Optimal Liquidation in a Finite Time Regime Switching Model with Permanent and Temporary Pricing Impact

Baojun Bian* and Nan Wu†and Harry Zheng‡

Abstract. In this paper we discuss the optimal liquidation over a finite time horizon until the exit time. The drift and diffusion terms of the asset price are general functions depending on all variables including control and market regime. There is also a local nonlinear transaction cost associated to the liquidation. The model deals with both the permanent impact and the temporary impact in a regime switching framework. The problem can be solved with the dynamic programming principle. The optimal value function is the unique continuous viscosity solution to the HJB equation and can be computed with the finite difference method.

Keywords. Optimal liquidation, permanent and temporary pricing impact, regime switching, viscosity solution.

1 Introduction

Optimal liquidation has attracted active research in recent years due to the liquidity risk. In a frictionless and competitive market an asset can be traded with any amount at any rate without affecting the market price of the asset. The optimal liquidation then becomes an optimal stopping problem which maximizes the expected liquidation value at the optimal stopping time. In an incomplete market with trading constraints on the volume and the rate and with the liquidation impact on the underlying asset price, the optimal liquidation is difficult to model and to solve.

Despite the wide recognition of the importance of the liquidity risk, there is no universal agreement on the definition of liquidity. In the academic literature the liquidity is usually defined in terms of the bid-ask spread and/or the transaction cost whereas in the practitioner literature the illiquidity is often viewed as the inability of buying and selling securities. Black [2] classifies the following four major properties of the liquidity: the immediacy of the transaction, the tightness of the spread, the resiliency of the market, and the depth of the market. The concept of liquidity can be summarized as the ability for traders to execute large trades rapidly at a price close to current market price. The liquidity risk refers to the loss stemming from the cost of liquidating a position.

Due to lack of universal agreement on the definition of liquidity, there are many different forms of mathematical characterizations. Apart from commonly used transaction cost and bid-ask spread and trading constraints (Cvitanic and Karatzas [5], Jouini [9], etc.), the other descriptions include, for example, that the order of a large investor adversely affects the stock price before being exercised (Bank and Baum [1]), that the market has a supply curve that depends on the order size of investors (Çetin et al. [3]), that trading can only happen at jump times of a Cox process (Gassiat et al. [8]), that the asset price is affected by the permanent and temporary impact of liquidation (Schied and Schöneborn [15]), etc. Once the mathematical framework is chosen for the liquidity risk one can then study specific problems such as the arbitrage pricing theory, the optimal investment and consumption, etc., see [1, 3, 5, 8, 9, 15] and references therein.

---

*Department of Mathematics, Tongji University, Shanghai 200092, China. `bianbj@tongji.edu.cn`. This work was supported in part by National Science Foundation of China(No.11371280, No.71090404).

†Department of Mathematics, Imperial College, London SW7 2BZ, UK. `nan.wu07@imperial.ac.uk`

‡Department of Mathematics, Imperial College, London SW7 2BZ, UK. `h.zheng@imperial.ac.uk`.

This paper studies the optimal liquidation in the presence of liquidity risk. There are several variations in the problem formulation in the literature, including finite or infinite time horizon, continuous trading or optimal stopping, geometric Brownian motion (GBM) asset price process or Markov modulated process, etc. Pemy et al. [13] study the optimal liquidation over an infinite time horizon. The stock price follows a GBM process with an extra term that reflects the permanent impact of liquidation on the asset price and there is no temporary impact. It is a constrained control problem which implicitly assumes that the stock holdings will never be sold out for any admissible trading strategies. The value function is the unique continuous viscosity solution to the Hamilton-Jacobi-Bellman (HJB) equation (two state variables and no time variable). In the continuous time finite state Markov chain framework Pemy and Zhang [12] study an optimal stopping problem of liquidation in finite time horizon. Pemy et al. [14] discuss the optimal liquidation over an infinite time, similar to that in [13]. The main difference is that the asset price follows a GBM process in which the drift and diffusion coefficients are determined by market regimes and the temporary impact of liquidation is reflected in the payoff function and there is no permanent impact. The assumptions and the conclusions are basically the same as those in [13].

In this paper we discuss the optimal liquidation over a finite time horizon until the exit time. The drift and diffusion coefficients μ and σ of the asset price are general functions depending on all variables including control (see (2)), which implies the trading may cause the permanent impact on the asset price. There are also nonlinear transaction costs associated to the trading through the temporary pricing impact function ϕ and the block liquidation impact function g (see (7)). The model deals with both the permanent impact and the temporary impact in a regime switching framework. We can apply the dynamic programing principle to derive the HJB equation that involves time variable as well as state variables, which makes the proofs more involved than those in [13, 14]. Our main contribution is that we show the optimal value function is the unique continuous viscosity solution to the HJB equation, which opens the way to solving the problem with the finite difference method.

The paper is organized as follows. Section 2 formulates the optimal liquidation problem and states the main results of the paper. Section 3 gives a numerical example. Section 4 proves that the optimal value function is continuous (Theorem 3). Section 5 proves that the value function is the viscosity solution to the HJB equation (Theorem 5). Section 6 proves the comparison theorem for the uniqueness of the viscosity solution (Theorem 6).

2　Model and Main Results

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $(\mathcal{F}_r)_{0 \leq r \leq T}$ be the natural filtration generated by a standard Brownian motion process W and a continuous time Markov chain process α, augmented by all P-null sets. Assume W and α are independent to each other. Assume that the Markov chain has a finite state space $\mathbb{M} = \{1, \ldots, m\}$ and is generated by the generator $Q = \{q_{ij}\}$, where $q_{ij} \geq 0$ for $i, j \in \mathbb{M}$, $j \neq i$ and $\sum_{j=1}^{m} q_{ij} = 0$ for each $i \in \mathbb{M}$. The transitional probability is given by

$$P\{\alpha(t + \Delta) = j | \alpha(t) = i\} = \begin{cases} q_{ij}\Delta + o(\Delta) & \text{if } j \neq i, \\ 1 + q_{ii}\Delta + o(\Delta) & \text{if } j = i \end{cases} \tag{1}$$

for small time interval $\Delta > 0$. The continuous time Markov chain $\alpha(r)_{0 \leq r \leq T}$ models the economic environment which affects the growth rate and the volatility of the asset price.

Let $r \in [t, T]$ be the time variable, where T is the fixed terminal time and $t \in [0, T)$ is the starting time. Let $S(r)_{0 \leq r \leq T}$ denote the stock price and $X(r)_{0 \leq r \leq T}$ the number of shares of stock. Let $u(r)_{0 \leq r \leq T}$ denote the rate of selling the stock, which is a control variable decided by the trader. We call $u = \{u(r)\}_{0 \leq r \leq T}$ is admissible if it is progressively measurable and $u(r) \in U$ for a compact set $U \subset [0, \infty)$ for all $t \leq r \leq T$. The stock price $S(r)$ follows a stochastic differential equation with regime switching

$$dS(r) = \mu(r, S(r), u(r), \alpha(r))dr + \sigma(r, S(r), u(r), \alpha(r))dW(r) \tag{2}$$

and the stock holding $X(r)$ follows the dynamics

$$dX(r) = -u(r)dr.$$

Since the drift and the diffusion terms of S are affected by the trading strategy u there is the permanent impact of liquidation on the asset price. Such an impact may be negligible for a small trader (when u is small) but can be significant for a large trader (when u is large). We implicitly assume that the asset price $S(r)$ is positive for all $t \leq r \leq T$. A sufficient condition that guarantees this is that S follows a geometric Brownian motion process with drift and diffusion coefficients depending on time, control and Markov state. We denote by K some generic positive constant which may take different values at different places.

Assumption 1. *Functions $f = \mu, \sigma$ satisfy, for all $t, s \in [0, T]$, $x, y \in \mathbb{R}$, $v \in [0, \infty)$ and $\ell \in \mathbb{M}$, that*

$$|f(t, x, v, \ell) - f(s, y, v, \ell)| \leq K(|t - s| + |x - y|) \quad and \quad |f(t, x, v, \ell)| \leq K(1 + |x|). \tag{3}$$

It can be shown, with Assumption 1, that for any admissible control process $u \in \mathcal{U}$ and any initial values $(t, s, \ell) \in [0, T) \times (0, \infty) \times \mathbb{M}$, there exists a unique solution, denoted by $\{S^u_{t,s,\ell}(r), t \leq r \leq T\}$, to equation (2), and that the following inequalities hold:

$$E\left[\sup_{r \in [t,T]} |S^u_{t,s,\ell}(r)|^p\right] \leq K(1 + s^p), \quad p = 1, 2 \tag{4}$$

$$E\left[|S^u_{t,s,\ell}(t_2) - S^u_{t,s,\ell}(t_1)|\right] \leq K(1 + s)|t_2 - t_1|^{1/2}, \quad t_1, t_2 \in [t, T] \tag{5}$$

$$E\left[\sup_{r \in [t,T]} |S^u_{t,s_1,\ell}(r) - S^u_{t,s_2,\ell}(r)|\right] \leq K|s_1 - s_2|, \quad s_1, s_2 \in (0, \infty). \tag{6}$$

The proofs of (4), (5) and (6) can be found in Mao and Yuan [10] with some minor changes to include control processes, see [10], Theorem3.23, Theorem 3.24 and Lemma 3.3.

Similarly, $\{X^u_{t,x}(r), t \leq r \leq T\}$ denotes the stock holding and $\{\alpha_{t,\ell}(r), t \leq r \leq T\}$ the Markov chain process.

Suppose a trader starts from time t, endowed with initial values $(X(t), S(t), \alpha(t)) = (x, s, \ell) \in (0, \infty) \times (0, \infty) \times \mathbb{M}$. Define a stopping time

$$\tau_0 = \inf\{r \geq t : X^u_{t,x}(r) = 0\} \wedge T.$$

This is the first time that $X^u_{t,x}(r)$ exits from $(0, \infty)$ before or at time T. Since the model is to study the liquidation strategy, the trader is only allowed to sell stock without buying back. When the number of shares reaches zero before time T the liquidation stops. Otherwise, it stops at time T.

The expected discounted total payoff associated with a strategy $u \in \mathcal{U}$ is defined by

$$J(t, x, s, \ell; u) = E\left[\int_t^{\tau_0} e^{-\beta(r-t)} \phi(u(r)) S^u_{t,s,\ell}(r) dr + e^{-\beta(\tau_0 - t)} g(X(\tau_0)) S(\tau_0)\right], \tag{7}$$

where $\beta > 0$ is a discount rate, ϕ a function measuring the temporary liquidation effect, g a function measuring the block liquidation effect, and E the conditional expectation given the information set $\mathcal{F}_t$ which is equivalent to given $X(t) = x$, $S(t) = s$ and $\alpha(t) = \ell$ since the model is Markov. The first term is the expected discounted accumulated cash value from the stock liquidation and the second term is the expected discounted cash value from the block liquidation at time T for any remaining shares of the stock.

Assumption 2. *Functions $f = \phi, g$ are continuous concave increasing on $\mathbb{R}$ and satisfy $f(0) = 0$ and $f'(0) = 1$. Furthermore, function g is continuously differentiable and satisfies, for all $x, y \in \mathbb{R}$, that*

$$|g(x) - g(y)| \leq K|x - y| \quad and \quad |g'(x) - g'(y)| \leq K|x - y|.$$

3

Note that in a completely liquid market $\phi(v) = v$ and $g(x) = x$, and that $f(0) = 0$ and $f'(0) = 1$ imply $f(x)$ is approximately equal to x when x is close to 0, which means when the trading rate u is small or the amount of stock X is small then there is essentially no transaction cost and the liquidity impact can be ignored. The objective of the trader is to maximize the expected discounted revenue from stock liquidation. The value function is defined by

$$V(t, x, s, \ell) = \sup_{u \in \mathcal{U}} J(t, x, s, \ell; u).$$

For $v \in U$ define operators $\mathcal{L}^v$ and $\mathcal{Q}$ of the value function V by

$$\mathcal{L}^v V(t, x, s, \ell) = -v \frac{\partial V}{\partial x}(t, x, s, \ell) + \mu(t, s, v, \ell) \frac{\partial V}{\partial s}(t, x, s, \ell) + \frac{1}{2} \sigma^2(t, s, v, \ell) \frac{\partial^2 V}{\partial s^2}(t, x, s, \ell),$$

and

$$\mathcal{Q}V(t, x, s, \ell) = \sum_{j \neq \ell} q_{\ell j} \left(V(t, x, s, j) - V(t, x, s, \ell) \right).$$

The HJB equation for the optimal control problem is, for $(t, x, s, \ell) \in [0, T) \times (0, \infty) \times (0, \infty) \times \mathbb{M}$,

$$\beta V(t, x, s, \ell) - \frac{\partial V}{\partial t}(t, x, s, \ell) - \sup_{v \in U} \{ \mathcal{L}^v V(t, x, s, \ell) + \phi(v)s \} - \mathcal{Q}V(t, x, s, \ell) = 0, \tag{8}$$

with the boundary condition

$$V(t, 0, s, \ell) = 0$$

and the terminal condition

$$V(T, x, s, \ell) = g(x)s.$$

It is easy to check that the value function is an increasing function with respect to the asset price and the stock holding. It also has the following continuity property.

Theorem 3. *Assume Assumptions 1 and 2. Then the value function $V(\cdot, \cdot, \cdot, \ell)$ is continuous on $[0, T] \times [0, \infty) \times (0, \infty)$ for $\ell \in \mathbb{M}$.*

Since we do not know if the value function V is continuously differentiable and cannot discuss the solution to the HJB equation in the classical sense, we need to introduce the concept of the viscosity solution to the HJB equation.

Definition 4. *A system of continuous functions $V = \{V(\cdot, \cdot, \cdot, \ell)\}_{\ell \in \mathbb{M}}$ on $[0, T) \times (0, \infty) \times (0, \infty)$ is a viscosity subsolution (resp. supersolution) of the HJB equation (8) if, for any fixed $\ell \in \mathbb{M}$, $\varphi \in C^{1,1,2}([0, T) \times (0, \infty) \times (0, \infty))$ and $(\bar{t}, \bar{x}, \bar{s}) \in [0, T) \times (0, \infty) \times (0, \infty)$ such that $V(t, x, s, \ell) - \varphi(t, x, s)$ attains its maximum (resp. minimum) at $(\bar{t}, \bar{s}, \bar{x})$, we have*

$$\beta \varphi(\bar{t}, \bar{x}, \bar{s}) - \frac{\partial \varphi}{\partial t}(\bar{t}, \bar{x}, \bar{s}) - \sup_{v \in U} \{ \mathcal{L}^v \varphi(\bar{t}, \bar{x}, \bar{s}) + \phi(v)\bar{s} \} - \mathcal{Q}V(\bar{t}, \bar{x}, \bar{s}, \ell) \leq 0; \quad (resp. \geq 0). \tag{9}$$

The system of continuous functions V is a viscosity solution if it is both a viscosity subsolution and a viscosity supersolution.

We have the following result for the value function.

Theorem 5. *Assume Assumptions 1 and 2. Then the value function V is a viscosity solution to the HJB equation (8).*

One in general has to use some numerical scheme to find the value function. To ensure the numerical solution to the HJB equation is indeed the value function one has to show that the value function is the unique viscosity solution to the HJB equation, which can be achieved by the following comparison theorem.

Theorem 6. *Assume Assumptions 1 and 2. Let U be a viscosity subsolution and V a viscosity supersolution to the HJB equation* (8) *and satisfy the polynomial growth condition and $U(T, x, s, \ell) \le V(T, x, s, \ell)$ for all $(x, s, \ell) \in (0, \infty) \times (0, \infty) \times \mathbb{M}$. Then $U \le V$ on $[0, T) \times (0, \infty) \times (0, \infty) \times \mathbb{M}$.*

The proofs of Theorems 3, 5, and 6 are given in Sections 4, 5, and 6, respectively. The proofs are technical and lengthy as one would expect with the viscosity solution method. The further complication in the proofs over the standard diffusion model is that we need to deal with the Markov chain process α and its relation with the diffusion process S.

3 A Numerical Example

In this section we give a numerical example to find the approximation of the value function and the optimal selling strategy. The finite difference method is one of the most common approximation schemes for viscosity solutions due to its well-known consistency, stability, convergence analysis, in particular in the presence of the monotonicity property, see [7] for numerical solutions of HJB equation and [12] for a regime switching optimal stopping problem which results in a system of HJB variational inequalities. We may apply the numerical scheme of [12] to solve our optimal liquidation problem. The numerical example is to provide a snapshot of the optimal trading strategy at a given specific time.

Assume that there are only two regimes. Regime 1 represents the strong economy and regime 2 the weak economy and assume that the stock price $S(r)$ follows a GBM process with $\mu(r, s, u, \alpha) = \mu(\alpha)s$ and $\sigma(r, s, u, \alpha) = \sigma(\alpha)s$. Define variables $z = \log s$ and $\tau = T - t$ and a function $W(\tau, x, z, \ell) = V(t, x, s, \ell)$. The HJB equation (8) becomes

$$\beta W(\tau, x, z, \ell) + \frac{\partial W}{\partial \tau}(\tau, x, z, \ell) - \sup_{v \in U} \left\{ -v \frac{\partial W}{\partial x}(\tau, x, z, \ell) + \mu(\ell) \frac{\partial W}{\partial z}(\tau, x, z, \ell) \right.$$

$$\left. + \frac{1}{2}\sigma^2(\ell) \left(\frac{\partial^2 W}{\partial z^2}(\tau, x, z, \ell) - \frac{\partial W}{\partial z}(\tau, x, z, \ell) \right) + \phi(v)e^z \right\} - \mathcal{Q}W(\tau, x, z, \ell) = 0, \qquad (10)$$

with the boundary condition $W(\tau, 0, z, \ell) = 0$ and the terminal condition $W(0, x, z, \ell) = g(x)e^z$.

To approximate the solution to (10) we discretize variables τ, x and z with stepsizes $\Delta\tau, \Delta x, \Delta z$, respectively. The value of W at a grid point (τ_n, x_i, z_j) in the regime ℓ is denoted by $W_{i,j}^n(\ell)$. The derivatives of W are approximated by $W_\tau = (W_{i,j}^{n+1}(\ell) - W_{i,j}^n(\ell))/\Delta\tau$, $W_x = (W_{i+1,j}^n(\ell) - W_{i-1,j}^n(\ell))/(2\Delta x)$, $W_z = (W_{i,j+1}^n(\ell) - W_{i,j-1}^n(\ell))/(2\Delta z)$, and $W_{zz} = (W_{i,j+1}^n(\ell) + W_{i,j-1}^n(\ell) - 2W_{i,j}^n(\ell))/\Delta z^2$. Discretizing equation (10) and rearranging the terms, we have

$$W_{i,j}^{n+1}(\ell) = \Delta\tau \left[\left(-\beta + \frac{1}{\Delta\tau} - q_{\ell\ell'} - \frac{\sigma(\ell)^2}{\Delta z^2} \right) W_{i,j}^n(\ell) + \left(\frac{\mu(\ell) - \frac{1}{2}\sigma(\ell)^2}{2\Delta z} + \frac{\sigma(\ell)^2}{2\Delta z^2} \right) W_{i,j+1}^n(\ell) \right.$$

$$+ \left(-\frac{\mu(\ell) - \frac{1}{2}\sigma(\ell)^2}{2\Delta z} + \frac{\sigma(\ell)^2}{2\Delta z^2} \right) W_{i,j-1}^n(\ell) + q_{\ell\ell'} W_{i,j}^n(\ell')$$

$$\left. + \sup_{v \in U} \left\{ -v \frac{W_{i+1,j}^n(\ell) - W_{i-1,j}^n(\ell)}{2\Delta x} + \phi(v)e^z \right\} \right], \qquad (11)$$

where $\ell, \ell' = 1, 2$ and $\ell \ne \ell'$. Assume that the temporary liquidation impact function is given by

$$\phi(v) = \frac{1}{\alpha}(1 - e^{-\alpha v}),$$

where $\alpha > 0$, and the block liquidation impact function is given by

$$g(x) = \begin{cases} x, & \text{if } 0 \le x \le 5, \\ -0.01x^2 + 1.1x - 0.25, & \text{if } 5 < x \le 15, \\ 10 + 0.8(x - 10), & \text{if } 15 < x \le 40, \\ -0.0075x^2 + 1.4x - 10, & \text{if } 40 < x \le 60, \\ 42 + 0.5(x - 50), & \text{if } x > 60. \end{cases}$$

Functions ϕ and g satisfy Assumption 2. In fact, g is constructed as a smooth approximation to a function f defined by

$$f(x) = \begin{cases} x, & \text{if } 0 \leq x \leq 10, \\ 10 + 0.8(x - 10), & \text{if } 10 < x \leq 50, \\ 42 + 0.5(x - 50), & \text{if } x > 50. \end{cases}$$

Function f captures the block liquidation effect at time T but is not differentiable at $x = 10$ and 50 and does not satisfy Assumption 2.

Data used for numerical tests are $\alpha = 0.005$, $\beta = 0.01$, $\{\mu(1), \mu(2)\} = \{0.3, -0.1\}$, $\{\sigma(1), \sigma(2)\} = \{0.2, 0.4\}$, $q_{12} = 0.5$, $q_{21} = 1$, $v \in U = [0, 100]$, $t \in [0, 1]$, $x \in [0, 100]$, $s \in [e^{-1}, e^2]$.
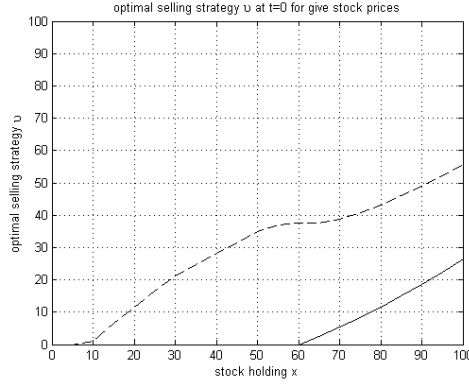
(a) $\ell = 1$

Figure 1: The optimal control at time $t = 0$ against stock holding x. The solid line is for regime 1 and the dashed line for regime 2.

Figure 1 demonstrates the relationship between the optimal selling strategy and the stock holding. It is clear that the more shares one holds, the sooner and the more one wants to sell to avoid the potential large transaction cost during the whole period. The market regime determines at what level of stock holding one should start to sell. In a rising market (regime 1) the trader is willing to keep the stock for a longer period in the hope for a higher price, which results in a lower optimal selling rate, whereas in a falling market (regime 2) the trader wants to liquidate the stock quickly to avoid a lower price. This is consistent with the general market phenomenon. The optimal trading strategy is independent of initial asset price in the numerical test, which is not surprising as the asset price follows a GBM process and depends on the initial asset price linearly. In general, the optimal trading strategy should also depend on the asset price. The particular shape of the curve in Figure 1 is determined by the tradeoff between function ϕ that captures the liquidity effect from 'flow' trading and function g that reflects the transaction cost for the block liquidation at the terminal time. Note that if there is no temporary pricing impact on liquidation, i.e., $\phi(v) = v$, then the optimal liquidation strategy is a "bang-bang" control with either no trading $v = 0$ or selling at maximum rate $v = 100$ due to the linear dependence of control v in the Hamiltonian function.

4 Proof of Theorem 3

We first convert the original control problem into a problem without terminal bequest function. Since function g is continuously differentiable, we can apply Dynkin's formula to $e^{\beta(\tau_0 - t)} g(X_{t,x}^u(\tau_0)) S_{t,s,\ell}^u(\tau_0)$ and rewrite the total payoff J as

$$J(t, x, s, \ell; u) = g(x)s + E\left[\int_t^{\tau_0} L(r, X_{t,x}^u(r), S_{t,s,\ell}^u(r), u(r), \alpha_{t,\ell}(r))dr\right],$$

6

where
$$L(r, x, s, v, \alpha) = e^{-\beta(r-t)} \left[\phi(v)s - \beta g(x)s - v g'(x)s + \mu(r, s, v, \alpha)g(x) \right].$$

Define a new value function by
$$\widetilde{V}(t, x, s, \ell) = \sup_{u \in \mathcal{U}} E \left[\int_t^{\tau_0} L\left(r, X_{t,x}^u(r), S_{t,s,\ell}^u(r), u(r), \alpha_{t,\ell}(r)\right) dr \right].$$

Since $V(t, x, s, \ell) = \widetilde{V}(t, x, s, \ell) + g(x)s$, we know $V(t, x, s, \ell)$ is continuous as long as $\widetilde{V}(t, x, s, \ell)$ is continuous. From now on in this section we work on the value function $\widetilde{V}$.

To prove the continuity of $\widetilde{V}$ we need to define some perturbed problems and show their corresponding value functions are continuous and converge quasi-uniformly to $\widetilde{V}$, which establishes Theorem 3.

For $0 < \epsilon < 1$ define the stopping time
$$\tau_\epsilon = \inf\{r \geq t : X_{t,x}^u(r) = -\epsilon\} \wedge T,$$

which is the first time $X_{t,x}^u(r)$ exits from $(-\epsilon, \infty)$. A control process $u = \{u(r)\}_{0 \leq r \leq T}$ is admissible if it is progressively measurable and $u(r) \in U_\epsilon(X_{t,x}^u(r))$, where $U_\epsilon(x) = U$ if $x \geq 0$ and $U_\epsilon(x) = \hat{U}$, a compact subset of U in $(0, \infty)$, if $x < 0$. The key here is to rule out zero from the compact set $\hat{U}$ after $X(r)$ reaches zero. The admissible control set is the collection of all admissible controls, denoted by $\mathcal{U}_\epsilon$. Note that when we only look at the control process before τ_0, the two admissible control sets, $\mathcal{U}$ and $\mathcal{U}_\epsilon$, are the same.

To simplify the notation denote by
$$L_{t,x,s,\ell}^u(r) := L\left(r, X_{t,x}^u(r), S_{t,s,\ell}^u(r), u(r), \alpha_{t,\ell}(r)\right).$$

Since U is a compact set in $[0, \infty)$, say $[0, N]$, we know that $X_{t,x}^u(r) \in [x - NT, x]$ for $t \leq r \leq T$, which implies that $|g(X_{t,x}^u(r))|$ and $|g'(X_{t,x}^u(r))|$ are bounded by some constant K_x depending on x due to continuity of g and g'. Assumptions 1 and 2 imply that, for $t \leq r \leq T$,
$$|L_{t,x,s,\ell}^u(r)| \leq K_x \left(1 + S_{t,s,\ell}^u(r)\right) \tag{12}$$

and
$$\left| L_{t,x_1,s_1,\ell}^u(r) - L_{t,x_2,s_2,\ell}^u(r) \right| \leq K_{x_1} |S_{t,s_1,\ell}^u(r) - S_{t,s_2,\ell}^u(r)| + K \left(1 + S_{t,s_2,\ell}^u(r)\right) |x_1 - x_2| \tag{13}$$

for some constant K_{x_1} depending on x_1.

Remark 7. In the proof we need to estimate $|L_{t,x,s,\ell}^u(r)|$ several times for different x. One case is that $x = -\epsilon$ for $0 < \epsilon < 1$. Then $X_{t,x}^u(r) \in [-1 - NT, 0]$ and constant K_x can be replaced by a generic constant K independent of x. The other case is that x is within a distance d of another point x_1. Then $X_{t,x}^u(r) \in [x_1 - d - NT, x_1 + d]$ and constant K_x can be written as K_{x_1} depending on x_1 for all such x.

For $\epsilon \in (0, 1)$ define a perturbed value function by
$$\widetilde{V}^\epsilon(t, x, s, \ell) = \sup_{u \in \mathcal{U}_\epsilon} E \left[\int_t^{\tau_\epsilon} L_{t,x,s,\ell}^u(r) dr \right].$$

For $\rho > 0$ define an auxiliary function
$$\Gamma_{t,x}^{\epsilon,\rho,u}(r) = \exp\left(-\frac{1}{\rho} \left(X_{t,x}^u(r) + \epsilon \right)^- \right),$$

where $x^- = \max(0, -x)$. Clearly, we have $\Gamma_{t,x}^{\epsilon,\rho,u}(r) \leq 1$ and, by the definition of the stopping time τ_ϵ, $\Gamma_{t,x}^{\epsilon,\rho,u}(r) = 1$ for $r \in [t, \tau_\epsilon]$. The auxiliary value function $\widetilde{V}^{\epsilon,\rho}$ is defined by
$$\widetilde{V}^{\epsilon,\rho}(t, x, s, \ell) = \sup_{u \in \mathcal{U}_\epsilon} \widetilde{J}^{\epsilon,\rho}(t, x, s, \ell; u) := E \left[\int_t^T \Gamma_{t,x}^{\epsilon,\rho,u}(r) L_{t,x,s,\ell}^u(r) dr \right].$$

From (12) and (4) we have that
$$|\widetilde{V}^{\epsilon,\rho}(t, x, s, \ell)| \leq K_x(1 + s). \tag{14}$$

7

Lemma 8. $\widetilde{V}^{\epsilon,\rho}(t,x,s,\ell)$ converges to $\widetilde{V}(t,x,s,\ell)$ quasi-uniformly as $\rho \to 0$ and $\epsilon \to 0$.

Proof. Step 1. Fix a point $(t,x,s) \in [0,T) \times \{-\epsilon\} \times (0,\infty)$. Since $X_{t,x}^u(t) = -\epsilon$ we have $\tau_\epsilon = t$ and for $r > t$ the admissible control $u(r)$ is in a compact set $U_\epsilon(x) := [N_0, N] \subset U$ with $N_0 > 0$, which implies that $X_{t,-\epsilon}^u(r) < -\epsilon$ and

$$\exp\left(-\frac{N}{\rho}(r-t)\right) \leq \Gamma_{t,-\epsilon}^{\epsilon,\rho,u}(r) = \exp\left(-\frac{1}{\rho}\int_t^r u(s)ds\right) \leq \exp\left(-\frac{N_0}{\rho}(r-t)\right) \tag{15}$$

and $\lim_{\rho\to 0}\Gamma_{t,-\epsilon}^{\epsilon,\rho,u}(r) = 0$. (15), (12) and (4) imply that, also noting Remark 7,

$$\widetilde{J}^{\epsilon,\rho}(t,-\epsilon,s,\ell;u) \leq K\int_t^T e^{-\frac{N_0}{\rho}(r-t)}\left(1 + E\left[S_{t,s,\ell}^u(r)\right]\right)dr \leq K(1+s)\frac{\rho}{N_0}\left(1 - e^{-\frac{N_0}{\rho}T}\right).$$

Similarly, we have

$$\widetilde{J}^{\epsilon,\rho}(t,-\epsilon,s,\ell;u) \geq -K(1+s)\frac{\rho}{N}\left(1 - e^{-\frac{N}{\rho}T}\right).$$

Combining the above two inequalities and taking the supremum, we have

$$-K(1+s)\frac{\rho}{N}\left(1 - e^{-\frac{N}{\rho}T}\right) \leq \widetilde{V}^{\epsilon,\rho}(t,-\epsilon,s,\ell) \leq K(1+s)\frac{\rho}{N_0}\left(1 - e^{-\frac{N_0}{\rho}T}\right).$$

Applying the dynamic programming principle (see [4]), for $(t,x,s,\ell) \in [0,T) \times [0,\infty) \times (0,\infty) \times \mathbb{M}$, we have

$$\widetilde{V}^{\epsilon,\rho}(t,x,s,\ell) = \sup_{u\in\mathcal{U}_\epsilon} E\left[\int_t^{\tau_\epsilon} L_{t,x,s,\ell}^u(r)dr + e^{-\beta(\tau_\epsilon-t)}\widetilde{V}^{\epsilon,\rho}\left(\tau_\epsilon, -\epsilon, S_{t,s,\ell}^u(\tau_\epsilon), \alpha_{t,\ell}(\tau_\epsilon)\right)\right]$$

$$\leq \sup_{u\in\mathcal{U}_\epsilon} E\left[\int_t^{\tau_\epsilon} L_{t,x,s,\ell}^u(r)dr + K(1 + S_{t,s,\ell}^u(\tau_\epsilon))\frac{\rho}{N_0}\left(1 - e^{-\frac{N_0}{\rho}T}\right)\right]$$

$$\leq \widetilde{V}^\epsilon(t,x,s,\ell) + K(1+s)\frac{\rho}{N_0}\left(1 - e^{-\frac{N_0}{\rho}T}\right).$$

Similarly, we have

$$\widetilde{V}^{\epsilon,\rho}(t,x,s,\ell) \geq \widetilde{V}^\epsilon(t,x,s,\ell) - K(1+s)\frac{\rho}{N}\left(1 - e^{-\frac{N}{\rho}T}\right).$$

The above two inequalities imply that $\widetilde{V}^{\epsilon,\rho}(t,x,s,\ell)$ converges to $\widetilde{V}^\epsilon(t,x,s,\ell)$ quasi-uniformly as $\rho \to 0$, independent of ϵ.

Step 2. By the definition of the perturbed value function, the Cauchy-Schwartz inequality, (12) and (4), we have

$$\widetilde{V}^\epsilon(t,x,s,\ell) = \sup_{u\in\mathcal{U}_\epsilon} E\left[\int_t^{\tau_0} L_{t,x,s,\ell}^u(r)dr + \int_{\tau_0}^{\tau_\epsilon} L_{t,x,s,\ell}^u(r)dr\right]$$

$$\leq \widetilde{V}(t,x,s,\ell) + \sup_{u\in\mathcal{U}_\epsilon} E\left[\int_t^T \mathbb{1}_{\{\tau_0 < r < \tau_\epsilon\}} L_{t,x,s,\ell}^u(r)dr\right]$$

$$\leq \widetilde{V}(t,x,s,\ell) + \sup_{u\in\mathcal{U}_\epsilon} \sqrt{E[\tau_\epsilon - \tau_0]}\sqrt{E\left[\int_t^T L_{t,x,s,\ell}^u(r)^2 dr\right]}$$

$$\leq \widetilde{V}(t,x,s,\ell) + K_x(1+s)\left(\frac{\epsilon}{N_0}\right)^{1/2}.$$

for some constant K_x depending on x. Similarly, we have

$$\widetilde{V}^\epsilon(t,x,s,\ell) \geq \widetilde{V}(t,x,s,\ell) - K_x(1+s)\left(\frac{\epsilon}{N_0}\right)^{1/2}.$$

As $\epsilon \to 0$, $\widetilde{V}^\epsilon(t,x,s,\ell)$ converges to $\widetilde{V}(t,x,s,\ell)$ quasi-uniformly. Combining the results of Steps 1 and 2, we conclude that $\widetilde{V}^{\epsilon,\rho}(t,x,s,\ell)$ converges to $\widetilde{V}(t,x,s,\ell)$ quasi-uniformly as $\rho \to 0$ and $\epsilon \to 0$. $\square$

Lemma 9. $\widetilde{V}^{\epsilon,\rho}(\cdot,\cdot,\cdot,\ell)$ *is continuous on* $[0,T] \times [0,\infty) \times (0,\infty)$ *for* $\ell \in \mathbb{M}$ *and arbitrary constants* $\epsilon > 0$ *and* $\rho > 0$.

Proof. Step 1. Let $(x_1, s_1), (x_2, s_2) \in [0,\infty) \times (0,\infty)$ satisfying $|x_2 - x_1| \le 1$ and $|s_2 - s_1| \le 1$ and $t \in [0,T]$ and $\ell \in \mathbb{M}$. Consider the auxiliary value functions $\widetilde{V}^{\epsilon,\rho}(t, x_1, s_1, \ell)$ and $\widetilde{V}^{\epsilon,\rho}(t, x_2, s_2, \ell)$.

Since $|e^{-a} - e^{-b}| \le |a - b|$ for any $a, b \ge 0$, we have

$$\left| \Gamma_{t,x_1}^{\epsilon,\rho,u}(r) - \Gamma_{t,x_2}^{\epsilon,\rho,u}(r) \right| \le \frac{1}{\rho} \left| (X_{t,x_1}^u(r) + \epsilon)^- - (X_{t,x_2}^u(r) + \epsilon)^- \right| \le \frac{1}{\rho} |x_1 - x_2|. \tag{16}$$

By the definition of $\widetilde{V}^{\epsilon,\rho}$ and the relation $|\sup A - \sup B| \le \sup |A - B|$ we have

$$\left| \widetilde{V}^{\epsilon,\rho}(t, x_1, s_1, \ell) - \widetilde{V}^{\epsilon,\rho}(t, x_2, s_2, \ell) \right|$$

$$\le \sup_{u \in \mathcal{U}_\epsilon} E \left[\int_t^T \left| \Gamma_{t,x_1}^{\epsilon,\rho,u}(r) L_{t,x_1,s_1,\ell}^u(r) - \Gamma_{t,x_2}^{\epsilon,\rho,u}(r) L_{t,x_2,s_2,\ell}^u(r) \right| dr \right]$$

$$\le \sup_{u \in \mathcal{U}_\epsilon} E \left[\int_t^T \left(\left| L_{t,x_1,s_1,\ell}^u(r) - L_{t,x_2,s_2,\ell}^u(r) \right| + \left| L_{t,x_2,s_2,\ell}^u(r) \left(\Gamma_{t,x_1}^{\epsilon,\rho,u}(r) - \Gamma_{t,x_2}^{\epsilon,\rho,u}(r) \right) \right| \right) dr \right]$$

$$\le K_{x_1} |s_1 - s_2| + K(1 + s_2)|x_1 - x_2| + \frac{1}{\rho}|x_1 - x_2| K_{x_1}(1 + s_2)$$

$$\le K_{x_1,s_1}(|x_1 - x_2| + |s_1 - s_2|), \tag{17}$$

where K_{x_1,s_1} is some constant depending on x_1 and s_1. In the second last inequality we have used (13), (6), (4), (12), (16) and Remark 7. This shows that the auxiliary value function $\widetilde{V}^{\epsilon,\rho}(t, x, s, \ell)$ is continuous in (x, s), uniformly in t.

Step 2. We prove that the auxiliary value function $\widetilde{V}^{\epsilon,\rho}$ is continuous in t. Let $0 \le t_1 < t_2 \le T$ and $(x, s, \ell) \in [0,\infty) \times (0,\infty) \times \mathbb{M}$. By the dynamic programming principle, for any $\delta > 0$, there exists an admissible control $u_\delta \in \mathcal{U}_\epsilon$ such that

$$\widetilde{V}^{\epsilon,\rho}(t_1, x, s, \ell) - \delta$$

$$\le E \left[\int_{t_1}^{t_2} \Gamma_{t_1,x}^{\epsilon,\rho,u_\delta}(r) L_{t_1,x,s,\ell}^{u_\delta}(r) dr + e^{-\beta(t_2 - t_1)} \widetilde{V}^{\epsilon,\rho}(t_2, X_{t_1,x}^{u_\delta}(t_2), S_{t_1,s,\ell}^{u_\delta}(t_2), \alpha_{t_1,\ell}(t_2)) \right]$$

$$\le \widetilde{V}^{\epsilon,\rho}(t_1, x, s, \ell).$$

Rearranging the above inequalities, we have

$$\left| \widetilde{V}^{\epsilon,\rho}(t_1, x, s, \ell) - \widetilde{V}^{\epsilon,\rho}(t_2, x, s, \ell) \right| - \delta$$

$$\le \left| E \left[\int_{t_1}^{t_2} \Gamma_{t_1,x}^{\epsilon,\rho,u_\delta}(r) L_{t_1,x,s,\ell}^{u_\delta}(r) dr + e^{-\beta(t_2 - t_1)} \widetilde{V}^{\epsilon,\rho}(t_2, X_{t_1,x}^{u_\delta}(t_2), S_{t_1,s,\ell}^{u_\delta}(t_2), \alpha_{t_1,\ell}(t_2)) \right] - \widetilde{V}^{\epsilon,\rho}(t_2, x, s, \ell) \right|$$

$$\le E \left[\int_{t_1}^{t_2} \left| L_{t_1,x,s,\ell}^{u_\delta}(r) \right| dr \right] + E \left[\left| e^{-\beta(t_2 - t_1)} \widetilde{V}^{\epsilon,\rho}(t_2, X_{t_1,x}^{u_\delta}(t_2), S_{t_1,s,\ell}^{u_\delta}(t_2), \ell) - \widetilde{V}^{\epsilon,\rho}(t_2, x, s, \ell) \right| \right]$$

$$+ E \left[\left| \widetilde{V}^{\epsilon,\rho}(t_2, X_{t_1,x}^{u_\delta}(t_2), S_{t_1,s,\ell}^{u_\delta}(t_2), \alpha_{t_1,\ell}(t_2)) - \widetilde{V}^{\epsilon,\rho}(t_2, X_{t_1,x}^{u_\delta}(t_2), S_{t_1,s,\ell}^{u_\delta}(t_2), \ell) \right| \right]$$

$$= I_1 + I_2 + I_3.$$

(12) and (4) imply that

$$I_1 \le K_x(1 + s)(t_2 - t_1).$$

(14) and Remark 7 imply that

$$E \left[\widetilde{V}^{\epsilon,\rho} \left(t_2, X_{t_1,x}^{u_\delta}(t_2), S_{t_1,s,\ell}^{u_\delta}(t_2), \alpha_{t_1,\ell}(t_2) \right) \right] \le E \left[K_x \left(1 + S_{t_1,s,\ell}^{u_\delta}(t_2) \right) \right] \le K_{x,s}$$

9

for some constant $K_{x,s}$ depending on x and s. Noting that the term inside the expectation of I_3 is zero when $\alpha_{t_1,\ell}(t_2) = \ell$, using Cauchy-Schwartz inequality and combining the above inequality, we have

$$I_3 \leq K_{x,s} \sqrt{P\left[\alpha_{t_1,\ell}(t_2) \neq \ell\right]}.$$

Using (17) and (5), we have

$$\begin{aligned}
I_2 &\leq E\left[\left|\widetilde{V}^{\epsilon,\rho}\left(t_2, X_{t_1,x}^{u_\delta}(t_2), S_{t_1,s,\ell}^{u_\delta}(t_2), \ell\right) - \widetilde{V}^{\epsilon,\rho}\left(t_2, x, s, \ell\right)\right| + \left|(e^{-\beta(t_2-t_1)} - 1)\widetilde{V}^{\epsilon,\rho}(t_2, x, s, \ell)\right|\right] \\
&\leq K_{x,s}\left(E\left[\left|X_{t_1,x}^{u_\delta}(t_2) - x\right|\right] + E\left[\left|S_{t_1,s,\ell}^{u_\delta}(t_2) - s\right|\right]\right) + E\left[\widetilde{V}^{\epsilon,\rho}(t_2, x, s, \ell)\right]\left|e^{-\beta(t_2-t_1)} - 1\right| \\
&\leq K_{x,s}\left((t_2 - t_1) + (t_2 - t_1)^{1/2} + \left|e^{-\beta(t_2-t_1)} - 1\right|\right)
\end{aligned}$$

for some constant $K_{x,s}$ depending on x, s. The above estimates for I_1, I_2, I_3 show that they all tend to 0 as $t_2 - t_1$ tends to 0, independent of δ and control u_δ but dependent on x and s. Therefore,

$$\left|\widetilde{V}^{\epsilon,\rho}(t_1, x, s, \ell) - \widetilde{V}^{\epsilon,\rho}(t_2, x, s, \ell)\right| - \delta \to 0 \text{ as } t_2 - t_1 \to 0.$$

The arbitrariness of δ confirms that $\widetilde{V}^{\epsilon,\rho}(t, x, s, \ell)$ is continuous in t.

Combining the results of Steps 1 and 2, we conclude that $\widetilde{V}^{\epsilon,\rho}(\cdot, \cdot, \cdot, \ell)$ is continuous in (t, x, s) for each $\ell \in \mathbb{M}$. $\qquad \square$

By Lemmas 8 and 9, the auxiliary value function $\widetilde{V}^{\epsilon,\rho}(t, x, s, \ell)$ converges quasi-uniformly to the value function $\widetilde{V}(t, x, s, \ell)$ as $\epsilon \to 0$ and $\rho \to 0$ and $\widetilde{V}^{\epsilon,\rho}(t, x, s, \ell)$ is continuous in (t, x, s), which shows that $\widetilde{V}(t, x, s, \ell)$ is continuous on $[0, T] \times [0, \infty) \times (0, \infty)$ for each $\ell \in \mathbb{M}$. We have proved Theorem 3.

5 Proof of Theorem 5

We first show that V is a viscosity supersolution.

Theorem 10. *Given Assumption 1, the value function $V = \{V(t, x, s, \ell)\}_{\ell \in \mathbb{M}}$ is a viscosity supersolution of the HJB equation (8).*

Proof. Let $\ell \in \mathbb{M}$, $(\bar{t}, \bar{x}, \bar{s}) \in [0, T) \times (0, \infty) \times (0, \infty)$. Let the test function $\varphi(t, x, s) \in C^{1,1,2}([0, T) \times (0, \infty) \times (0, \infty))$ such that $V(t, x, s, \ell) - \varphi(t, x, s)$ attains its minimum at $(\bar{t}, \bar{x}, \bar{s})$ and, without loss of generality, $V(\bar{t}, \bar{x}, \bar{s}, \ell) - \varphi(\bar{t}, \bar{x}, \bar{s}) = 0$. Choose a constant control $\bar{u}(t) \equiv \upsilon \in U$ for $t \in [0, \tau_0]$. Let the state variables X and S start from time $\bar{t}$ with initial values $\bar{x}$ and $\bar{s}$.

Define $\hat{\tau}_1$ as the first jump time of the regime $\alpha_{\bar{t},\ell}(\cdot)$. Without loss of generality, assume that η is small enough such that $B_\eta(\bar{x}, \bar{s}) \subset (0, \infty) \times (0, \infty)$. Define $\hat{\tau}_2$ by

$$\hat{\tau}_2 := \inf\left\{r \geq \bar{t} : \left(X_{\bar{t},\bar{x}}^{\bar{u}}(r), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r)\right) \notin B_\eta(\bar{x}, \bar{s})\right\}.$$

For $h < T - \bar{t}$, define the stopping time $\tau := (\bar{t} + h) \wedge \hat{\tau}_1 \wedge \hat{\tau}_2$. Note that $\tau < \tau_0$. By dynamic programming principle,

$$V(\bar{t}, \bar{x}, \bar{s}, \ell) \geq E\left[\int_{\bar{t}}^{\tau} e^{-\beta(r-\bar{t})}\phi(\bar{u}(r))S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r)dr + e^{-\beta(\tau-\bar{t})}V\left(\tau, X_{\bar{t},\bar{x}}^{\bar{u}}(\tau), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(\tau), \alpha_{\bar{t},\ell}(\tau)\right)\right]. \quad (18)$$

Define

$$\psi(t, x, s, i) = \begin{cases} \varphi(t, x, s) & \text{if } i = \ell, \\ V(t, x, s, i) & \text{if } i \neq \ell. \end{cases} \quad (19)$$

10

Applying Dynkin's formula at point $(\bar{t}, \bar{x}, \bar{s}, \ell)$, also noting $\psi(t,x,s,\ell) = \varphi(t,x,s)$, we have

$$
E\left[e^{-\beta(\tau-\bar{t})} \psi\left(\tau, X^{\bar{u}}_{\bar{t},\bar{x}}(\tau), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(\tau), \alpha_{\bar{t},\ell}(\tau) \right) \right]
$$
$$
= \varphi\left(\bar{t}, \bar{x}, \bar{s} \right) + E\left[\int_{\bar{t}}^{\tau} \left\{ (-\beta)e^{-\beta(r-\bar{t})} \varphi\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) \right) \right. \right.
$$
$$
\left. \left. + e^{-\beta(r-\bar{t})} \left(\frac{\partial}{\partial t} + \mathcal{L}^{v} \right) \varphi\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{u}_{\bar{t},\bar{s},\ell}(r) \right) + \mathcal{Q}\psi\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r), \ell \right) \right\} dr \right], \qquad (20)
$$

which implies, from the choice of $(\bar{t}, \bar{x}, \bar{s}, \ell)$ and the definition of ψ, that

$$
E\left[e^{-\beta(\tau-\bar{t})} \psi\left(\tau, X^{\bar{u}}_{\bar{t},\bar{x}}(\tau), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(\tau), \alpha_{\bar{t},\ell}(\tau) \right) \right]
$$
$$
\geq V\left(\bar{t}, \bar{x}, \bar{s}, \ell \right) + E\left[\int_{\bar{t}}^{\tau} \left\{ (-\beta)e^{-\beta(r-\bar{t})} \varphi\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) \right) \right. \right.
$$
$$
\left. \left. + e^{-\beta(r-\bar{t})} \left(\frac{\partial}{\partial t} + \mathcal{L}^{v} \right) \varphi\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{u}_{\bar{t},\bar{s},\ell}(r) \right) + \mathcal{Q}V\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r), \ell \right) \right\} dr \right]. \qquad (21)
$$

Substitute (21) into (18) and divide both sides by $-h$ we get

$$
0 \leq E\left[\frac{1}{h} \int_{\bar{t}}^{\tau} \left\{ e^{-\beta(r-\bar{t})} \left(\beta\varphi\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) \right) - \left(\frac{\partial}{\partial t} + \mathcal{L}^{v} \right) \varphi\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) \right) \right. \right. \right.
$$
$$
\left. \left. \left. - \phi\left(\bar{u}(r) \right) S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) \right) - \mathcal{Q}V\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r), \ell \right) \right\} dr \right]
$$
$$
\leq E\left[\frac{1}{h} \int_{\bar{t}}^{\bar{t}+h} \left\{ e^{-\beta(r-\bar{t})} \left(\beta\varphi\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) \right) - \left(\frac{\partial}{\partial t} + \mathcal{L}^{v} \right) \varphi\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) \right) \right. \right. \right.
$$
$$
\left. \left. \left. - \phi\left(\bar{u}(r) \right) S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) \right) - \mathcal{Q}V\left(r, X^{\bar{u}}_{\bar{t},\bar{x}}(r), S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r), \ell \right) \right\} dr \right| \hat{\tau}_1 \wedge \hat{\tau}_2 > \bar{t} + h \right] P\left[\hat{\tau}_1 \wedge \hat{\tau}_2 > \bar{t} + h \right]
$$
$$
+ K \frac{E\left[(\hat{\tau}_1 \wedge \hat{\tau}_2 - \bar{t}) | \hat{\tau}_1 \wedge \hat{\tau}_2 \leq \bar{t} + h \right]}{h} P\left[\hat{\tau}_1 \wedge \hat{\tau}_2 \leq \bar{t} + h \right] \qquad (22)
$$

for some constant K, due to continuity of the function on the left hand side of (9) and the boundedness of state variable on the time interval $[0, \hat{\tau}_1 \wedge \hat{\tau}_2]$.

By definition of $\hat{\tau}_1$, we have

$$
P\left[\hat{\tau}_1 \leq \bar{t} + h \right] = 1 - P\left[\alpha_{\bar{t},\ell}(r) = \ell, \ r \in (\bar{t}, \bar{t}+h] \right] = -q_{\ell\ell}h.
$$

So as $h \to 0$, $P\left[\hat{\tau}_1 \leq \bar{t} + h \right]$ goes to zero. By Chebyshev's inequality, we have

$$
P\left[\hat{\tau}_2 \leq \bar{t} + h \right] = P\left[\sup_{r \in [\bar{t},\bar{t}+h]} \left\{ \left| X^{\bar{u}}_{\bar{t},\bar{x}}(r) - \bar{x} \right|^2 + \left| S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) - \bar{s} \right|^2 \geq \eta^2 \right\} \right]
$$
$$
\leq \frac{E\left[\sup_{r \in [\bar{t},\bar{t}+h]} \left| X^{\bar{u}}_{\bar{t},\bar{x}}(r) - \bar{x} \right|^2 \right] + E\left[\sup_{r \in [\bar{t},\bar{t}+h]} \left| S^{\bar{u}}_{\bar{t},\bar{s},\ell}(r) - \bar{s} \right|^2 \right]}{\eta^2}. \qquad (23)
$$

Since each term on the numerator of (23) converges to zero as $h \to 0$ and $\lim_{h \to 0} P\left[\hat{\tau}_2 \leq \bar{t} + h \right] = 0$, we have

$$
\lim_{h \to 0} P\left[\hat{\tau}_1 \wedge \hat{\tau}_2 \leq \bar{t} + h \right] \leq \lim_{h \to 0} \left(P\left[\hat{\tau}_1 \leq \bar{t} + h \right] + P\left[\hat{\tau}_2 \leq \bar{t} + h \right] \right) = 0. \qquad (24)
$$

Let $h \to 0$ in (22). By the mean value theorem and the dominated convergence theorem, we have

$$
\beta\varphi\left(\bar{t}, \bar{x}, \bar{s} \right) - \left(\frac{\partial}{\partial t} + \mathcal{L}^{v} \right) \varphi\left(\bar{t}, \bar{x}, \bar{s} \right) - \phi(v)\bar{s} - \mathcal{Q}V\left(\bar{t}, \bar{x}, \bar{s}, \ell \right) \geq 0.
$$

11

Since $\bar{u}(r) \equiv \upsilon \in U$ is chosen arbitrarily, we take the supremum over U and get

$$\beta\varphi\left(\bar{t},\bar{x},\bar{s}\right) - \frac{\partial}{\partial t}\varphi\left(\bar{t},\bar{x},\bar{s}\right) - \sup_{\upsilon\in U}\left\{\mathcal{L}^{\upsilon}\varphi\left(\bar{t},\bar{x},\bar{s}\right) + \phi(\upsilon)\bar{s}\right\} - \mathcal{Q}V\left(\bar{t},\bar{x},\bar{s},\ell\right) \geq 0.$$

Therefore, V is a viscosity supersolution of the HJB equation (8). □

For $\ell \in \mathbb{M}$, define the Hamiltonian function $\mathcal{H}$ by

$$\mathcal{H}(t,x,s,p,q,M,\ell) := \sup_{\upsilon\in U}\left\{-\upsilon p + \mu(t,s,\upsilon,\ell)q + \frac{1}{2}\sigma^2(t,s,\upsilon,\ell)M + \phi(\upsilon)s\right\}. \tag{25}$$

Lemma 11. *For all $\ell \in \mathbb{M}$, the Hamiltonian $\mathcal{H}(t,x,s,p,q,M,\ell)$ is continuous in $(t,x,s,p,q,M) \in [0,T) \times (0,\infty) \times (0,\infty) \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.*

Proof. Let the point $(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M}) \in [0,T) \times (0,\infty) \times (0,\infty) \times \mathbb{R}^3$ and $B_\eta(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M})$ the ball with the center $(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M})$ and the radius η, a small constant. By the definition of the Hamiltonian function, for an arbitrary given $\delta > 0$, there exists a $\bar{\upsilon} \in U$ such that

$$\mathcal{H}\left(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M},\ell\right) - \delta \leq -\bar{\upsilon}\bar{p} + \mu\left(\bar{t},\bar{s},\bar{\upsilon},\ell\right)\bar{q} + \frac{1}{2}\sigma^2\left(\bar{t},\bar{s},\bar{\upsilon},\ell\right)\bar{M} + \phi\left(\bar{\upsilon}\right)\bar{s}. \tag{26}$$

For any point $(t',x',s',p',q',M') \in B_\eta(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M})$ we also have

$$\mathcal{H}\left(t',x',s',p',q',M',\ell\right) \geq -\bar{\upsilon}p' + \mu\left(t',s',\bar{\upsilon},\ell\right)q' + \frac{1}{2}\sigma^2\left(t',s',\bar{\upsilon},\ell\right)M' + \phi\left(\bar{\upsilon}\right)s'. \tag{27}$$

Subtracting (27) from (26), we have

$$\mathcal{H}(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M},\ell) - \mathcal{H}(t',x',s',p',q',M',\ell) - \delta$$
$$+ \frac{1}{2}\sigma^2(\bar{t},\bar{s},\bar{\upsilon},\ell)\left|\bar{M}-M'\right| + \frac{1}{2}\left|\bar{M}+\eta\right|\left|\sigma^2(\bar{t},\bar{s},\bar{\upsilon},\ell) - \sigma^2(t',s',\bar{\upsilon},\ell)\right| + |\phi(\bar{\upsilon})|\left|\bar{s}-s'\right|. \tag{28}$$

Taking the limit inferior and then letting δ tend to zero in (28) we get

$$\mathcal{H}(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M},\ell) \leq \liminf_{\substack{(t',x',s',p',q',M') \\ \to(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M})}} \mathcal{H}(t',x',s',p',q',M',\ell). \tag{29}$$

Similarly, we can show, using the uniform continuity of $\mu(\cdot,\cdot,\cdot,\ell)$ and $\sigma(\cdot,\cdot,\cdot,\ell)$ and the boundedness of the control set U, that

$$\limsup_{\substack{(t',x',s',p',q',M') \\ \to(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M})}} \mathcal{H}(t',x',s',p',q',M',\ell) \leq \mathcal{H}(\bar{t},\bar{x},\bar{s},\bar{p},\bar{q},\bar{M},\ell). \tag{30}$$

(29) and (30) imply that the Hamiltonian $\mathcal{H}(t,x,s,p,q,M,\ell)$ is continuous in (t,x,s,p,q,M). □

For $\varphi \in C^{1,1,2}$ Theorem 3 and Lemma 11 imply that the mapping

$$(t,x,s) \mapsto \beta\varphi(t,x,s) - \frac{\partial\varphi}{\partial t}(t,x,s) - \sup_{\upsilon\in U}\left\{\mathcal{L}^{\upsilon}\varphi(t,x,s) + \phi(\upsilon)s\right\} - \mathcal{Q}V(t,x,s,\ell) \tag{31}$$

is continuous.

Theorem 12. *For each $\ell \in \mathbb{M}$, the value function $V = \{V(t,x,s,\ell)\}_{\ell\in\mathbb{M}}$ is a viscosity subsolution of the HJB equation (8).*

Proof. Assume, for contradiction, that V is not a viscosity subsolution. Then there exists $\ell \in \mathbb{M}$, $(\bar{t}, \bar{x}, \bar{s}) \in [0, T) \times (0, \infty) \times (0, \infty)$ and a test function $\varphi(t, x, s) \in C^{1,1,2}([0, T) \times (0, \infty) \times (0, \infty))$ such that

$$\beta\varphi(\bar{t}, \bar{x}, \bar{s}) - \frac{\partial\varphi}{\partial t}(\bar{t}, \bar{x}, \bar{s}) - \sup_{v \in U}\{\mathcal{L}^v\varphi(\bar{t}, \bar{x}, \bar{s}) + \phi(v)\bar{s}\} - \mathcal{Q}V(\bar{t}, \bar{x}, \bar{s}, \ell) > 0, \tag{32}$$

where $V(t, x, s, \ell) - \varphi(t, x, s)$ attains its maximum at $(\bar{t}, \bar{x}, \bar{s})$. Without loss of generality, assume that $V(\bar{t}, \bar{x}, \bar{s}, \ell) - \varphi(\bar{t}, \bar{x}, \bar{s}) = 0$.

By the continuity of the mapping in (31), for $\delta > 0$, there exists $\eta > 0$ such that

$$\beta\varphi(t, x, s) - \frac{\partial\varphi}{\partial t}(t, x, s) - \sup_{v \in U}\{\mathcal{L}^v\varphi(t, x, s) + \phi(v)s\} - \mathcal{Q}V(t, x, s, \ell) \geq \delta \tag{33}$$

for all $(t, x, s) \in B_\eta(\bar{t}, \bar{x}, \bar{s})$. Let η be small enough such that $B_\eta(\bar{t}, \bar{x}, \bar{s}) \subset [0, T) \times (0, \infty) \times (0, \infty)$.

Let $h > 0$ be small enough such that $(\bar{t}, \bar{t} + h) \subset [0, T)$. By dynamic programming principle, there exists a control process $\bar{u} \in \mathcal{U}$ such that

$$V(\bar{t}, \bar{x}, \bar{s}, \ell) - \frac{\delta}{2}h \leq E\left[\int_{\bar{t}}^{\tau} e^{-\beta(r-\bar{t})}\phi(\bar{u}(r))S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r)dr + e^{-\beta(\tau-\bar{t})}V\left(\tau, X_{\bar{t},\bar{x}}^{\bar{u}}(\tau), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(\tau), \alpha_{\bar{t},\ell}(\tau)\right)\right], \tag{34}$$

where $\tau \geq t$ is any stopping time. Let $\hat{\tau}_1$ be the first jump time of $\alpha_{\bar{t},\ell}(\cdot)$ and define the exit time

$$\hat{\tau}_3 := \inf\left\{r \geq \bar{t} : \left(r, X_{\bar{t},\bar{x}}^{\bar{u}}(r), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r)\right) \notin B_\eta(\bar{t}, \bar{x}, \bar{s})\right\}.$$

Let $\tau := (\bar{t} + h) \wedge \hat{\tau}_1 \wedge \hat{\tau}_3$ and define a function $\psi(t, x, s, i)$ as in (19). We have

$$\psi\left(\tau, X_{\bar{t},\bar{x}}^{\bar{u}}(\tau), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(\tau), \alpha_{\bar{t},\ell}(\tau)\right) \geq V\left(\tau, X_{\bar{t},\bar{x}}^{\bar{u}}(\tau), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(\tau), \alpha_{\bar{t},\ell}(\tau)\right)$$

and

$$\mathcal{Q}\psi\left(r, X_{\bar{t},\bar{x}}^{\bar{u}}(r), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r), \ell\right) \leq \mathcal{Q}V\left(r, X_{\bar{t},\bar{x}}^{\bar{u}}(r), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r), \ell\right). \tag{35}$$

So equation (34) turns into

$$\varphi(\bar{t}, \bar{x}, \bar{s}) - \frac{\delta}{2}h \leq E\left[\int_{\bar{t}}^{\tau} e^{-\beta(r-\bar{t})}\phi(\bar{u}(r))S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r)dr + e^{-\beta(\tau-\bar{t})}\psi\left(\tau, X_{\bar{t},\bar{x}}^{\bar{u}}(\tau), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(\tau), \alpha_{\bar{t},\ell}(\tau)\right)\right]. \tag{36}$$

Combining (20), (35) and (36), we divide both sides of the equation by h,

$$0 \geq -\frac{\delta}{2} + E\left[\frac{1}{h}\int_{\bar{t}}^{\tau}\left\{e^{-\beta(r-\bar{t})}\left[\beta\varphi\left(r, X_{\bar{t},\bar{x}}^{\bar{u}}(r), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r)\right) - \left(\frac{\partial}{\partial t} + \mathcal{L}^{\bar{u}(r)}\right)\right.\right.$$
$$\left.\left. \varphi\left(r, X_{\bar{t},\bar{x}}^{\bar{u}}(r), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r)\right) - \phi(\bar{u}(r))S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r)\right] - \mathcal{Q}V\left(r, X_{\bar{t},\bar{x}}^{\bar{u}}(r), S_{\bar{t},\bar{s},\ell}^{\bar{u}}(r), \ell\right)\right\}dr\right]. \tag{37}$$

Substituting (33) into (37), we have

$$0 \geq -\frac{\delta}{2} + \frac{\delta}{h}E[\tau - \bar{t}]. \tag{38}$$

By (24), we have

$$1 \geq \frac{1}{h}E[\tau - \bar{t}] \geq \frac{1}{h}E[h | \hat{\tau}_1 \wedge \hat{\tau}_3 > \bar{t} + h]P[\hat{\tau}_1 \wedge \hat{\tau}_2 > \bar{t} + h] = P[\hat{\tau}_1 \wedge \hat{\tau}_3 > \bar{t} + h] \to 1$$

as $h \to 0$, which implies that

$$\lim_{h \to 0}\frac{1}{h}E[\tau - \bar{t}] = 1.$$

Letting $h \to 0$ in (38), we get $\delta/2 \leq 0$, a contradiction. The inequality in (32) therefore holds, which completes the proof.

$\square$

Since the value function V is both a viscosity subsolution and a viscosity supersolution, we conclude that it is a viscosity solution of the HJB equation (8). We have proved Theorem 5.

6 Proof of Theorem 6

In this section vectors (t, x, s) and (r, y, v) and their specific values such as $(\bar{t}, \bar{x}, \bar{s})$ appear many times. To simplify the expressions we denote by $\mathbf{x} = (t, x, s)$ and $\mathbf{y} = (r, y, v)$. Their specific values are defined similarly, for example, $\bar{\mathbf{x}} = (\bar{t}, \bar{x}, \bar{s})$.

To prove the uniqueness, we need an alternative definition of viscosity solution in terms of superjets and subjets. The second-order superjet of an upper-semicontinuous function U at a point $\bar{\mathbf{x}} \in \Sigma := [0, T) \times (0, \infty) \times (0, \infty)$, denoted by $\mathcal{P}^{2,+}U(\bar{\mathbf{x}})$, is defined as a set of elements $(\bar{b}, \bar{p}, \bar{q}, \bar{M}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ such that

$$U(\mathbf{x}) \leq U(\bar{\mathbf{x}}) + (\bar{b}, \bar{p}, \bar{q}) \cdot (\mathbf{x} - \bar{\mathbf{x}}) + \frac{1}{2}\bar{M}(s - \bar{s})^2 + e(\mathbf{x} - \bar{\mathbf{x}}), \tag{39}$$

where $e(\mathbf{x} - \bar{\mathbf{x}}) = o(|t - \bar{t}| + |x - \bar{x}| + |s - \bar{s}|^2)$ is a higher order error term. The limiting superjet $\overline{\mathcal{P}}^{2,+}U(\mathbf{x})$ is the set of elements $(b, p, q, M) \in \mathbb{R}^4$ for which there exists a sequence $(\mathbf{x}_\epsilon)$ in Σ and $(b_\epsilon, p_\epsilon, q_\epsilon, M_\epsilon) \in \mathcal{P}^{2,+}U(\mathbf{x}_\epsilon)$ such that $(\mathbf{x}_\epsilon, U(\mathbf{x}_\epsilon), b_\epsilon, p_\epsilon, q_\epsilon, M_\epsilon) \to (\mathbf{x}, U(\mathbf{x}), b, p, q, M)$.

The second-order subjet of a lower-semicontinuous function V at a point $\bar{\mathbf{x}} \in \Sigma$, denoted by $\mathcal{P}^{2,-}V(\bar{\mathbf{x}})$, is defined as in (39) with a greater than or equal ($\geq$) inequality. The set $\overline{\mathcal{P}}^{2,-}V(\mathbf{x})$ is defined similarly.

Note that since x is a state variable superjets and subjets should normally also have second order terms with respect to x. However, since the HJB equation (8) only involves the first order derivative of the value function with respect to x, the second order expansion in x is not needed.

Assume that U is upper-semicontinuous and $\varphi \in C^{1,1,2}(\Sigma)$. Then $\bar{\mathbf{x}} \in \Sigma$ is a maximum point of $U - \varphi$ if and only if $(D_{\mathbf{x}}\varphi(\bar{\mathbf{x}}), D_s^2\varphi(\bar{\mathbf{x}})) \in \mathcal{P}^{2,+}U(\bar{\mathbf{x}})$, where $D_{\mathbf{x}}\varphi(\bar{\mathbf{x}}) = (D_t\varphi(\bar{\mathbf{x}}), D_x\varphi(\bar{\mathbf{x}}), D_s\varphi(\bar{\mathbf{x}}))$. Similar conclusion holds for the minimum point and the subjet.

Lemma 13. *([6, Theorem 8.3]) An m-tuple $V = \{V(\cdot, \cdot, \cdot, \ell)\}_{\ell \in \mathbb{M}}$ of continuous functions on Σ is a viscosity subsolution (resp. supersolution) of the HJB equation (8) if and only if for $\mathbf{x} \in \Sigma$ such that $(b, p, q, M) \in \overline{\mathcal{P}}^{2,+}V(\mathbf{x}, \ell)$ (resp. $\in \overline{\mathcal{P}}^{2,-}V(\mathbf{x}, \ell)$) for any fixed $\ell \in \mathbb{M}$, we have*

$$\beta V(\mathbf{x}, \ell) - b - \mathcal{H}(\mathbf{x}, p, q, M, \ell) - \mathcal{Q}V(\mathbf{x}, \ell) \leq 0 \quad (resp. \geq 0),$$

where $\mathcal{H}(\mathbf{x}, p, q, M, \ell)$ is the Hamiltonian define in (25). The m-tuple V is a viscosity solution if it is both a viscosity subsolution and a viscosity supersolution.

The uniform polynomial growth condition for U and V implies that there exists a constant $p > 1$ such that, for each $\ell \in \mathbb{M}$

$$\sup_{[0,T] \times (0,\infty) \times (0,\infty)} \frac{|U(\mathbf{x}, \ell)| + |V(\mathbf{x}, \ell)|}{1 + |x|^p + |s|^p} < \infty.$$

Define functions $\theta(x, s) := (1 + |x|^{2p} + |s|^{2p})$ and $\kappa(t, x, s) := e^{-\gamma t}\theta(x, s)$ for $\gamma > 0$. Due to the linear growth condition (3) and the boundedness of set U, there exists a positive constant c such that, for all $\ell \in \mathbb{M}$,

$$\beta\kappa - \frac{\partial\kappa}{\partial t} - \sup_{v \in U}\{\mathcal{L}^v\kappa\}$$
$$= \beta\kappa - \frac{\partial\kappa}{\partial t} - \sup_{v \in U}\left\{-vD_x\kappa + \mu(t, s, v, \ell)D_s\kappa + \frac{1}{2}\sigma^2(t, s, v, \ell)D_s^2\kappa + \mathcal{Q}\kappa\right\}$$
$$= e^{-\gamma t}\left[(\beta + \gamma)\theta - \sup_{v \in U}\left\{-vD_x\theta + \mu(t, s, v, \ell)D_s\theta + \frac{1}{2}\sigma^2(t, s, v, \ell)D_s^2\theta + \mathcal{Q}\theta\right\}\right]$$
$$\geq e^{-\gamma t}(\beta + \gamma - c)\theta,$$

which is nonnegative as long as we choose the constant γ large enough such that $(\beta + \gamma - c) > 0$. Therefore, for any $\epsilon > 0$, $\widetilde{V}^\epsilon(\mathbf{x}, \ell) := V(\mathbf{x}, \ell) + \epsilon\kappa(\mathbf{x})$ is a supersolution to the HJB equation (8). To

check this, let $\varphi(\mathbf{x}, \ell)$ be the test function for $\widetilde{V}^\epsilon(\mathbf{x}, \ell)$. So $\varphi(\mathbf{x}, \ell) - \epsilon\kappa(\mathbf{x})$ is the test function for the supersolution $V(\mathbf{x}, \ell)$. We have

$$\beta\varphi - \frac{\partial\varphi}{\partial t} - \sup_{v\in U}\{\mathcal{L}^v\varphi + \phi(v)s\}$$

$$\geq \beta(\varphi - \epsilon\kappa) - \frac{\partial}{\partial t}(\varphi - \epsilon\kappa) - \sup_{v\in U}\{\mathcal{L}^v(\varphi - \epsilon\kappa) + \phi(v)s\} + \epsilon\left(\beta\kappa - \frac{\partial\kappa}{\partial t} - \sup_{v\in U}\{\mathcal{L}^v\kappa\}\right)$$

$$\geq 0.$$

By the polynomial growth condition of U, V and the definition of κ, we have

$$\lim_{x,s\to\infty}\sup_{t\in[0,T]}(U - \widetilde{V}^\epsilon)(\mathbf{x}, \ell) = -\infty$$

for all $\epsilon > 0$. We can assume that the maximum of $(U - V)(\mathbf{x}, \ell)$ over $\ell \in \mathbb{M}$ and $\mathbf{x} \in [0, T] \times (0, \infty) \times (0, \infty)$ is attained (up to a penalization) at $\ell \in \mathbb{M}$ and $\mathbf{x} \in \Sigma_1 := [0, T] \times O_1 \times O_2$ for some compact set $O_1 \subset (0, \infty)$ and $O_2 \subset (0, \infty)$. Let $\mathcal{M}$ denote this maximum.

Suppose, for contradiction, that there exists $\ell \in \mathbb{M}$ and $\mathbf{x} \in \Sigma$ such that $U(\mathbf{x}, \ell) > V(\mathbf{x}, \ell)$. We have

$$\mathcal{M} := \max_{i\in\mathbb{M}}\sup_{[0,T]\times(0,\infty)^2}(U - V)(\mathbf{x}, i) = \max_{i\in\mathbb{M},\mathbf{x}\in\Sigma_1}(U - V)(\mathbf{x}, i) > 0. \tag{40}$$

For any $\epsilon > 0$, define a function Ψ^ϵ by

$$\Psi^\epsilon(\mathbf{x}, \mathbf{y}, \ell) := U(\mathbf{x}, \ell) - V(\mathbf{y}, \ell) - \psi^\epsilon(\mathbf{x}, \mathbf{y}),$$

where ψ^ϵ is defined by

$$\psi^\epsilon(\mathbf{x}, \mathbf{y}) := \frac{1}{2\epsilon}|\mathbf{x} - \mathbf{y}|^2. \tag{41}$$

For each $\ell \in \mathbb{M}$, $\Psi^\epsilon(\cdot, \cdot, \ell)$ is continuous. Hence its maximum, denoted by $\mathcal{M}_\ell^\epsilon$, over the compact set $\Sigma_1 \times \Sigma_1$ can be attained at $(\mathbf{x}_\ell^\epsilon, \mathbf{y}_\ell^\epsilon)$. Assume that the maximum $\mathcal{M}^\epsilon := \max_{\ell\in\mathbb{M}}\mathcal{M}_\ell^\epsilon$ is attained at $\ell^\epsilon \in \mathbb{M}$ and $(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon)$. We have

$$\mathcal{M} \leq \mathcal{M}^\epsilon = \Psi^\epsilon(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) \leq U(\mathbf{x}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) - V(\mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon). \tag{42}$$

As $\epsilon \to 0$, the bounded sequence $(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon)$ converges, up to a subsequence, to a limit $(\bar{\mathbf{x}}, \bar{\mathbf{y}}) \in \Sigma_1 \times \Sigma_1$. By assumption, $\mathbb{M}$ is finite. For each $\ell \in \mathbb{M}$, the sequence $(\mathbf{x}_\ell^\epsilon, \mathbf{y}_\ell^\epsilon)$ converges, up to a subsequence, to its limit, respectively. Therefore, for ϵ small enough, $\ell^\epsilon = \bar{\ell}$ for $\bar{\ell} \in \mathbb{M}$.

Since $\{U(\cdot, \ell)\}_{\ell\in\mathbb{M}}$ and $\{V(\cdot, \ell)\}_{\ell\in\mathbb{M}}$ are continuous and $\mathbb{M}$ is a finite set, $U(\mathbf{x}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) - V(\mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon)$ is bounded for all $\epsilon > 0$. From (42), $\psi^\epsilon(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon)$ is also bounded, which implies that

$$\lim_{\epsilon\to 0}(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon) = (\bar{\mathbf{x}}, \bar{\mathbf{x}}), \quad \lim_{\epsilon\to 0}\mathcal{M}^\epsilon = \mathcal{M} = (U - V)(\bar{\mathbf{x}}, \bar{\ell}), \quad \lim_{\epsilon\to 0}\psi^\epsilon(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon) = 0. \tag{43}$$

By applying Ishii's Lemma (see [11, Lemma 4.4.6, Remark 4.4.9]) to function Ψ^ϵ at its maximum point $(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon)$ with $\ell = \ell^\epsilon$, we can find $M^\epsilon, N^\epsilon \in \mathbb{R}$ such that

$$\left(\frac{1}{\epsilon}(\mathbf{x}_{\ell^\epsilon}^\epsilon - \mathbf{y}_{\ell^\epsilon}^\epsilon), M^\epsilon\right) \in \overline{\mathcal{P}}^{2,+}U(\mathbf{x}_{\ell^\epsilon}^\epsilon, \ell^\epsilon), \quad \left(\frac{1}{\epsilon}(\mathbf{x}_{\ell^\epsilon}^\epsilon - \mathbf{y}_{\ell^\epsilon}^\epsilon), N^\epsilon\right) \in \overline{\mathcal{P}}^{2,-}V(\mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon)$$

and, for any $c, d \in \mathbb{R}$,

$$c^2 M^\epsilon - d^2 N^\epsilon \leq \frac{3}{\epsilon}(c - d)^2. \tag{44}$$

Denote by

$$(\eta_1^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon) := \frac{1}{\epsilon}(\mathbf{x}_{\ell^\epsilon}^\epsilon - \mathbf{y}_{\ell^\epsilon}^\epsilon) = \left(\frac{1}{\epsilon}(t_{\ell^\epsilon}^\epsilon - r_{\ell^\epsilon}^\epsilon), \frac{1}{\epsilon}(x_{\ell^\epsilon}^\epsilon - y_{\ell^\epsilon}^\epsilon), \frac{1}{\epsilon}(s_{\ell^\epsilon}^\epsilon - v_{\ell^\epsilon}^\epsilon)\right).$$

15

Since U is a viscosity subsolution and V a supersolution, by the definition of viscosity solutions in terms of superjets and subjets, we have

$$\beta U(\mathbf{x}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) - \eta_1^\epsilon - \mathcal{Q}U(\mathbf{x}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) - \mathcal{H}\left(\mathbf{x}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, M^\epsilon, \ell^\epsilon\right) \leq 0 \tag{45}$$

$$\beta V(\mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) - \eta_1^\epsilon - \mathcal{Q}V(\mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) - \mathcal{H}\left(\mathbf{y}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, N^\epsilon, \ell^\epsilon\right) \geq 0, \tag{46}$$

where the Hamiltonian $\mathcal{H}$ is defined in (25). By the definition of operator $\mathcal{Q}$ we have

$$\begin{aligned}
&\mathcal{Q}\left(U(\mathbf{x}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) - V(\mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon)\right) \\
&= \sum_{j \neq \ell^\epsilon} q_{\ell^\epsilon j}\left[(U(\mathbf{x}_{\ell^\epsilon}^\epsilon, j) - V(\mathbf{y}_{\ell^\epsilon}^\epsilon, j)) - (U(\mathbf{x}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) - V(\mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon))\right] \\
&= \sum_{j \neq \ell^\epsilon} q_{\ell^\epsilon j}\left[\Psi^\epsilon(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon, j) - \Psi^\epsilon(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon)\right] \\
&\leq 0.
\end{aligned} \tag{47}$$

The last line is from the fact that $\Psi^\epsilon(\mathbf{x}_{\ell^\epsilon}^\epsilon, \mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon)$ is the maximum of $\Psi^\epsilon(\mathbf{x}, \mathbf{y}, \ell)$ over $\ell \in \mathbb{M}$ and $(\mathbf{x}, \mathbf{y}) \in \Sigma_1 \times \Sigma_1$.

Subtracting (46) from (45) and rearranging, also noting (47), we have

$$\beta\left(U(\mathbf{x}_{\ell^\epsilon}^\epsilon, \ell^\epsilon) - V(\mathbf{y}_{\ell^\epsilon}^\epsilon, \ell^\epsilon)\right) \leq \mathcal{H}\left(\mathbf{x}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, M^\epsilon, \ell^\epsilon\right) - \mathcal{H}\left(\mathbf{y}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, N^\epsilon, \ell^\epsilon\right). \tag{48}$$

By the definition of the Hamiltonian function, for any $\delta > 0$, there exists a $v^\delta \in U$ such that

$$\mathcal{H}\left(\mathbf{x}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, M^\epsilon, \ell^\epsilon\right) - \delta \leq -v^\delta \eta_2^\epsilon + \mu\left(t_{\ell^\epsilon}^\epsilon, s_{\ell^\epsilon}^\epsilon, v^\delta, \ell^\epsilon\right)\eta_3^\epsilon + \frac{1}{2}\sigma^2\left(t_{\ell^\epsilon}^\epsilon, s_{\ell^\epsilon}^\epsilon, v^\delta, \ell^\epsilon\right)M^\epsilon + \phi(v^\delta)s_{\ell^\epsilon}^\epsilon. \tag{49}$$

We also have

$$\mathcal{H}\left(\mathbf{y}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, N^\epsilon, \ell^\epsilon\right) \geq -v^\delta \eta_2^\epsilon + \mu\left(r_{\ell^\epsilon}^\epsilon, v_{\ell^\epsilon}^\epsilon, v^\delta, \ell^\epsilon\right)\eta_3^\epsilon + \frac{1}{2}\sigma^2\left(s_{\ell^\epsilon}^\epsilon, v_{\ell^\epsilon}^\epsilon, v^\delta, \ell^\epsilon\right)N^\epsilon + \phi(v^\delta)v_{\ell^\epsilon}^\epsilon. \tag{50}$$

Subtracting (50) from (49), we get

$$\begin{aligned}
&\mathcal{H}\left(\mathbf{x}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, M^\epsilon, \ell^\epsilon\right) - \mathcal{H}\left(\mathbf{y}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, N^\epsilon, \ell^\epsilon\right) - \delta \\
&\leq \eta_3^\epsilon\left[\mu\left(t_{\ell^\epsilon}^\epsilon, s_{\ell^\epsilon}^\epsilon, v^\delta, \ell^\epsilon\right) - \mu\left(r_{\ell^\epsilon}^\epsilon, v_{\ell^\epsilon}^\epsilon, v^\delta, \ell^\epsilon\right)\right] \\
&\quad + \frac{3}{2\epsilon}(\sigma\left(t_{\ell^\epsilon}^\epsilon, s_{\ell^\epsilon}^\epsilon, v^\delta, \ell^\epsilon\right) - \sigma\left(r_{\ell^\epsilon}^\epsilon, v_{\ell^\epsilon}^\epsilon, v^\delta, \ell^\epsilon\right))^2 + \phi(v^\delta)\left(s_{\ell^\epsilon}^\epsilon - v_{\ell^\epsilon}^\epsilon\right).
\end{aligned} \tag{51}$$

Here we have used (44).

By Assumption 1 on μ and σ, (43) and the boundedness of $\phi(v^\delta)$, the right side of (51) tends to 0 as $\epsilon \to 0$. Since $\delta > 0$ is chosen arbitrarily, we have

$$\limsup_{\epsilon \to 0}\left\{\mathcal{H}\left(\mathbf{x}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, M^\epsilon, \ell^\epsilon\right) - \mathcal{H}\left(\mathbf{y}_{\ell^\epsilon}^\epsilon, \eta_2^\epsilon, \eta_3^\epsilon, N^\epsilon, \ell^\epsilon\right)\right\} \leq 0. \tag{52}$$

Combining (43), (48) and (52), we have

$$\beta\left(U\left(\bar{\mathbf{x}}, \bar{\ell}\right) - V\left(\bar{\mathbf{x}}, \bar{\ell}\right)\right) \leq 0,$$

which contradicts (40). Therefore $U \leq V$ on $[0, T) \times (0, \infty) \times (0, \infty) \times \mathbb{M}$. We have proved Theorem 6.

References

[1] Bank, P. and Baum, D., Hedging and portfolio optimization in financial markets with a large trader, *Mathematical Finance* 14, 1-18, 2004.

[2] Black, F., Towards a fully automated exchange: Part 1, *Financial Analyst Journal* 27, 29-34, 1971.

[3] Çetin, U., Jarrow, R.A. and Protter, P., Liquidity risk and arbitrage pricing theory, *Finance and Stochastics* 8, 311-341, 2004.

[4] W.H. Fleming, and H.M. Soner *Controlled Markov Processes and Viscosity Solutions*, Springer, 2006

[5] Cvitanic, J. and Karatzas, I., Hedging and portfolio optimization under transaction costs: a martingale approach, *Mathematical Finance* 6, 370-398, 1996.

[6] Crandall, M.G., Ishii, H. and Lions, P.L., User's guide to viscosity solutions of second order partial differential equations, *Bulletin American Mathematical Society* 27, 1-67, 1992.

[7] Crandall, M.G. and Lions, P.L., Two apprximations of solutions of Hamilton-Jacobi equations, *Mathematics of Computation* 43, 1-19, 1984.

[8] Gassiat, P., Gozzi, F. and Pham, H., Investment/consumption problem in illiquid markets with regimes switching, it SIAM J. Control Optimization, to appear, 2012.

[9] Jouini, E., Price functionals with bid-ask spreads: an axiomatic approach, *J. Mathematical Economics* 34, 547-558, 2000.

[10] Mao, X. and Yuan, C., *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, 2006.

[11] Pham, H., *Continuous-time Stochastic Control and Optimization with Financial Applications*, Springer, 2010

[12] Pemy, M. and Zhang, Q., Optimal stock liquidation in a regime switching model with finite time horizon, *J. Mathematical Analysis & Applications* 321, 537-552, 2006.

[13] Pemy, M., Zhang, Q. and Yin, G., Liquidation of a large block of stock, *J. Banking & Finance* 31, 1295-1305, 2007.

[14] Pemy, M., Zhang, Q. and Yin, G., Liquidation of a large block of stock with regime switching, *Mathematical Finance* 18, 629-648, 2008.

[15] Schied, A. and Schöneborn, T., Optimal portfolio liquidation for CARA investors, working paper, 2007.