7)\) such that \(v(\bar{y})/w\) and \(u(\bar{y})/w\) are bounded on \(\mathcal{Y}\) and such that \(v = u = h'\) on the active boundary \(\Gamma'\), then \(v \leq u\) on \(\mathcal{Y}\). It follows that \(V'\) is the only continuous viscosity solution of \((5.7)\).

Proof. First, by Theorem 4.9, \(V'\) is continuous. Barles [8] Theorem 5.2 implies that \(V'\) is a viscosity solution of \((5.4)\) with \(\Psi = V'\) and hence of equation \((5.7)\); for the details we refer to Davis and Farid [31, Lemma 7.5].

Next we turn to the comparison principle. Note that in order to establish the inequality \(v \leq u\) we cannot directly apply the comparison principle obtained in Barles [8] for \((5.4)\), since we have to compare equations with different \(\Psi\). Instead we use an inductive argument based on the monotonicity of \(T'\). Let \(u_0 := u\) and define \(u_1 = T'u_0\). It follows from Barles [8] Theorem 5.2 that \(u_1\) is a viscosity solution of \((5.4)\) with \(\Psi = u_0\). Moreover, \(u_1(\bar{y})/w\) is bounded on \(\mathcal{Y}\) so that \(u_1 = h'\) on \(\Gamma'\). Since \(u_0\) is a supersolution of \((5.7)\) it is also a supersolution of \((5.4)\) with \(\Psi = u_0\). It follows from Barles [8] Theorem 5.3 (Comparison Principle for \((5.4)\)\), that \(u_1 \leq u_0\) this theorem applies since \(-q_{jk} > 0\) for all \(k\) and since \(u_0 = u_1 = h'\) on \(\Gamma'\). Define now inductively \(u_n = T'u_{n-1}\), and suppose that \(u_n \leq u_{n-1}\). Then, using the monotonicity of \(T'\), we have

\[
u_{n+1} = T'u_n \leq T'u_{n-1} = u_n.
\]

This proves that \(u_{n+1} \leq u_n\) for every \(n\). Moreover, the sequence \(\{u_n\}\) converges to \(V'\) since \(T'\) is a contraction (see the discussion preceding Theorem 5.4), so that \(u_n \geq V'\) for all \(n\).

Analogously we can construct a sequence of functions \(\{v^n\}\) with \(v_0 = v\) such that \(v^n \uparrow V'\), and we conclude that \(v \leq V' \leq u\). The remaining statements are clear. \(\square\)

We briefly comment on the relation of Theorem 5.3 to the results in Davis and Farid [31]. Their analysis implies that \(V'\) is the unique continuous viscosity solution of \((5.7)\). However, they do not give a comparison principle for \((5.7)\). This is crucial since we need comparison to establish the convergence of the numerical scheme for \(V'\). Finally write the HJB equation \((5.7)\) more explicitly. To this, we use the fact that \(\lambda(\bar{y},\nu) = \sum_{k=1}^K \pi^k\eta^k(t,\nu,\mathbb{R})\), the definition of \(g\) in (4.1) and the definition of \(V'\) in (5.2) to obtain the equation

\[
0 = \frac{\partial V'}{\partial t}(t, w, \pi) + \sup \left\{ H(\nu, t, w, \pi, V', \nabla V') : \nu \in [0, \nu_{\text{max}}] \right\},
\]

with \(H\) given by

\[
H(\nu, t, w, \pi, V', \nabla V') = -\rho V' + \nu(1 - f(\nu)) - \nu \frac{\partial V'}{\partial w}(t, w, \pi) + \sum_{k,j=1}^K \frac{\partial V'}{\partial \pi^k}(t, w, \pi)(q_{jk}^k - \pi^k \int_{\mathbb{R}} u^k(t, \nu, \pi, z)\eta^j(t, \nu, dz))
\]

\[
+ \sum_{j=1}^K \pi^j \int_{\mathbb{R}} \Delta V'(t, w, \pi, z)\eta^j(t, \nu, dz),
\]

where

\[
\Delta V'(t, w, \pi, z) := (1 + z) V'(t, w, (\pi^i(1 + u^i(t, \nu, \pi, z)))_{i=1,...,K}) - V'(t, w, \pi).
\]

This equation coincides with the standard HJB equation associated with the controlled Markov process \((W, \pi)\). The advantage of using viscosity solution theory is that we are able to give a
precise mathematical meaning to this equation, even in those cases where the value function is merely continuous. Moreover, in the current setting we cannot expect the existence of a classical solution to the HJB equation. Indeed, in the next section we provide an example showing that the solution to the HJB equation only exists in the viscosity sense.

5.2 A counterexample We work in the context of Example 2.2 (linear permanent price impact and no Markov chain). For simplicity we let \( \rho = 0, s_0 = 1, h(\nu) \equiv 0, f(\nu) \equiv 0 \) (no temporary price impact). Moreover, we assume that \( e^{\text{up}} < e^{\text{down}} \). Recalling the definition of \( \hat{\eta}^P \) in equation (2.4), we have that for \( \nu \in [0, \nu^{\max}] \),

\[
\hat{\eta}^P(\nu) := \theta(e^{\text{up}} - e^{\text{down}}(1 + a\nu)) < 0.
\]

By the semimartingale decomposition of \( S \) given in equation (2.5) we get that \( S \) is a supermartingale. We conjecture that it is optimal to sell as fast as possible to reduce the impact of the falling price, that is \( \nu^*_t = \nu^{\max}1_{[0,w_0/\nu^{\max} \wedge T]}(t) \). In that case, for \( t < w_0/\nu^{\max} \wedge T \) it holds that \( \hat{\eta}^P(\nu_t) = \hat{\eta}^P(\nu^{\max}) \) and the expected bid price is \( \mathbb{E}(S^*_t) = e^{\hat{\eta}^P(\nu^{\max})} \). By Fubini’s Theorem we hence get

\[
\mathbb{E}\left( \int_0^T \nu^*_s S^*_s ds \right) = \int_0^{w_0/\nu^{\max} \wedge T} \nu^{\max}_s e^{\hat{\eta}^P(\nu^{\max})} ds
\]

and this integral is equal to

\[
V'(t, w) := \frac{\nu^{\max}}{\hat{\eta}^P(\nu^{\max})} \left( e^{\hat{\eta}^P(\nu^{\max})(w/\nu^{\max} \wedge (T-t))} - 1 \right).
\]

(5.9)

In order to verify that \( V' \) is in fact the value function, we show that \( V' \) is a viscosity solution of the corresponding HJB equation (5.1). Note that in the current setting this equation reads as

\[
-\frac{\partial V'}{\partial t} - \sup \left\{ \nu - \frac{\partial V'}{\partial w} - \hat{\eta}^P(\nu)V' : \nu \in [0, \nu^{\max}] \right\} = 0.
\]

(5.10)

First note that \( V' \) cannot be a classical solution as it is not differentiable for points \((t, w)\) with \( w = \nu^{\max}(T-t) \) and that \( V' \) satisfies the correct terminal and boundary conditions. Next we verify that \( V' \) is a viscosity solution. Fix some point \((\bar{t}, \bar{w})\). There are three different cases.

Case 1: \((T - \bar{t})\nu^{\max} < \bar{w} \). In that case \( \frac{\partial V'}{\partial w} = 0 \) and

\[
\frac{\partial V'}{\partial t} = \frac{\nu^{\max}}{\hat{\eta}^P(\nu^{\max})} \frac{\partial}{\partial t} \left( e^{\hat{\eta}^P(\nu^{\max})(T-t)} - 1 \right) = -\nu^{\max} e^{\hat{\eta}^P(\nu^{\max})(T-t)}.
\]

Moreover, comparing the cases \( \nu = 0 \) and \( \nu = \nu^{\max} \), we get that the supremum in (5.10) is attained for \( \nu = \nu^{\max} \) (selling at the maximum rate is optimal), and a direct computation shows that the HJB equation (5.10) is satisfied in the classical sense.

Case 2: \((T - \bar{t})\nu^{\max} > \bar{w} \). In that case \( \frac{\partial V'}{\partial w} = 0 \) and

\[
\frac{\partial V'}{\partial t} = \frac{\nu^{\max}}{\hat{\eta}^P(\nu^{\max})} \frac{\partial}{\partial t} \left( e^{\hat{\eta}^P(\nu^{\max})(w/\nu^{\max})} - 1 \right) = e^{\hat{\eta}^P(\nu^{\max})(w/\nu^{\max})}.
\]

Again, comparing the cases \( \nu = 0 \) and \( \nu = \nu^{\max} \) we get that the supremum in (5.10) is attained for \( \nu = \nu^{\max} \), (as \( \hat{\eta}^P(\nu) \leq 0 \) for \( \nu = 0 \) due to the assumption \( e^{\text{up}} < e^{\text{down}} \)), and a direct computation shows that (5.10) holds in the classical sense.
Case 3: \((T - t)\nu^\text{max} = \overline{w}\). Here we have to consider the sub and supersolution property. For the supersolution property there is nothing to show as there is no \(C^1\)-function \(\phi\) such that \(V' - \phi\) has a local minimum in \((\bar{t}, \overline{w})\). For the subsolution, consider \(\phi \in C^1\) such that \(V' - \phi\) has a local maximum in \((\bar{t}, \overline{w})\). Then we get that \(-\nu^\text{max} e^{\eta P(\nu^\text{max})(T-t)} \leq \frac{\partial \phi}{\partial t} \leq 0\) and that \(0 \leq \frac{\partial \phi}{\partial w} \leq e^{\eta P(\nu^\text{max})(T-t) - 1}\). Differentiating with respect to \(t\) gives that

\[
\left(\frac{\partial \phi}{\partial t} - \nu^\text{max} \frac{\partial \phi}{\partial w}\right)(\bar{t}, \overline{w}) = -\nu^\text{max} e^{\eta P(\nu^\text{max})(T-t)}.
\]

Using the inequalities for \(\frac{\partial \phi}{\partial w}\) we get that

\[
\sup \{\nu - \nu \frac{\partial \phi}{\partial w} + \eta P(\nu)V' : \nu \in [0, \nu^\text{max}]\} = \nu^\text{max}\left( -\frac{\partial \phi}{\partial w} + e^{\eta P(\nu^\text{max})(T-t)}\right).
\]

As \(\frac{\partial \phi}{\partial t} = \nu^\text{max}\left( -\frac{\partial \phi}{\partial w} - e^{\eta P(\nu^\text{max})(T-t)}\right)\), this gives

\[
-\frac{\partial \phi}{\partial t} - \sup \{\nu - \nu \frac{\partial \phi}{\partial w} + \eta P(\nu)V' : \nu \in [0, \nu^\text{max}]\} = 0
\]

and hence the subsolution property.

Note that the counterexample also works for nonzero temporary price impact, i.e. \(f(\nu) \neq 0\), see Figure 1 below. However, taking \(f(\nu) \equiv 0\) allows us to compute the value function and the optimal strategy explicitly.

---

**Figure 1:** Value function in Example 1 for nonzero temporary price impact with parameters \(c^\text{up} = 990\), \(c^\text{down} = 1000\), \(w = 6000\), and the other parameters as in Table 1.

---

**Unbounded liquidation rate.** Finally we consider the limit of the value function as \(\nu^\text{max} \to \infty\). Note that \(\lim_{\nu \to \infty} \frac{\nu}{\eta P(\nu)} = -\frac{1}{\theta c^\text{down}}\). Denote by \(V'(t, w; \nu^\text{max})\) the value function in (5.9).

Then it holds that

\[
\lim_{\nu^\text{max} \to \infty} V'(t, w; \nu^\text{max}) = V'(t, w; \infty) := -\frac{1}{\theta c^\text{down}}\left( e^{-w \theta a c^\text{down}} - 1\right).
\]

\(V'(:, \infty)\) is a strict (classical) supersolution of equation (5.10) since

\[
\nu - \nu \frac{\partial V'}{\partial w} - \eta P(\nu)V' = \left(1 - e^{-w \theta a c^\text{down}}\right)\left(2\nu - \frac{c^\text{up} - c^\text{down}}{a c^\text{down}}\right) > 0.
\]
This shows that $V'(t, w; \infty)$ is the value function of the optimal liquidation problem for $\nu^{\max} = \infty$, and we conclude that in the absence of a bound on the liquidation rate the viscosity-solution characterization from Theorem 5.3 ceases to hold in general.

6 Examples and numerical results

In this section we study in detail the value function and the optimal liquidation rate for Example 2.3. Since the HJB equation cannot be solved analytically, we resort to numerical methods.

Numerical method and parameters. In the following we briefly sketch the approach we use: We apply an explicit finite difference scheme to solve the HJB equation and to compute the corresponding liquidation strategy. First, we turn the HJB equation into an initial value problem via time reversal. Given a time discretization $0 = t_0 < \cdots < t_k < \cdots < t_m = T$ we set $V'_{t_0} = h$, and given $V'_{t_k}$, we approximate the liquidation strategy as follows:

$$
\nu^*_t(w, \pi) := \arg\max_{\nu \in [0, \nu^{\max}]} H(\nu, t_k, w, \pi, V'_{t_k}, \nabla^{\text{disc}} V'_{t_k}),
$$

where $\nabla^{\text{disc}}$ is the gradient operator with derivatives replaced by suitable finite differences. With this we obtain the next time iterate of the value function,

$$
V'_{t_{k+1}} = V'_{t_k} + (t_{k+1} - t_k) H(\nu^*_t, t_k, w, \pi, V'_{t_k}, \nabla^{\text{disc}} V'_{t_k}).
$$

Since we showed in Theorem 5.3 that the comparison principle holds and that the value function is the unique viscosity solution of our HJB equation, we get convergence of the proposed procedure to the value function by similar arguments as in Barles and Souganidis [9], Dang and Forsyth [29]; details are omitted.

In our numerical experiments – unless stated otherwise – we work with the parameter set given in Table 1. The maximum rate of selling $\nu^{\max}$ is chosen so that it is possible to liquidate the whole portfolio before $T$. We model the temporary price impact as $f(\nu) = c_f \nu^\varsigma$, with $\varsigma > 0$. Moreover, we set the liquidation value $h(w) \equiv 0$, i.e. we do not allow for a block sale at maturity $T$. Without loss of generality we set $s = 1$, so that the expected liquidation profit is equal to $V'$.

<table>
<thead>
<tr>
<th>$w_0$</th>
<th>$\nu^{\max}$</th>
<th>$T$</th>
<th>$\rho$</th>
<th>$\theta$</th>
<th>$a$</th>
<th>$c_f$</th>
<th>$\varsigma$</th>
<th>$q^{12}$</th>
<th>$q^{21}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>60000</td>
<td>18000</td>
<td>1 day</td>
<td>0.00005</td>
<td>0.001</td>
<td>$4 \times 10^{-6}$</td>
<td>$5 \times 10^{-5}$</td>
<td>0.6</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1: Parameter values used in numerical experiments.

Example 2.3 re-revisited. In this example, $\eta^P$ depends on the 2-state Markov chain $Y$ and is of the form

$$
\eta^P(e_i, \nu, z) = (1 + a \nu t)(c_1^{\text{down}} c_2^{\text{down}}) e_i \delta_{-\theta}(dz) + (c_1^{\text{up}} c_2^{\text{up}}) e_i \delta_{\theta}(dz).
$$

Since $\pi_1^1 + \pi_2^1 = 1$, we can eliminate $\pi_2^1$ from the set of state variables. In the sequel we denote by $V'(t, w, \pi^1)$ the value function evaluated at the point $(t, w, \pi^1, 1 - \pi^1)$. Using the functions
$u^i$ given in (3.6), the Hamiltonian $H$ in (5.8) becomes

$$H(\nu, t, w, \pi^1, V^i, \nabla V^i) = -\rho V^i(t, w, \pi^1) + \nu(1 - f(\nu)) - \nu \frac{\partial V'}{\partial w}(t, w, \pi^1)$$

$$+ \frac{\partial V'}{\partial \pi^1}(t, w, \pi^1)(\pi^1 q^{11} + (1 - \pi^1)q^{21})$$

$$- \frac{\partial V'}{\partial \pi^1}(t, w, \pi^1)^1(1 - \pi^1)((1 + \nu)(c_{1\text{down}}^{\text{down}} - c_{1\text{down}}^{\text{down}}) + c_{1\text{up}}^{\text{up}} - c_{2\text{up}}^{\text{up}})$$

$$+ [(1 - \theta) V'(t, w, \pi^1) \pi^1 c_{1\text{down}}^{\text{down}} + (1 - \pi^1) c_{2\text{down}}^{\text{down}}) - V'(t, w, \pi^1)]$$

$$\times (\pi^1 c_{1\text{down}}^{\text{down}} + (1 - \pi^1) c_{2\text{down}}^{\text{down}})(1 + \nu)$$

$$+ [(1 + \theta) V'(t, w, \pi^1) \pi^1 c_{1\text{up}}^{\text{up}} + (1 - \pi^1) c_{2\text{up}}^{\text{up}}) - V'(t, w, \pi^1)] (\pi^1 c_{1\text{up}}^{\text{up}} + (1 - \pi^1) c_{2\text{up}}^{\text{up}}).$$

The liquidation strategy is given by

$$\nu^*_k = \min \left\{ \max \left\{ \frac{1}{c_f(s + 1)} (1 - \partial_{\pi^1} V'_{k} - \partial_{\pi_1} V'_{k} \pi^1 (1 - \pi^1)) a(c_{1\text{down}}^{\text{down}} - c_{2\text{down}}^{\text{down}}) \right\}, 0 \right\}^{1/\kappa}, \nu^{\text{max}} \right\},$$

where $\Delta V'(t_k, w, \pi^1) = (1 - \theta) V'(t_k, w, \pi^1) \pi^1 c_{1\text{down}}^{\text{down}}(1 - \pi^1) + (1 - \pi^1) c_{2\text{down}}^{\text{down}}) - V'(t_k, w, \pi^1),$ and $\partial_{\pi^1}$ denotes the discretized partial derivative in direction $\pi^1$.

We focus on three different research questions: (i) the impact of $\nu^{\text{max}}$ on the expected proceeds from liquidation; (ii) the additional liquidation profit from the use of stochastic filtering; (iii) the influence of the temporary and permanent price impact parameters on the form of the liquidation strategy.

(i) Impact of $\nu^{\text{max}}$. Table 2 displays the value function $V'(0, w_0, \pi^1; \nu^{\text{max}})$ for varying $\nu^{\text{max}}$ (expressed as multiple of the initial inventory $w_0$) and for fixed $w_0 = 6000$, $\pi^1 = \frac{1}{2}$. The value grows in $\nu^{\text{max}}$, but for $\nu^{\text{max}} > 2w_0$, the additional gain is small. This supports the argument from Proposition 2.4 that imposing an upper bound on $\nu$ is not a severe restriction.

<table>
<thead>
<tr>
<th>$\nu^{\text{max}}$</th>
<th>$w_0$</th>
<th>$2w_0$</th>
<th>$3w_0$</th>
<th>$5w_0$</th>
<th>$7w_0$</th>
<th>$10w_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V'(0, w_0, \pi^1; \nu^{\text{max}})$</td>
<td>5783.40</td>
<td>5876.85</td>
<td>5876.93</td>
<td>5876.97</td>
<td>5876.99</td>
<td>5877.00</td>
</tr>
</tbody>
</table>

Table 2: The expected proceeds from liquidation $V'(0, w_0, \pi^1; \nu^{\text{max}})$ for varying $\nu^{\text{max}}$ and fixed $w_0 = 6000$, $\pi^1 = 1/2$.

(ii) Gain from stochastic filtering. We compare the optimal expected proceeds from liquidation in this setup to the expected proceeds in case that the trader (mistakenly) uses a model with deterministic $\eta^P$ and $c_{1\text{up}} = 0.5c_{1\text{up}} + 0.5c_{2\text{up}}$, $c_{1\text{down}} = 0.5c_{1\text{down}} + 0.5c_{2\text{down}}$, i.e. she works with the stationary distribution of $Y$ throughout. For $c_{1\text{up}} = c_{2\text{down}} = 1000$, $c_{1\text{up}} = c_{2\text{down}} = 800$, the expected gain from the use of filtering is € 61.99. For $q^{12} = q^{21} = 2$ the effect becomes even stronger with € 128.12. This shows that the additional complexity of using filtering methods to learn about the state of the market may be worthwhile.
(iii) **Form of the liquidation strategy for varying price impact parameters.** Throughout we fix $c_{1}^{up} = c_{2}^{down} = 1000$, $c_{2}^{up} = c_{1}^{down} = 950$. First, we discuss the form of the liquidation rate for varying size of the temporary price impact, that is for varying $c_f$ and for $a$ as in Table 1 (moderate value of $a$). Figure 2 shows the liquidation rate at $t = 0$ for the cases of no, intermediate, and large temporary price impact as a function of $w$ and $\pi^1$. The figure is a contour plot: white areas correspond to $\nu_0 = 0$; black areas correspond to selling at maximum speed ($\nu_0 = 18000$); grey areas correspond to selling at a moderate speed, see also the color bars below the graphs. We observe that the liquidation rate is decreasing in $\pi^1$ (the probability that $Y$ is in the good state) and increasing in the inventory level. Comparing the graphs we see that for higher temporary price impact the trader tends to trade more evenly over the state space to keep the cost due to the temporary price impact small. Note that for $c_f \rightarrow 0$ the liquidation strategy converges to a *bang-bang type strategy*. The optimal policy is then characterized by two regions: a *sell* region, where the trader sells at the maximum speed, and a *wait* region, where she does not sell at all.

![Figure 2: Contour plot of the liquidation policy as a function of $w$ ($x$-axis) and $\pi^1$ ($y$-axis) for $c_f = 5 \times 10^{-11}$ (left), $c_f = 10^{-5}$ (middle), and $c_f = 5 \times 10^{-5}$ (right) and $t = 0$ for Example 2.3. Note the different scale in the graphs.](image)

Now we study the situation where the permanent price impact becomes large. The left plot in Figure 3 shows the sell and wait regions under partial information in dependence of the inventory level $w$ and the filter probability $\pi^1$ for $a = 10^{-4}$, and (essentially) without temporary price impact. The sell region forms a band from low values of $w$ and $\pi^1$ to high values of $w$ and $\pi^1$. In the presence of a temporary price impact (right plot of Figure 3) the qualitative behaviour of the liquidation rate is similar to the case without temporary price impact, but the transition from waiting to selling at the maximum rate is smooth.

![Figure 3: Contour plot of the liquidation policy as a function of $w$ ($x$-axis) and $\pi^1$ ($y$-axis) for $c_f = 5 \times 10^{-11}$ (left) and $c_f = 10^{-5}$ (right) for $a = 10^{-4}$ and $t = 0$ for Example 2.3.](image)
We offer the following interpretation. The qualitative behaviour of the liquidation rate is determined by an interplay of two effects. The price anticipation effect makes the trader ‘wait’ in the good state where prices are increasing on average (for $\mu_t \equiv 0$), and ‘sell’ in the bad state where prices are trending downward. On the other hand, there is a liquidity effect. In our model the permanent price impact is proportional to $e^{\text{down}}$, which is small in the good state and large in the bad state. Hence the trader has an incentive to sell in the good state and wait in the bad state to reduce the permanent price impact. As we see in Figure 2, for small values of $a$ the price anticipation effect dominates, explaining the observed monotonicity behaviour in $\pi^1$. For large $a$ and large $w$ the liquidity effect dominates (‘sell’ for $\pi^1$ close to one and ‘wait’ for $\pi^1$ close to zero). The observation that for $w$ small the price anticipation effect dominates, whereas for $w$ large the liquidity effect dominates, is due to the fact that for small $w$ the trader can liquidate (most of) her inventory before the negative impact on the stock price materializes. For large $w$ on the other hand selling can be prohibitively expensive and the trader prefers to waits for better market conditions. This leads to an interesting feature for the case where $w$ is large and $\pi^1$ is small: there is a gambling region where the trader prefers to wait for better market liquidity, even if he is confronted with a downward-trending price.

A Setup and filtering: proofs and additional results

Lemma A.1. Suppose that Assumption 2.1 holds. Fix $m > w_0/T$ and consider some $\mathbb{F}^S$-adapted strategy $\nu$ with values in $[0, m]$. Define

$$C := 0 \lor \sup \left\{ \int_{\mathbb{R}} (z^2 + 2z)\eta^p(t, e, \nu, dz) : (t, e, \nu) \in [0, T] \times \mathcal{E} \times [0, m] \right\}.$$ 

Then $C < \infty$, $E((S^\nu_t)^2) \leq S_0^2 e^{Ct}$, and $(\int_0^t S^\nu_s dM^R_s)_{0 \leq t \leq T}$ is a true martingale.

Proof. To ease the notation we write $S_t$ for $S^\nu_t$. We begin with the bound on $S^2_t$. First note that $C$ is finite by Assumption 2.1. At a jump time $T_n$ of $R$ it holds that $S_{T_n} = S_{T_n-} (1 + \Delta R_{T_n})$ and therefore

$$S^2_{T_n} - S^2_{T_n-} = S^2_{T_n-} \Delta R^2_{T_n} + 2S_{T_n-} \Delta R_{T_n}.$$ 

Hence $S^2_t = S^2_0 + \int_0^t \int_{\mathbb{R}} S^2_{s-}(z^2 + 2z)\mu^R(dz, ds)$ and we get

$$E(S^2_t) = S_0^2 + E\left( \int_0^t \int_{\mathbb{R}} S^2_{s-}(z^2 + 2z)\eta^p(s, Y_s-\nu_s-, dz) ds \right)$$

$$\leq S_0^2 + C \int_0^t E(S^2_s) ds,$$

so that $E((S^\nu_t)^2) \leq S_0^2 e^{Ct}$ by the Gronwall inequality. To show that $\int_0^t S^\nu_s dM^R_s$ is a true martingale we show that this process has integrable quadratic variation. Since $[\int_0^t S^\nu_s dM^R_s]_t = \int_0^t \int_{\mathbb{R}} S^2_{s-} z^2 \mu^R(dz, ds)$, we have

$$E\left( \int_0^t S^\nu_s dM^R_s \right)_t = E\left( \int_0^t S^2_s \int_{\mathbb{R}} z^2 \eta^p(s, Y_s-\nu_s-, dz) ds \right) \leq S_0^2 \tilde{C} \int_0^t e^{Ct} ds,$$

for every $t \in [0, T]$, where $\tilde{C} = \sup \left\{ \int_{\mathbb{R}} z^2 \eta^p(t, e, \nu, dz) : t \in [0, T], e \in \mathcal{E}, \nu \in [0, m] \right\}$ is finite by Assumption 2.1. \qed
Lemma 3.1. Conditions (2.6) and (2.7) imply that \( \tilde{Z} \) is a true martingale, see Protter and Shimbo [18]. Moreover, \( \beta(t, Y_{t-}, \nu_{t-}, z) > -1 \), since \( \text{d}p(t, e_i, \nu)/\text{d}q(t, z) > 0 \) by assumption. This implies that \( \tilde{Z}_T > 0 \), and hence the equivalence of \( P \) and \( Q \). Next we identify the \( P \)-compensator of \( \mu^{(Y, R)} \).

Define the function \( \tilde{\beta} : [0, T] \times \mathcal{E} \times [0, \nu_{\max}] \times E \times \mathbb{R} \to (-1, \infty) \) by

\[
\tilde{\beta}(t, e, \nu, \gamma, z) = \begin{cases}
\beta(t, e, \nu, z) & \text{for } \gamma = 0 \\
0 & \text{for } \gamma \neq 0,
\end{cases}
\]

and note that the stochastic differential equation for \( \tilde{Z} \) can be written in the form

\[
\text{d}\tilde{Z}_t = \tilde{Z}_{t-} \int_{\Gamma \times \mathbb{R}} \tilde{\beta}(t, Y_{t-}, \nu_{t-}, \gamma, z) \left( \mu^{(Y, R)}(dt, dz, dt) - \tilde{\eta}_t^Q(d\gamma, dz)dt \right).
\]

The Girsanov theorem for random measures (see Brémaud [18, VIII, Theorem T10]) shows that under \( P \), \( \mu^{(Y, R)}(dt, dz) \) has the predictable compensator \( (\beta(t, Y_{t-}, \nu_{t-}, z) + 1)\tilde{\eta}_t^Q(dz)dt \). By definition of \( \beta \) this is equal to \( \tilde{\eta}_t^P(t, Y_{t-}, \nu_t, dz)dt \).

Conversely, in order to construct \( Q \) define the process \( Z \) by

\[
\text{d}Z_t = Z_{t-} \int_{\mathbb{R}} \left( \frac{\text{d}\eta_t^Q}{\text{d}p(t, Y_{t-}, \nu_{t-})}(z) - 1 \right) \left( \mu^R(dt, dz) - \tilde{\eta}_t^P(t, Y_{t-}, \nu_{t-}, dz) \right),
\]

with initial condition \( Z_0 = 1 \). Then \( t \in [0, T] \) is a true \( (\mathbb{F}, P) \) martingale and we may define a measure \( Q \) by \( \text{d}Q/\text{d}P \big|_{\mathcal{F}_T} = Z_T \). The Girsanov theorem for random measures ensures that under \( Q \), the measure \( \mu^{(Y, R)}(dt, dz) \) has the predictable compensator \( \tilde{\eta}_t^Q(dz)dt \). Using the relation

\[
\frac{\text{d}\eta_t^Q}{\text{d}p(t, Y_{t-}, \nu_{t-})}(z) - 1 = \frac{1}{\beta(t, Y_{t-}, \nu_{t-}, z) + 1} - 1 = \frac{-\beta(t, Y_{t-}, \nu_{t-}, z)}{\beta(t, Y_{t-}, \nu_{t-}, z) + 1},
\]

for every \( t \in [0, T] \), Itô’s formula for jump processes finally gives that \( Z_t = (\tilde{Z}_t)^{-1} \). Therefore, the two measure changes are inverse to each other.

In the sequel we want to prove Theorem 3.2. For this we need the following result.

Lemma A.2. For \( t \in [0, T] \) let \( U \) be an integrable \( \mathcal{F}_t \)-measurable random variable and let \( \mathcal{F}^S := \bigvee_{t \geq 0} \mathcal{F}_t^S \). Then \( \mathbb{E}^Q \left( U \mid \mathcal{F}_t^S \right) = \mathbb{E}^Q \left( U \mid \mathcal{F}^S \right) \).

Proof. First note that for \( t \in [0, T] \), \( \mathcal{F}_t^S = \mathcal{F}_t^R \), as the jumps of \( R \) are greater than \(-1\). Define \( \mathcal{F}_t^R := \sigma\left( R_{t+u} - R_u, \ u \geq 0 \right) \). Then \( \mathcal{F}^S = \sigma\left( \mathcal{F}_t^S, \mathcal{F}_t^R \right) \). Under the probability measure \( Q \), the \( \sigma \)-algebra \( \mathcal{F}_t^S \subset \mathcal{F}^S \) is independent of \( \mathcal{F}_t \). This is a consequence of the fact that under \( Q \), \( \mu^R \) is a Poisson random measure, and then \( R \) has independent increments. Therefore, since \( U \) is \( \mathcal{F}_t \)-measurable, we get that \( \mathbb{E}^Q \left( U \mid \mathcal{F}_t^S \right) = \mathbb{E}^Q \left( U \mid \sigma\left( \mathcal{F}_t^S, \mathcal{F}_t^R \right) \right) = \mathbb{E}^Q \left( U \mid \mathcal{F}^S \right) \).

Proposition 3.3. In order to obtain the Kushner-Stratonovich equation we apply Itô’s formula to compute the dynamics of \( p(f)/p(1) \). From Theorem 3.2 we have \( \text{d}p_t(1) = \int_{\mathbb{R}} p_{t-}(\beta(z))(\mu^R(dt, dz) - \eta_t^Q(dz)dt) \), and hence

\[
\text{d}\left( \frac{1}{p_t(1)} \right) = -\frac{1}{(p_t(1))^2} \int_{\mathbb{R}} p_{t-}(\beta(z)) \eta_t^Q(dz)dt - \frac{1}{p_t(1)} \int_{\mathbb{R}} \frac{p_{t-}(\beta(z))}{p_{t-}(1 + \beta(z))} \mu^R(dt, dz).
\]
Since \( \pi(f) = \frac{p(f)}{P(\tau)} \), by the product rule we get

\[
\begin{align*}
\pi_t(f) &= \pi_t(Qf)dt \\
&+ \int_{\mathbb{R}} \left( \frac{\pi_t(\beta(z)f)}{\pi_t(\beta(z) + 1)} - \pi_t(f) \frac{\pi_t(\beta(z))}{\pi_t(\beta(z) + 1)} \right) (\mu^R(dt, dz) - \pi_t(\eta^P(dz))dt).
\end{align*}
\]

The integrand above is equal to

\[
(\pi_t((\beta(z) + 1)f) - \pi_t(\beta(z) + 1)\pi_t(f)) / \pi_t(\beta(z) + 1).
\]

It follows from the definition of \( \beta \) in (3.1) that \( \pi_t(1 + \beta(z)) = \pi_t(d\eta^P/d\eta^Q(z)) \). Making the necessary substitutions and cancellations we get the result. \( \Box \)

### B Optimization via Markov decision models

#### B.1 Markov Decision Models

For the convenience of the reader we recall a few basics of MDMs. Our presentation is largely based on Bäuerle and Rieder [13, Chapter 7].

**Definition B.1.**

1. A **stationary infinite horizon MDM** consists of the data \((\tilde{X}, A, Q', r)\), where the Borel set \( \tilde{X} \) is the state space, the Borel set \( A \) is the set of control actions, \( r : \tilde{X} \times A \to \mathbb{R}^+ \) is the reward function and \( Q'(dl' \mid l, \alpha) \) is a transition kernel from \( B(\tilde{X}) \times \tilde{X} \times A \to [0, 1] \).

2. A **control policy** \( \nu^n = \nu^0, \nu^1, \ldots \) is a sequence of mappings \( \nu^n : \tilde{X} \to A \). If \( \nu^n = \nu \) for all \( n \), then the policy \( \nu^n \) is called **stationary**.

Given a control policy \( \nu^n \) and an initial state \( l_1 \in \tilde{X} \), one can construct a sequence \( \{L_n\}_{n \in \mathbb{N}} \) of random variables with \( L_1 = l_1 \) and

\[
P(L_{n+1} \in B \mid L_n) = Q'(B \mid L_n, \nu^n(L_n)).
\]

The reward associated with the policy \( \nu^n \) is then

\[
J^{\nu^n}(l) = \mathbb{E}^{\nu^n}_{l_1} \left[ \sum_{n=0}^{\infty} r(L_n, \nu^n(L_n)) \right],
\]

and the value function is \( J_{\infty}(l) = \sup \left\{ J^{\nu^n}(l), \nu^n \right\} \) control policy \( \nu^n \).

**Definition B.2.** For \( \alpha \in A \) and a measurable function \( v : \tilde{X} \to \mathbb{R}^+ \) we define the function \( \mathcal{L}v(\cdot, \alpha) \) by

\[
\mathcal{L}v(l, \alpha) = r(l, \alpha) + \int v(l', \alpha)Q'(dl' \mid l, \alpha), \ l \in \tilde{X}.
\]

The **maximal reward operator** \( \mathcal{T} \) is then given by \( \mathcal{T}v(l) = \sup_{\alpha \in A} \mathcal{L}v(l, \alpha) \).

Dynamic programming suggests that \( J_{\infty} \) is a solution of the so-called optimality equation \( J_{\infty} = \mathcal{T} J_{\infty} \). In Theorem B.4 below we provide conditions which ensure that this is in fact true; this theorem is a crucial tool for our analysis of the optimal liquidation problem via MDMs in Section 4.3. To state the result we introduce several definitions.
Lemma 4.1. To establish the claim we show that the first derivatives of the vector field \( \nu \) bounded, uniformly in \( \nu \). Then the following assertions hold:

1. The value function \( J_\infty \) is continuous and lies in the set \( B_b \).
2. \( J_\infty = T J_\infty \), and \( J_\infty \) is the unique fixed point of \( T \) in \( B_b \).
3. There exists an optimal strategy. More precisely, any stationary strategy \( \{ \nu^{*,n} \} \) with \( \nu^{*,n} = \nu^* \) for every \( n \in \mathbb{N} \) and \( \nu^*(\bar{x}) \in \arg\max \{ L v(\bar{x}, \alpha) : \alpha \in \bar{A} \} \) is optimal.

B.2 Additional proofs

Lemma [7.1]. To establish the claim we show that the first derivatives of the vector field \( g \) are bounded, uniformly in \( \nu \). The components of \( \frac{\partial g}{\partial \pi^i} \) and \( \frac{\partial g}{\partial s} \) are all 0, and, using Assumption 2.1, the nonzero components of \( \frac{\partial g}{\partial \pi^i} \), \( i = 1, \ldots, K \), can be estimated by

\[
\left| \frac{\partial g^{k+3}}{\partial \pi^i} \right| = q^{ik} - \pi^k \int u^k(t, \nu, \pi, z)\eta^P(t, e_i, \nu, dz) - \pi^k \sum_{j=1}^{K} \pi^j \int u^k(t, \nu, \pi, z) \frac{\partial^k}{\partial \pi^i} \eta^P(t, e_j, \nu, dz) \leq \max_{i,k} q^{ik} + \pi^k \int u^k(t, \nu, \pi, z)\eta^P(t, e_i, \nu, dz) + \pi^k \sum_{j=1}^{K} \pi^j \int \frac{d\eta^P(t, e_i, \nu)}{d\eta^Q(z)} \frac{d\eta^P(t, e_k, \nu)}{d\eta^Q(z)} \eta^P(t, e_j, \nu, dz) \leq \max_{i,k} q^{ik} + (M^4 + M^2) \lambda^{\max}
\]
for $i \neq k$. For $i = k$ we get
\[
\left| \frac{\partial g^{i+3}}{\partial n^i} \right| = \left| q^{ii} - 2q^{ii} \int_{\mathbb{R}} u^i(t, u, \pi, z) \eta^P(t, e_i, u, dz) - \sum_{j \neq i} q^{ji} \int_{\mathbb{R}} u^j(t, u, \pi, z) \eta^P(t, e_j, u, dz) \right| \leq \max_i q^{ii}(M^4 + 3M^2)\lambda^{\max}.
\]

\[\square\]

**Lemma 4.7.** To simplify notation, without loss of generality we set $t = 0$, and we let $\rho = 0$. Since $h(0) = 0$, we get
\[
V(0, x, \{\nu^n\}) = E_{\{0, x\}} \left( \sum_{n=0}^{\infty} \left( \int_{T_{n+1} \wedge \tau} \nu^s S_s ds + 1_{\{T_n < \tau \wedge T_{n+1} \leq \tau\}} S_{\tau} h(W_{\tau}) \right) \right)
\]
\[
= \sum_{n=0}^{\infty} \left\{ E_{\{0, x\}} \left( \int_{T_{n+1} \wedge \tau} \nu^s S_s ds \mid \mathcal{G}_n \right) + E_{\{0, x\}} \left( 1_{\{T_n < \tau \wedge T_{n+1} \leq \tau\}} S_{\tau} h(W_{\tau}) \mid \mathcal{G}_n \right) \right\},
\]
where we let $\mathcal{G}_n = \sigma \{ (T_1 \wedge \tau, X_{T_1 \wedge \tau}), \ldots, (T_n \wedge \tau, X_{T_n \wedge \tau}) \}$. Using the transition kernel $Q_L$ of the sequence $\{L_n\}$ we get that
\[
E_{\{0, x\}} \left( \int_{T_{n+1} \wedge \tau} \nu^s S_s ds \mid \mathcal{G}_n \right) + E_{\{0, x\}} \left( 1_{\{T_n < \tau \wedge T_{n+1} \leq \tau\}} S_{\tau} h(W_{\tau}) \mid \mathcal{G}_n \right)
\]
\[
= 1_{\{\tau > T_n\}} \left( \int_0^{\tau} \nu^n(s, L) S_{T_n} e^{-\Lambda^\alpha_{\tau\nu}(L)} ds + e^{-\Lambda^\alpha_{\tau\nu}(L)} h(w_{\tau\nu}) S_{T_n} \right)
\]
\[
= 1_{\{\tau > T_n\}} r(L_n, \nu^n) = r(L_n, \nu^n),
\]
where the last equality follows from the fact that $L_n = \bar{\Delta}$ for $\tau \leq T_n$ and $r(\bar{\Delta}, \alpha) = 0$. \[\square\]

**Lemma 4.8.** First we estimate the reward function introduced in (4.8). Since $f \geq 0$, $e^{-pt} < 1$, and $h(w) \leq w$, we get with $\bar{\alpha}_u = \int_0^{\mu_{\max}} \nu\alpha_u(du)$,
\[
r(\bar{x}, \alpha) \leq s \int_0^{\tau} e^{-\Lambda^\alpha_u \bar{\alpha}_u} du + se^{-\Lambda^\alpha_u \bar{\alpha}_u} w_{\tau\nu}^\alpha.
\]
(B.1)
Using partial integration we get
\[
\int_0^{\tau} e^{-\Lambda^\alpha_u \bar{\alpha}_u} du = \left[ - w_{\tau\nu}^\alpha e^{-\Lambda^\alpha_u} \right]_0^{\tau} - \int_0^{\tau} \Lambda^\alpha_u e^{-\Lambda^\alpha_u} w_{\tau\nu}^\alpha du \leq w - e^{-\Lambda^\alpha_{\tau\nu}} w_{\tau\nu}^\alpha.
\]
Substituting this into (B.1) gives $r(\bar{x}, \alpha) \leq sw$. Next we estimate $Q_L b(\bar{x}, \alpha)$. Define $c_\eta := \sup \left\{ \int_{\mathbb{R}} (1 + z) \eta(t, e, 0, dz) : t \in [0, T], e \in \mathcal{E} \right\}$. It holds that
\[
Q_L b(\bar{x}, \alpha) = \int_0^{\tau} e^{\gamma(T-(u+t))} e^{-\Lambda^\alpha_u}
\]
\[
\sum_{j=1}^{K} \pi_j \int_0^{\mu_{\max}} \int_{\mathbb{R}} s(1 + z) w_{\tau\nu}^\alpha \eta(t + u, e_j, \nu, dz) \alpha_u(du) dt
\]
\[
\leq s w e^{\gamma(T-t)} c_\eta \int_0^{\tau} e^{-\gamma u} du.
\]
where we have used that \( w_n^\alpha \leq w \) and \( e^{-\Lambda_n^\alpha} < 1 \). The last term is bounded by \( b(\bar{x}) \frac{\max_i}{\gamma} \), and the MDM is contracting for \( \gamma > c_\eta \).

The following lemma is needed in the proof of Proposition 4.8.

**Lemma B.5.** Consider a continuous function \( v: \tilde{X} \rightarrow \mathbb{R} \) with \( v(\bar{x}) \leq C_v s \). Then for all \( j = 1, \ldots, K \) the mapping

\[
(t, w, s, \pi, \nu) \mapsto \int_{\mathbb{R}} v(t, s(1 + z), \pi^1(1 + u^1(t, \nu, \pi, z)), \ldots, \pi^K(1 + u^K(t, \nu, \pi, z))) \eta^j(t, \nu, dz)
\]

is continuous on \( \tilde{X} \times [0, \nu^\max] \).

**Proof.** Consider a sequence with elements \( \{s^n(1+z): z \in \text{supp}(\eta)\} \) contained in a compact subset \([\bar{z}, \bar{s}] \subset (0, \infty)\). Moreover, \( v \) is uniformly continuous on the compact set \([0, T] \times [0, w_n] \times [\bar{z}, \bar{s}] \times \mathcal{S}_K \times [0, \nu^\max] \).

Then, Assumption 4.5-(2) implies that the sequence \( \{v^n\} \) with

\[
v^n(z) := v(t_n, s_n(1 + z), \pi^1_n(1 + u^1(t_n, \nu_n, \pi_n, z)), \ldots, \pi^K_n(1 + u^K(t_n, \nu_n, \pi_n, z)))
\]

converges uniformly in \( z \in \text{supp}(\eta) \) to \( v(z) := v(t, s, \pi, \nu, z) \). Hence the following estimate holds:

\[
\left| \int_{\text{supp}(\eta)} v^n(z) \eta^j(t_n, \nu_n, dz) - \int_{\text{supp}(\eta)} v(z) \eta^j(t, \nu, dz) \right| \\
\leq \int_{\text{supp}(\eta)} |v^n(z) - v(z)| \eta^j(t_n, \nu_n, dz) + \int_{\text{supp}(\eta)} v(z) \eta^j(t_n, \nu_n, dz) - \int_{\text{supp}(\eta)} v(z) \eta^j(t, \nu, dz) \right|.
\]

Finally, the first term in (B.2) can be estimated by \( \lambda^\max \sup \{|v^n(z) - v(z)|: z \in \text{supp}(\eta)\} \), which converges to zero as \( v^n \) converges to \( v \) uniformly; the second term in (B.2) converges to zero by Assumption 4.5-(1) (continuity of the mapping \( (t, \nu) \mapsto \eta^j(t, \nu, dz) \) in the weak topology).

**References**


