Optimal control of trading algorithms: a general impulse control approach

Bruno Bouchard[*], Ngoc-Minh Dang [†] and Charles-Albert Lehalle [‡]

October 1, 2010

Abstract

We propose a general framework for intra-day trading based on the control of trading algorithms. Given a set of generic parameterized algorithms (which have to be specified by the controller ex-ante), our aim is to optimize the dates $(\tau_i)_i$ at which they are launched, the length $(\delta_i)_i$ of the trading period and the value of the parameters $(\mathcal{E}_i)_i$ kept during the time interval $[\tau_i, \tau_i + \delta_i)$. This provides to the financial agent a decision tool for selecting which algorithm (and for which set of parameters and time length) should be used in the different phases of the trading period. From the mathematical point of view, this gives rise to a non-classical impulse control problem where not only the regime $\mathcal{E}_i$ but also the period $[\tau_i, \tau_i + \delta_i)$ have to be determined by the controller at the impulse time τ_i. We adapt the *weak dynamic programming principle* of Bouchard and Touzi (2009) to our context to provide a characterization of the associated value function as a discontinuous viscosity solution of a system of PDEs with appropriate boundary conditions, for which we prove a comparison principle. We also propose a numerical scheme for the resolution of the above system and show that it is convergent. We finally provide a simple example of application to a problem of optimal stock trading with a non-linear market impact function. This shows how parameters adapt to the market.

Key words: optimal impulse control, discontinuous viscosity solutions, intra-day trading.

Mathematical subject classifications: 93E20, 49L25, 91B28.

1 Introduction

Trading algorithms are nowadays widely spread among financial agents. They are typically used for high frequency intra-day trading purposes, e.g. for "statistical arbitrage" or for the execution of large orders by brokers. In both cases, the use of robots is justified by the fact that orders have to be executed very quickly, in order to make profit of "good prices", and, typically for brokers, by the large size of the portfolios to be handled by a limited number of traders.

A lot of efforts have been devoted these last years to build efficient trading algorithms, taking all possible market features into account, and in particular the so-called market impact effect. Some of the most popular ones have been proposed by academics, see e.g. [1], [2] [3], [5] and

[*]CEREMADE, Université Paris Dauphine and CREST-ENSAE, bouchard@ceremade.dauphine.fr
[†]CEREMADE, Université Paris Dauphine and CA Cheuvreux, dang@ceremade.dauphine.fr
[‡]CA Cheuvreux, clehalle@cheuvreux.com

[11], a lot of them are developed in research divisions of brokers or investment banks, and are somehow kept secret as fundamental building blocks of their everyday business.
One could expect that these algorithms correspond to global optimization problems and are run on the whole trading period without interruption. In practice, this is often not the case, in particular among brokers.
In fact the way brokers execute large orders is typically as follows: they split the global number of assets to be bought or sold into small pieces, called slices, and then use robots to execute these different slices one by one. Each time a new slice is launched, the trader tunes the parameters of the algorithm (and possibly the size of the slice) depending on the evolution of the market's conditions. He can even decide to use a different algorithm from one slice to the other (which can actually also be viewed as changing the parameters of a single robot, at least from the mathematical point of view).
The real life situation is thus the following: given a bunch of trading algorithms, the trader decides how to slice the order (i.e. divide the global order in smaller parts), at what time each slice is launched, and with which algorithm and values of the parameters. Moreover, he also decides how long he should let the algorithm run. In practice, there exists a minimal time period during which each algorithm is executed. One reason for this is simply that the trader can not practically monitor all the algorithms that are running for different purposes at the same time. Another one is that it makes no sense (in practice) to launch an algorithm for less than, say, one second.
The aim of this paper is to provide a decision tool for traders given the above described practical situation. Namely, we propose a rigorous framework for the optimal control of trading algorithms: how does one decide optimally of the times at which algorithms are launched, what parameters to use, for how long? Obviously, one could argue that it would be better to discuss a global optimization problem, i.e. discuss the optimal control problem associated to the global trading agenda. This is not the aim of this paper. Most practitioners have their own algorithms and do not want to have the same as the others. Our approach allows to adapt to the algorithms each trader wants to use (optimal or not), and help them to use these algorithms in an optimal way, and in particular to tune the parameters. Exactly like the computation of greeks serves as a reference for the hedging policy of an option, which is then more or less followed by the trader depending on his own feeling of the market's conditions (and because he can in practice only discretely rebalanced his portfolio), we aim at providing on-line values of the parameters that should be optimal in some sense and therefore should serve as references for the trader in the way he launches the different slices.

We are therefore not interested here in the trading algorithms themselves. The aim of this paper is not to come up with a new algorithm but to provide a rigorous decision tool for the use of already existing trading algorithms.

From the mathematical point of view, it gives rise to a non-classical impulse control problem. As for standard impulse control problems, see e.g. the reference book [6], one chooses the times at which "impulses" are given (times at which an algorithm is launched) and the size of the impulse (the value of the parameters). The novelty comes from the fact that one also chooses a time period during which no new impulse can be given (the period during which the algorithm runs). To the best of our knowledge, such problems have never been discussed in the mathematical literature on optimal impulse control.
In this paper, we provide a rigorous characterization of the value function as a discontinuous viscosity solution of a partial differential equation (PDE), together with suitable boundary conditions. To this end, we adapt the approach of [7] who proposes a weak version of the

dynamic programming principle. The main advantage of this weak formulation is that it does not require any *a-priori* continuity of the value function. It is the first time this approach is used in the context of impulse control problems, and this requires some non-trivial modifications of the arguments of [7].

Our PDE characterization reads as follows. When the current regime is the *passive* one, i.e. no trading algorithm is running, the controller can launch one at any moment τ_i with a given set of parameters $\mathcal{E}_i$ and for a period of length δ_i. This gives rise to a standard quasi-variational inequality in the region corresponding to the passive regime. However, once the algorithm is launched, no change in the value of the parameters can be made before the end of the period $[\tau_i, \tau_i + \delta_i)$. This implies that the value function satisfies a linear parabolic equation on the *active* region. We also provide a comparison principle for the above equations and construct a finite difference numerical scheme, which we prove to be convergent.

The rest of the paper is organized as follows. The model is described in Section 2. In Section 3, we provide the PDE characterization of the value function and the associated comparison principle. The proofs of these results are given in Section 4. The numerical scheme is studied in Section 5. In the last section, we discuss a very simple numerical example of application to a particular model of optimal stock acquisition. It shows how the proposed framework naturally allows a real-time adaptive control of the trading algorithm, by switching optimally after the end of each slice given the current state of the market.

Notation: All over this paper, we shall use the following notations. Given $x \in \mathbb{R}^k$, for k given by the context, we denote by $|x|$ its Euclidean norm and by $B_r(x)$ the open ball of center x and radius $r > 0$. The scalar product is denoted by $\langle \cdot, \cdot \rangle$. Given a set $A \subset \mathbb{R}^k$, we denote by ∂A its boundary. Given $d \in \mathbb{N}$, we denote by $\mathbb{M}^d$ the set of d-dimensional square matrices. For $M \in \mathbb{M}^d$, M^* is the associated transposed matrix. For a function $(t, x, y) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^k \mapsto \varphi(t, x, y)$, we denote by $D\varphi$ and $D^2\varphi$ its gradient and Hessian matrix with respect to x, whenever they are well defined. The other partial derivatives will be written by using standard notations.

2 Problem formulation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space supporting a d-dimensional Brownian motion W, $d \geq 1$. Let $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$ denote the right-continuous complete filtration generated by W, and let $T > 0$ be a finite time horizon.

2.1 Control policies

We first describe how the trading algorithms are controlled, precise dynamics will be imposed in Section 2.2 below. Note that different algorithms can be viewed as a single parameterized one. In what follow, we therefore consider that we have only one algorithm.

A control policy of the trading algorithm is described by a non-decreasing sequence of stopping times $(\tau_i)_{i \geq 1}$ and a sequence of $[\underline{\delta}, \infty) \times E$-valued random variables $(\delta_i, \mathcal{E}_i)_{i \geq 1}$. The stopping times τ_i describe the times at which an order is given to the algorithm, $\mathcal{E}_i$ is the value of the parameters with which the algorithm is run and δ_i the length of the period (latency period) during which it is run with the value $\mathcal{E}_i$. The set E is a compact subset of $\mathbb{R}^d$, which represents the possible values of the parameters, the quantity

$$0 < \underline{\delta} < T$$

denotes the minimum length of the time period during which the algorithm can be run. To be consistent we impose that

$$\tau_i + \delta_i \le \tau_{i+1} \text{ and } (\delta_i > 0 \Rightarrow \tau_i + \delta_i \le T) \;,\; i \ge 1 \;. \tag{2.1}$$

The first condition expresses the fact that a new order can not be given before the end of the time period associated to the previous order. The second one means that an order should be given only if it ends before the final time horizon T.

Remark 2.1. 1. The minimal duration constraint, $\delta_i \ge \underline{\delta}$ with $\underline{\delta} > 0$, has been justified from a practical point of view in the introduction. From the mathematical point of view, the problem described in Sections 2.2 and 2.3 below would not make sense without this condition, if no additional cost related to launching the algorithm with new parameters is introduced. Indeed, for $\underline{\delta} = 0$, and without additional costs, the controller could, at the limit, control the parameters continuously, and this would actually certainly be optimal. The controller will then act, at the limit, as a trader acting continuously on the market.
2. Models with $\underline{\delta} = 0$ and with an additional cost (paid each time the algorithm is launched with new parameters) could be discussed by following the lines of this paper. Such a cost is actually already embedded in the general dynamics of Section 2.2, up to an additional assumption on the function β, see Example 2.10 below. This would however require to justify this cost, and to evaluate it in practice. Moreover, certain bounds like the one stated in Remark 2.2 below would not be true anymore and other conditions would be required to retrieve the estimates of Remark 2.5 below. For sake of simplicity, we therefore stick to the case $\underline{\delta} > 0$ which corresponds more, from our point of view, to practical situations.
3. In the absence of a cost penalizing frequent changes in the parameters, it may be optimal to choose $\delta_i = \underline{\delta}$ most of the time (depending whether $\underline{\delta}$ is small or not). However, other values of δ_i may also be optimal. Our algorithm provides a way to select the maximal optimal value, the one for which the trader changes the parameters the least often (which is desirable in practice).

As usual the value of the parameters and the size of the latency period can not be chosen in some anticipative way, i.e. we impose that

$$(\delta_i, \mathcal{E}_i) \text{ is } \mathcal{F}_{\tau_i}\text{-measurable} \;,\; i \ge 1 \;. \tag{2.2}$$

At time $t \in [\tau_i, \tau_i + \delta_i)$, the value of the parameter of the trading algorithm is denoted by ν_t. For $t \in A((\tau_i, \delta_i)_{i\ge1})$, defined as

$$A((\tau_i, \delta_i)_{i\ge1}) := \mathbb{R}_+ \setminus \left(\bigcup_{i\ge1} [\tau_i, \tau_i + \delta_i) \right) \;,$$

we set $\nu_t = \varpi$, where $\varpi \in \mathbb{R}^d \setminus E$ can be viewed as a cemetery point, recall that E is compact. It follows that the value of the parameters of the trading algorithm ν can be written as

$$\nu_t = \varpi \mathbf{1}_{t \in A((\tau_i,\delta_i)_{i\ge1})} + \sum_{i\ge1} \mathcal{E}_i \mathbf{1}_{t \in [\tau_i, \tau_i+\delta_i)} \;,\; t \in [0, T] \;, \tag{2.3}$$

where $\nu_t = \varpi$ means that the algorithm is not running at time t.

In the following, we denote by $\mathcal{S}$ the set of adapted processes ν that can be written in the form (2.3) for some sequence of stopping times $(\tau_i)_{i\geq 1}$ and of $[\underline{\delta},\infty)\times E$-valued random variables $(\delta_i,\mathcal{E}_i)_{i\geq 1}$ satisfying (2.1) and (2.2).
For ease of notations, we shall now write

$$(\tau_i^\nu,\delta_i^\nu,\mathcal{E}_i^\nu)_{i\geq 1} \text{ the sequence associated to } \nu\in\mathcal{S},$$

and define, for all stopping times ϑ_1 and ϑ_2 satisfying $\vartheta_1 \leq \vartheta_2$ $\mathbb{P}$-a.s., the set of indices corresponding to orders whose execution ends between ϑ_1 and ϑ_2:

$$\mathbb{I}^\nu_{\vartheta_1,\vartheta_2} := \{i\geq 1 : \vartheta_1 < \tau_i^\nu+\delta_i^\nu \leq \vartheta_2\}\ .$$

Remark 2.2. Note that the constraint $\delta_i^\nu \geq \underline{\delta}$ for all $i\geq 1$ and $\nu\in\mathcal{S}$ implies that

$$\mathrm{card}(\mathbb{I}^\nu_{0,T}) \leq \mathrm{card}\left(\{\tau_i^\nu\leq T,\ i\geq 1\}\right) \leq T/\underline{\delta}\ .$$

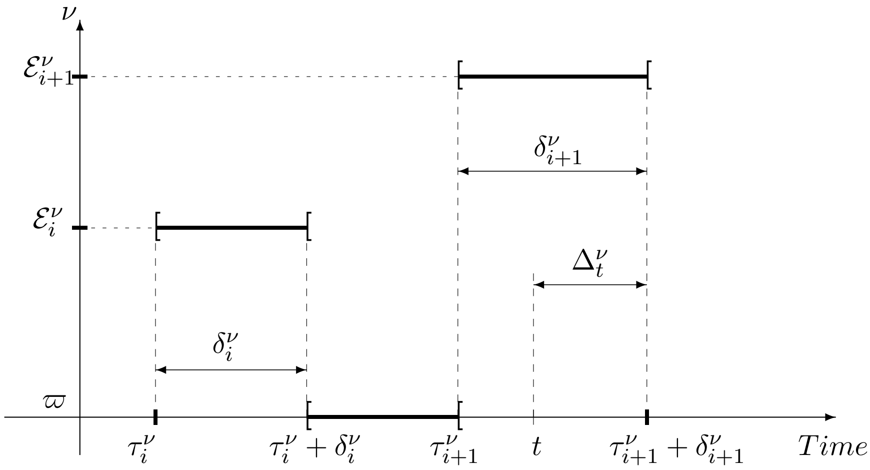
For ease of notations, we also set

$$\bar{E} := E\cup\{\varpi\}\subset\mathbb{R}^d\ ,$$

and introduce the process $\Delta^\nu_\cdot$

$$\Delta_t^\nu := \sum_{i\geq 1}[\tau_i^\nu+\delta_i^\nu-t]^+\mathbf{1}_{t\geq\tau_i^\nu}\ ,\ t\in[0,T]\ .$$

The quantity Δ_t^ν denotes the remaining latency period during which no new order can be passed to the algorithm. When $\Delta_t^\nu>0$, the algorithm is running with a value of the parameters ν_t. When $\Delta_t^\nu=0$, the algorithm is not running anymore, $\nu_t=\varpi$, and a new order can be passed.
The following picture sums up the dynamics of the control.



2.2 Output of the trading algorithm

Given some initial data $(t,x)\in[0,T]\times\mathbb{R}^d$, the output of the trading algorithm associated to some control policy $\nu\in\mathcal{S}$ is defined as the strong solution $X^\nu_{t,x}$ on $[0,T]$ of the stochastic

differential equation
(2.4)

$$X^{\nu}_{t,x}(s) = x + \mathbf{1}_{s\geq t}\left(\int_t^s b(X^{\nu}_{t,x}(r),\nu_r)dr + \int_t^s a(X^{\nu}_{t,x}(r),\nu_r)dW_r + \sum_{i\geq 1}\beta(X^{\nu}_{t,x}(\tau_i^{\nu}-),\mathcal{E}_i^{\nu},\delta_i^{\nu})\mathbf{1}_{t<\tau_i^{\nu}\leq s}\right) ,$$

where $b, \beta : \mathbb{R}^d \times \bar{E} \times [\underline{\delta}, T] \to \mathbb{R}^d$ and $a : \mathbb{M}^d \times \bar{E} \to \mathbb{M}^d$ are continuous functions such that there exists $K > 0$ for which, for all $x, x' \in \mathbb{R}^d$, $e, e' \in \bar{E}$, $\delta, \delta' \in [\underline{\delta}, T]$,

$$(2.5)\begin{cases} |\psi(x,e,\delta) - \psi(x',e,\delta)|) & \leq & K|x-x'| \\ |\psi(x,e,\delta)| & \leq & K(1+|x|) \\ |\psi(x,e,\delta) - \psi(x,e',\delta')| & \leq & K(1+|x|)(|e-e'|+|\delta-\delta'|) \end{cases} \quad \text{for } \psi = b, a, \beta .$$

We do not differentiate here between the components that correspond to real outputs of the algorithm (cumulated gains, cumulated volumes executed by the algorithm, etc...) and others that simply describe the evolution of financial data or market factors (prices of the traded assets, global traded volumes on the markets, volatilities, etc...).
The jumps on the dynamics are introduced to model the changes in the initial conditions on the variables of interest for the trading algorithm when it is launched (e.g. volume to be executed between τ_i^{ν} and $\tau_i^{\nu} + \delta_i^{\nu}$), see Example 2.7 below.
Moreover, there is no loss of generality in assuming that X, W and ν have the same dimension d. One can always reduce more general situations to this case by playing with the coefficients b, a, β and with the choice of E. Time dependent coefficients can similarly be considered by putting the first line of a and β equal to 0, and the first component of b equal to 1, so that the first component of X actually coincides with the time parameter.
We refer to Section 2.4 for the description of simple examples of application which illustrate the flexibility of the above model.

We conclude this section with two technical remarks that will be used later on.

Remark 2.3. Observe that X^{ν} does not jump when the regime is switched to the passive regime ϖ.

Remark 2.4. Note that (2.5), the fact that E is bounded and Remark 2.2 imply that, for all $t \in [0,T]$, $x \in \mathbb{R}^d$ and $\nu \in \mathcal{S}$,

$$\mathbb{E}\left[\sup_{s\in[t,T]} |X^{\nu}_{t,x}(s)|^p\right] \leq C^p_K(1+|x|^p) \tag{2.6}$$

where C^p_K depends only on K and $p \geq 1$.

2.3 Gain function

The aim of the controller is to maximize the expected value of the gain functional

$$\nu \in \mathcal{S} \mapsto \Pi_{t,x}(\nu) := g(X^{\nu}_{t,x}(T)) + \sum_{i\in\mathbb{I}^{\nu}_{t,T}} f(X^{\nu}_{t,x}(\tau_i^{\nu}+\delta_i^{\nu}-),\mathcal{E}_i^{\nu}) ,$$

with the usual convention $\sum_{\emptyset} = 0$, among the set

$$\mathcal{S}_{t,\delta,e} := \begin{cases} \{\nu \in \mathcal{S} \ : \ \nu_s = e \text{ for } s \in [t,t+\delta) \text{ and } \Delta^{\nu}_{t+\delta} = 0\} & \text{if } e \neq \varpi \text{ and } \delta > 0 \\ \{\nu \in \mathcal{S} \ : \ \nu_t = \varpi\} & \text{otherwise} \end{cases}$$

where $(\delta, e) \in \mathbb{R}_+ \times \bar{E}$ denotes the initial state of the remaining latency time and value of the parameters.

Here, g and f are assumed to be continuous on $\mathbb{R}^d \times \bar{E}$ and to satisfy, for some $\gamma > 0$,

$$\sup_{(x,e)\in\mathbb{R}^d\times E} \frac{|f(x,e)| \;+\; |g(x)|}{1+|x|^\gamma} < \infty \;\;,\;\; f(\cdot,\varpi) = 0\,. \tag{2.7}$$

In view of (2.6), this ensures that the quantity

$$J(t,x;\nu) \;:=\; \mathbb{E}\left[\Pi_{t,x}(\nu)\right]$$

is well defined for all $\nu \in \mathcal{S}$ and satisfies

$$|J(t,x;\nu)| \le C_K^\gamma\,(1+|x|^\gamma) \tag{2.8}$$

where C_K^γ depends only on K and γ.

For technical reason related to the dynamic programming principle, see [7] and the proof of Lemma 4.1 below, we shall restrict to admissible trading strategies $\nu \in \mathcal{S}_{t,\delta,e}$ such that ν is independent on $\mathcal{F}_t$, see Remark 5.2 in [7]. We denote by $\mathcal{S}^a_{t,\delta,e}$ the associated set of controls and therefore define the value function as:

$$V(t,x,\delta,e) := \sup_{\nu\in\mathcal{S}^a_{t,\delta,e}} \mathbb{E}\left[\Pi_{t,x}(\nu)\right]\;.$$

We refer to Section 2.4 for examples of application, and to Section 6 for a numerical illustration.

Remark 2.5. It follows from (2.8) that there exists $C_K^\gamma > 0$ which depends only on K and γ such that

$$|V(t,x,\delta,e)| \le C_K^\gamma\,(1+|x|^\gamma) \quad \text{for all } (t,x,\delta,e) \in [0,T]\times\mathbb{R}^d\times\mathbb{R}_+\times\bar{E} \text{ s.t. } \mathcal{S}^a_{t,\delta,e} \neq \emptyset\,.$$

Note that for $\delta = T-t$ and $e \in E$, (2.1) implies that

$$V(t,x,T-t,e) = \mathcal{V}(t,x,e) := \mathbb{E}\left[g(\mathcal{X}^e_{t,x}(T)) + f(\mathcal{X}^e_{t,x}(T),e)\right]\;, \tag{2.9}$$

where $\mathcal{X}^e_{t,x}$ is the solution of

$$\mathcal{X}^e_{t,x}(s) = x + \int_t^s b(\mathcal{X}^e_{t,x}(r),e)dr + \int_t^s a(\mathcal{X}^e_{t,x}(r),e)dW_r\;,\; s\in[t,T]\,. \tag{2.10}$$

Remark 2.6. Under (2.5), the continuity assumption on f, g and (2.7), it follows from standard arguments that the auxiliary value function $\mathcal{V}$ is continuous, and that, for each $e \in E$, it is a viscosity solution of

$$-\,\mathcal{L}^e\varphi(t,x) = 0 \;\text{ on } [0,T)\times\mathbb{R}^d\,,\; \varphi(T,x) = g(x) + f(x,e) \;\text{ on } \mathbb{R}^d\,, \tag{2.11}$$

where, for $e \in \bar{E}$, and a smooth function φ,

$$\mathcal{L}^e\varphi(t,x) := \frac{\partial}{\partial t}\varphi(t,x) \;+\; \langle b(x,e)\;,\; D\varphi(t,x)\rangle \;+\; \frac{1}{2}\mathrm{Tr}\left[aa^*(x,e)D^2\varphi(t,x)\right]\;. \tag{2.12}$$

2.4 Examples

Before to go on with the presentation of our general problem, let us describe simple examples of application which illustrate the flexibility of the above model. For sake of simplicity, we restrict here to the case of a single parameterized algorithm.

Example 2.7. As a first example, we consider the case where the aim of the controller is to sell a number Q_0 of one stock S between 0 and $T > 0$. We denote by V_t the global volume instantaneously traded on the market at time t. The dynamics of (S, V) is given by the strong solution of the SDE

$$
\begin{aligned}
S_t &= S_0 + \int_0^t \mu_S(r, S_r, V_r)\, dr + \int_0^t \sigma_S(r, S_r, V_r)\, dW_r \ , \\
V_t &= V_0 + \int_0^t \mu_V(r, S_r, V_r)\, dr + \int_0^t \sigma_V(r, S_r, V_r)\, dW_r \ ,
\end{aligned}
$$

where W denotes a two dimensional standard Brownian motion, and $(\mu_S, \sigma_S, \mu_V, \sigma_V)$ are Lipschitz continuous. We implicitly assume here that the above SDE has non-negative solutions whatever the initial conditions are.
A control $\nu \in S$ is identified to a sequence $(\tau_i^\nu, \delta_i^\nu, \mathcal{E}_i^\nu)_{i\geq 1} \in \mathcal{S}$ as in Section 2. Here $\mathcal{E}_i^\nu$ stands for the proportion of the remaining number of shares that have to be sold, $Q^\nu_{\tau_i^\nu}$, which will be traded on $[\tau_i^\nu, \tau_i^\nu + \delta_i^\nu)$. We assume that this quantity is sold uniformly on the corresponding time interval. Namely, we sell $\mathcal{E}_i^\nu Q^\nu_{\tau_i}/\delta_i^\nu dt$ on $[t, t+dt]$ for $t \in [\tau_i^\nu, \tau_i^\nu + \delta_i^\nu)$.
This means that the dynamics of the remaining number of stocks to sell is given by

$$
Q_t^\nu = Q_0 - \sum_{i\geq 1} \int_{\tau_i^\nu}^{\tau_i^\nu + \delta_i^\nu} \mathbf{1}_{s\leq t} \mathcal{E}_i^\nu Q^\nu_{\tau_i^\nu}/\delta_i^\nu ds \ .
$$

Clearly, E has to be contained in $[0, 1]$. Note that in order to ensure that the dynamics of the system is Markovian, we need to introduce two additional processes which coincide with $Q^\nu_{\tau_i^\nu}$ and δ_i^ν on $[\tau_i^\nu, \tau_i^\nu + \delta_i^\nu]$. This is done by considering the process $(\bar{Q}^\nu, \bar{\delta}^\nu)$ defined by the dynamics

$$
\bar{Q}_t^\nu := \sum_{i\geq 1} (Q^\nu_{\tau_i^\nu -} - \bar{Q}^\nu_{\tau_i^\nu -}) \mathbf{1}_{\tau_i^\nu \leq t} \ \text{ and } \ \bar{\delta}_t^\nu := \sum_{i\geq 1} (\delta_i^\nu - \bar{\delta}^\nu_{\tau_i^\nu -}) \mathbf{1}_{\tau_i^\nu \leq t}
$$

with the convention $\bar{Q}^\nu_{0-} = \bar{\delta}^\nu_{0-} = 0$. This explains why jumps have been introduced in the general dynamics (2.4). It follows that

$$
Q_t^\nu = Q_0 - \int_0^t q(\nu_s, \bar{\delta}_s^\nu) \bar{Q}_s^\nu ds \ .
$$

where $q(e, \delta) := e/\delta$ if $e \neq \varpi$ and $\delta > 0$, $q(e, \delta) = 0$ otherwise.
Due to the impact of the strategy on the market, the price obtained for the volume executed under the regime $\mathcal{E}_i^\nu$ is

$$
\tilde{S}_t = S_t - \eta(\mathcal{E}_i^\nu Q^\nu_{\tau_i^\nu}/\delta_i^\nu, S_t, V_t) \ , \ t \in [\tau_i^\nu, \tau_i^\nu + \delta_i^\nu)
$$

where η is some market impact function and is assumed to be Lipschitz continuous. It follows that the cumulated wealth's dynamic is

$$
Y_t^\nu = 0 + \int_0^t \tilde{S}_r q(\nu_r, \bar{\delta}_r^\nu) \bar{Q}_r^\nu dr \ .
$$

The remaining part Q_T^ν is instantaneously sold on the market at a price: $S_T - c(Q_T^\nu, S_T, V_T)$, for some Lipschitz continuous function c.
The total gain after the final transaction is thus given by:

$$Y_T^\nu + (S_T - c(Q_T^\nu, S_T, V_T))\,(Q_T^\nu)^+ \ .$$

The aim of the controller is to maximize the expectation of the quantity

$$g(Y_T^\nu + (S_T - c(Q_T^\nu, S_T, V_T))\,(Q_T^\nu)^+)$$

for some concave function g with polynomial growth.
In this model, the process X coincides with $(S, V, Y, Q, \bar{Q}, \bar{\delta})$. Only $\bar{Q}$ and $\bar{\delta}$ have jumps.

Example 2.8. As a second example, we consider the case where the aim of the controller is to buy a number Q_0 of one stock S between 0 and $T > 0$. The dynamics of (S, V) is given as in the previous example.
Here $\mathcal{E}_i^\nu$ stands for the intensity at which the stocks are bought, i.e. the algorithm buys a number $\mathcal{E}_i^\nu dt = \nu_t \mathbf{1}_{\nu_t \neq \varpi} dt$ of stocks on $[t, t+dt]$, $t \in [\tau_i^\nu, \tau_i^\nu + \delta_i^\nu)$. The dynamics of the remaining number of stocks to be bought before T is thus given by:

$$Q_t^\nu = Q_0 - \int_0^t q(\nu_s) ds$$

where q is now defined as $q(e) = e\mathbf{1}_{e \neq \varpi}$. It follows that the cumulated wealth's dynamic is

$$Y_t^\nu = 0 + \int_0^t \tilde{S}_r q(\nu_r) dr = 0 + \int_0^t (S_r + \eta(\nu_r, S_r, V_r)) q(\nu_r) dr \ ,$$

where η is a given market impact function.
If the number Q_0 of shares is not liquidated at time T, the remaining part Q_T^ν is instantaneously bought on the market at the price $S_T + c(Q_T^\nu, S_T, V_T)$, for some Lipschitz continuous function c.
The total cost after the final transaction is thus given by:

$$Y_T^\nu + (S_T + c(Q_T^\nu, S_T, V_T))\,(Q_T^\nu)^+ \ .$$

The aim of the controller is to minimize the expectation of the quantity

$$\ell(Y_T^\nu + (S_T + c(Q_T^\nu, S_T, V_T))\,(Q_T^\nu)^+)$$

for some convex function ℓ with polynomial growth.
A numerical application within the above framework is presented in Section 6.

Example 2.9. We now consider a similar situation as in the previous example except that the controller has an incentive to buy the shares more or less quickly. This can be modeled by adding a process Z^ν corresponding to the number of stocks sold during a trading period. Namely,

$$Z_t^\nu := \sum_{i \geq 1} \mathbf{1}_{t \in [\tau_i^\nu, \tau_i^\nu + \delta_i^\nu)} \int_{\tau_i^\nu}^t \nu_s ds = \int_0^t \nu_s \mathbf{1}_{\nu_s \neq \varpi} ds - \sum_{i \geq 1} Z_{\tau_i^\nu -}^\nu \mathbf{1}_{\tau_i^\nu \leq t} \ .$$

The aim of the controller is then to minimize

$$\mathbb{E}\left[\ell(Y_T^\nu + (S_T + c(Q_T^\nu, S_T, V_T))\,(Q_T^\nu)^+) + f(T, (Q_T^\nu)^+) + \sum_{i \geq 1} \mathbf{1}_{\tau_i^\nu + \delta_i^\nu \leq T} f(\tau_i^\nu + \delta_i^\nu, Z_{\tau_i + \delta_i}^\nu)\right]$$

where the dependence of f in time means that the controller prefers to buy quickly, i.e. f is increasing in time, or take more time, i.e. f is decreasing in time.

Example 2.10. We finally explain how to incorporate a cost paid each time the algorithm is launched with new parameters. It suffices to consider an additional process

$$C_t^\nu := \sum_{i\geq 1} \lambda \mathbf{1}_{\tau_i^\nu \leq t},$$

for some $\lambda > 0$, and to consider a reward function which is non-increasing in C_T^ν. Obviously, one can let λ depend on other variables of interest.

3 Viscosity characterization of the value function

The aim of this section is to provide a PDE characterization of the value function V. Before to state our main result, we need to introduce some additional notations and definitions.

In view of (2.1), (2.9) and the constraint that the latency period should be greater than $\underline{\delta}$, the natural domain of definition of the value function V is

$$D := \left\{(t,x,\delta,e) \in [0,T) \times \mathbb{R}^d \times (((0,\infty)\times E) \cup \{(0,\varpi)\}) \;:\; \underline{\delta} \leq t+\delta < T \text{ or } e = \varpi\right\},$$

which can be decomposed in two main regions. We call the *active region*, the region where $\delta > 0$ and $e \neq \varpi$:

$$D_{E,>0} := \left\{(t,x,\delta,e) \in [0,T) \times \mathbb{R}^d \times (0,\infty) \times E \;:\; \underline{\delta} \leq t+\delta < T\right\}. \tag{3.1}$$

It corresponds to the set of initial conditions where the algorithm is running and the controller has to wait the end of the latency period before passing a new order. We call the *passive region*, the region where $e = \varpi$, and therefore $\delta = 0$:

$$D_\varpi := [0,T) \times \mathbb{R}^d \times \{(0,\varpi)\}. \tag{3.2}$$

It corresponds to the set of initial conditions where the algorithm is not running and can be launched immediately with a new set of parameters. These two regions are complemented by the natural boundaries of the active region when $\delta \to 0$ and $t+\delta \to T$:

$$D_{E,0} := [\underline{\delta}, T) \times \mathbb{R}^d \times \{0\} \times E \tag{3.3}$$

$$D_{E,T} := \left\{(t,x,\delta,e) \in [0,T) \times \mathbb{R}^d \times (0,\infty) \times E \;:\; \underline{\delta} \leq t+\delta = T\right\}, \tag{3.4}$$

and by the time boundary:

$$D_T := \{T\} \times \mathbb{R}^d \times \mathbb{R}_+ \times \bar{E}. \tag{3.5}$$

The closure of the natural domain of definition of the value function V is therefore

$$\bar{D} := \left\{(t,x,\delta,e) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}_+ \times \bar{E} \;:\; \underline{\delta} \leq t+\delta \leq T \text{ or } e = \varpi\right\}.$$

As usual, we shall rely on the dynamic programming principle, see Lemma 4.1 below for a precise statement, to deduce the behavior of the value function on each component of $\bar{D}$:

$$V(t,x,\delta,e) = \sup_{\nu \in S^a_{t,\delta,e}} \mathbb{E}\left[V(\vartheta, X^\nu_{t,x}(\vartheta), \Delta^\nu_\vartheta, \nu_\vartheta) + \sum_{i\in \mathbb{I}^\nu_{t,\vartheta}} f(X^\nu_{t,x}(\tau^\nu_i + \delta^\nu_i), \mathcal{E}^\nu_i)\right], \tag{3.6}$$

for any $[t,T]$-valued stopping time ϑ.

In the passive region. For $(t,x,\delta,e) \in D_{\varpi}$, the controller can immediately launch the trading algorithm with a new set of parameters $(\delta',e') \in [\underline{\delta}, T-t] \times E$. Taking $\vartheta = t$ in (3.6) thus implies that

$$V(t,x,0,\varpi) \geq \mathcal{M}[V](t,x)$$

where

$$\mathcal{M}[V](t,x) := \sup_{(\delta',e') \in [\underline{\delta}, T-t] \times E} V(t, x + \beta(x,e',\delta'), \delta', e') ,$$

with the usual convention $\sup \emptyset = -\infty$. The controller can also decide to wait before passing a new order to the algorithm, i.e. choose $\nu = \varpi$ on some time interval $[t, t+\delta')$ with $\delta' > 0$. In view of (3.6) applied to an arbitrarily small stopping time $\vartheta < t + \delta'$, this implies that

$$-\mathcal{L}^{\varpi} V(t,x,0,\varpi) \geq 0 .$$

The dynamic programming principle (3.6) formally implies that one of the two above choices should be optimal, i.e.

$$\min \{ -\mathcal{L}^{\varpi} V(t,x,0,\varpi) \; ; \; V(t,x,0,\varpi) - \mathcal{M}[V](t,x) \} = 0 .$$

In the active region. For $(t,x,\delta,e) \in D_{E,>0}$, the controller can not change the parameter of the algorithm before the end of the initial latency period $\delta > 0$. Choosing ϑ arbitrarily small in (3.6) thus implies that V should satisfy

$$\left(-\mathcal{L}^{e} + \frac{\partial}{\partial \delta} \right) V(t,x,\delta,e) = 0 ,$$

where $(t,x) \mapsto \mathcal{L}^{e} V(t,x,\delta,e)$ is defined as in (2.12) for (δ, e) taken as parameters. It is naturally complemented with the boundary conditions

$$V(t,x,\delta,e) = V(t,x,0,\varpi) + f(x,e) \; , \;\; \text{if } (t,x,\delta,e) \in D_{E,0} ,$$

and

$$V(t,x,\delta,e) = \mathcal{V}(t,x,e) \; , \;\; \text{if } (t,x,\delta,e) \in D_{E,T} ,$$

recall (2.9).

Terminal boundary condition. As usual, the boundary condition as $t \uparrow T$ should be given by the terminal condition:

$$V(t,x,\delta,e) = g(x) + f(x,e) \; , \;\; \text{if } (t,x,\delta,e) \in D_{T} ,$$

where we recall that $f(\cdot, \varpi) = 0$ by convention.

The above discussion shows that V should solve the equation

$$\mathcal{H}\varphi = 0 \tag{3.7}$$

on $\bar{D}$, where, for a smooth function φ defined on $\bar{D}$,

$$\mathcal{H}\varphi(t,x,\delta,e) := \begin{cases} \left(-\mathcal{L}^{e} + \frac{\partial}{\partial \delta} \right) \varphi(t,x,\delta,e) & \text{on} \;\; D_{E,>0} , \\ \varphi(t,x,\delta,e) - \varphi(t,x,0,\varpi) - f(x,e) & \text{on} \;\; D_{E,0} , \\ \varphi(t,x,\delta,e) - \mathcal{V}(t,x,e) & \text{on} \;\; D_{E,T} , \\ \min \left\{ -\mathcal{L}^{\varpi} \varphi(t,x,\delta,e) \; ; \; \varphi(t,x,\delta,e) - \mathcal{M}[\varphi](t,x) \right\} & \text{on} \;\; D_{\varpi} , \\ \varphi(t,x,\delta,e) - g(x) - f(x,e) & \text{on} \;\; D_{T} . \end{cases}$$

However, since V may not be smooth, it has to be stated in terms of viscosity solutions, see [8], in the following sense.

Definition 3.1. *We say that a lower-semicontinuous (resp. upper-semicontinuous) function U on $\bar{D}$ is a viscosity supersolution (resp. subsolution) of (3.7) on $\bar{D}$ if for any function $\varphi \in C^{1,2,1,0}([0,T] \times \mathbb{R}^d \times \mathbb{R}_+ \times \bar{E})$ and $(t_0, x_0, \delta_0, e_0) \in \bar{D}$, which achieves a global minimum (resp. maximum) of $U - \varphi$ on $\bar{D}$ such that $(U - \varphi)(t_0, x_0, \delta_0, e_0) = 0$, we have*

$$\mathcal{H}\varphi(t_0, x_0, \delta_0, e_0) \geq 0 \quad (\text{ resp. } \mathcal{H}\varphi(t_0, x_0, \delta_0, e_0) \leq 0) \ .$$

If U is continuous, we say that it is a viscosity solution of (3.7) if it is a super- and a subsolution.

In all this paper, we shall say that a function φ is *smooth* if it belongs to $C^{1,2,1,0}([0,T] \times \mathbb{R}^d \times \mathbb{R}_+ \times \bar{E})$.

As usual, showing that V is *a-priori* continuous is a rather difficult task. As a first step, we shall therefore prove the super- and subsolution property only for the upper- and lower-semicontinuous enveloppes V^* and V_* of V defined as

$$\begin{aligned} V^*(t,x,\delta,e) &:= \limsup_{(t',x',\delta',e') \in D \to (t,x,\delta,e)} V(t',x',\delta',e') \\ V_*(t,x,\delta,e) &:= \liminf_{(t',x',\delta',e') \in D \to (t,x,\delta,e)} V(t',x',\delta',e') \ , \quad (t,x,\delta,e) \in \bar{D} \ . \end{aligned}$$

Theorem 3.2. *The function V_* (resp. V^*) is a viscosity supersolution (resp. subsolution) of (3.7) on $\bar{D}$.*

The following comparison result combined with Remark 2.5 insures *a-posteriori* that V is continuous and that it is the unique viscosity solution of (3.7) on $\bar{D}$ with polynomial growth.

Theorem 3.3. *Let u and v be respectively a lower semicontinuous viscosity supersolution of (3.7) on $\bar{D}$ and a upper-semicontinuous viscosity subsolution of (3.7) on $\bar{D}$. Assume that v^+ and u^- have polynomial growth. Then, $u \geq v$ on $\bar{D}$.*

4 Proof of the viscosity characterization

4.1 Dynamic programming

As usual, the derivation of the partial differential equation relies on the so-called dynamic programming principle, a formal version of which is given in (3.6) above. In this section, we provide a rigorous formulation which follows ideas introduced in [7]. Namely, we only provide a weak formulation in terms of test functions. The main advantage of this approach is that it does not require any regularity on the value function V itself, but only some lower-semicontinuity of the objective function $J(\cdot;\nu)$, see below. We refer to [7] for a general discussion.

Lemma 4.1 (Weak Dynamic Programming Principle). *Fix $(t,x,\delta,e) \in D$ and let $\{\vartheta^\nu,\ \nu \in \mathcal{S}^a_{t,\delta,e}\}$ be a family of $[t,T]$-valued stopping times independent of $\mathcal{F}_t$. Then, we have*

(4.1)
$$V(t,x,\delta,e) \leq \sup_{\nu \in \mathcal{S}^a_{t,\delta,e}} \mathbb{E}\left[[V^*, g](\vartheta^\nu, X^\nu_{t,x}(\vartheta^\nu), \Delta^\nu_{\vartheta^\nu}, \nu_{\vartheta^\nu}) + \sum_{i \in \mathbb{I}^\nu_{t,\vartheta^\nu}} f(X^\nu_{t,x}(\tau^\nu_i + \delta^\nu_i -), \mathcal{E}^\nu_i) \right] ,$$

where $[V^*, g](s,\cdot) := V^*(s,\cdot)\mathbf{1}_{s<T} + g\mathbf{1}_{s=T}$*, and*

$$\text{(4.2)}\quad \sup_{\nu\in\mathcal{S}^a_{t,\delta,e}} \mathbb{E}\left[\varphi(\vartheta^\nu, X^\nu_{t,x}(\vartheta^\nu), \Delta^\nu_{\vartheta^\nu}, \nu_{\vartheta^\nu}) + \sum_{i\in\mathbb{I}^\nu_{t,\vartheta^\nu}} f(X^\nu_{t,x}(\tau^\nu_i + \delta^\nu_i -), \mathcal{E}^\nu_i)\right] \le V(t,x,\delta,e)$$

for all upper semi-continuous function φ *such that* $V \ge \varphi$ *on* $\bar{D}$*.*

As in [7], the proof of the above result relies on some lower-semicontinuity property of the function J. Because of the latency time δ, we can however not apply their result directly and need to adapt their arguments by exploiting the lower-semicontinuity of the map

$$(t,x,\delta,e,\nu) \in \bar{D}\times\mathcal{S} \mapsto J(t,x,\mathscr{P}^e_{t,t+\delta}(\nu))$$

where, for $s_1 \le s_2 \in [0,T]$,

$$\mathscr{P}^e_{s_1,s_2}\ :\ \nu\in\mathcal{S} \mapsto \mathscr{P}^e_{s_1,s_2}(\nu) := e\mathbf{1}_{[0,s_1)} + \varpi\mathbf{1}_{[s_1,s_2)} + \nu\mathbf{1}_{[s_2,T]}\ .$$

Lemma 4.2. *Fix* $(t,x,\delta,e)\in D$ *and* $\nu\in\mathcal{S}_{t,\delta,e}$*. Let* $(t_n,x_n,\delta_n,e_n)_{n\ge1}$ *be a sequence in* D *such that* $(t_n,x_n,\delta_n,e_n)\to(t,x,\delta,e)$ *as* $n\to\infty$ *and*

$$\text{(4.3)}\qquad t_n \le t \text{ and } t_n+\delta_n \le t+\delta \quad \text{for all } n\ge 1.$$

Then, $\liminf_{n\to\infty} J(t_n,x_n;\mathscr{P}^{e_n}_{t_n+\delta_n,t+\delta}(\nu)) \ge J(t,x;\nu)$*.*

Proof. We only prove the result in the case where $\delta>0$. The case $\delta=0$ can be handled similarly. In this case we have $t_n \le t < t_n+\delta_n \le t+\delta$ for n large enough since $t_n\to t$ and $\delta_n\to\delta>0$. For ease of notations, we set $X := X^\nu_{t,x}$ and $X^n := X^{\nu^n}_{t_n,x_n}$ where $\nu^n := \mathscr{P}^{e_n}_{t_n+\delta_n,t+\delta}(\nu)$. In all this proof, we let $C>0$ denote a generic constant which does not depend on n but takes value which may change from line to line.
1. We first prove that, for any fixed integer $p\ge1$:

$$\text{(4.4)}\qquad \lim_{n\to\infty}\mathbb{E}\left[\sup_{t\le s\le T}|X^n(s)-X(s)|^{2p}\right] = 0\ .$$

First note that $\nu^n=\nu$ on $[t+\delta,T]$. Since the size of the possible jump of X and X^n at time $t+\delta$, recall (2.4), depends on $\nu_{t+\delta}=\nu^n_{t+\delta}$, standard computations based on Burkholder-Davis-Gundy's inequality, Gronwall's Lemma and the Lipschitz continuity of b, a and β, we thus deduce that

$$\mathbb{E}\left[\sup_{t+\delta\le s\le T}|X^n(s)-X(s)|^{2p}\right] \le C\,\mathbb{E}\left[|X^n(t+\delta-)-X(t+\delta-)|^{2p}\right]\ .$$

Since $(\nu^n,\nu)=(\varpi,e)$ on $[t_n+\delta_n,t+\delta)$ and no jump occurs on this time interval, the Lipschitz continuity of b and a implies that

$$\mathbb{E}\left[\sup_{t_n+\delta_n\le s<t+\delta}|X^n(s)-X(s)|^{2p}\right] \le C\left(|t+\delta-t_n-\delta_n|^p + \mathbb{E}\left[|X^n(t_n+\delta_n)-X(t_n+\delta_n)|^{2p}\right]\right)\ .$$

We similarly have, since $(\nu^n,\nu)=(e_n,e)$ on $[t,t_n+\delta_n)$ and (X^n,X) does not jump on this time interval, that:

$$\mathbb{E}\left[\sup_{t\le s\le t_n+\delta_n}|X^n(s)-X(s)|^{2p}\right] \le C\left(|e_n-e|^{2p} + \mathbb{E}\left[|X^n(t)-x|^{2p}\right]\right)\ .$$

Finally, by the linear growth condition on b and a, the fact that $\nu_n = e_n$ on $[t_n, t)$ and that X^n does not jump at time t,

$$\begin{aligned}\mathbb{E}\left[\sup_{t_n\le s\le t}|X^n(s)-x|^{2p}\right] &\le C\left(|x_n-x|^{2p}+\mathbb{E}\left[\sup_{t_n\le s\le t}|X^n(s)-x_n|^{2p}\right]\right)\\ &\le C\left(|x_n-x|^{2p}+|t_n-t|^p\right) .\end{aligned}$$

2. We now use the above estimate to conclude the proof. We first note that

$$\begin{aligned}\Pi_{t_n,x_n}(\nu^n)-\Pi_{t,x}(\nu) = &\Big(g(X^n(T))-g(X(T))\Big)\\ &+\Bigg(\sum_{i\in\mathbb{I}^{\nu^n}_{t_n,T}} f(X^n(\tau^n_i+\delta^n_i-),\mathcal{E}^n_i)-\sum_{i\in\mathbb{I}^{\nu}_{t,T}} f(X(\tau^\nu_i+\delta^\nu_i-),\mathcal{E}^\nu_i)\Bigg) ,\end{aligned} \tag{4.5}$$

with $(\tau^n,\delta^n,\mathcal{E}^n) := (\tau^{\nu^n},\delta^{\nu^n},\mathcal{E}^{\nu^n})$. In view of (4.4), we can assume that

$$\sup_{t\le s\le T}|X(s)-X^n(s)| \longrightarrow 0 \;\; \mathbb{P}\text{-a.s.} \tag{4.6}$$

after possibly passing to a subsequence. Similar estimates show that, after possibly passing to a subsequence,

$$|X^n(t_n+\delta_n-)-X^n(t+\delta-)| \longrightarrow 0 \;\; \mathbb{P}\text{-a.s. and in } L^p,\ p\ge 1 . \tag{4.7}$$

Since g is continuous, it follows that

$$\lim_{n\to\infty} g(X^n(T)) = g(X(T)) \;\; \mathbb{P}\text{-a.s.} \tag{4.8}$$

Moreover, by definition of ν^n, we have

$$\sum_{i\in\mathbb{I}^{\nu^n}_{t_n,T}} f(X^n(\tau^n_i+\delta^n_i-),\mathcal{E}^n_i) = f(X^n(t_n+\delta_n-),e_n)+\sum_{i\in\mathbb{I}^{\nu}_{t+\delta,T}} f(X^n(\tau^\nu_i+\delta^\nu_i-),\mathcal{E}^\nu_i) .$$

It then follows from the continuity of f, (4.6) and (4.7) that

$$\begin{aligned}\lim_{n\to\infty}\sum_{i\in\mathbb{I}^{\nu^n}_{t_n,T}} f(X^n(\tau^n_i+\delta^n_i-),\mathcal{E}^n_i) &= f(X(t+\delta-),e)+\sum_{i\in\mathbb{I}^{\nu}_{t+\delta,T}} f(X(\tau^\nu_i+\delta^\nu_i-),\mathcal{E}^\nu_i)\\ &= \sum_{i\in\mathbb{I}^{\nu}_{t,T}} f(X(\tau^\nu_i+\delta^\nu_i-),\mathcal{E}^\nu_i) \;\; \mathbb{P}\text{-a.s.}\end{aligned} \tag{4.9}$$

The required result then follows from (4.8), (4.9), (2.7), and Fatou's Lemma combined with (4.4) and (2.6) which insure that the sequence $(\Pi_{t_n,x_n}(\nu^n)^-)_{n\ge1}$ is uniformly integrable. □

We now turn to the proof of the dynamic programming principle.

Proof. [Lemma 4.1] In this proof, we consider $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ as the d-dimensional canonical filtered space equipped with the Wiener measure and denote by ω or $\tilde\omega$ a generic point. The Brownian motion is thus defined as $W(\omega)=(\omega_t)_{t\ge0}$. For $\omega\in\Omega$ and $r\ge0$, we set $\omega^r := \omega_{\cdot\wedge r}$ and $\mathbf{T}_r(\omega) := \omega_{\cdot+r}-\omega_r$. In the following, we omit the dependence of ϑ^ν with respect to ν and simply write ϑ, for ease of notations.

1. The proof of (4.1) is standard and is based on the observation that, for all $\nu \in \mathcal{S}^a_{t,\delta,e}$,

$$
(4.10) \qquad J(t,x;\nu) = \mathbb{E}\left[\mathbb{E}\left[g(X^\nu_{t,x}(T)) + \sum_{i\in\mathbb{I}^\nu_{\vartheta,T}} f(X^\nu_{t,x}(\tau^\nu_i+\delta^\nu_i-),\mathcal{E}^\nu_i) \mid \mathcal{F}_\vartheta\right]\right] + \mathbb{E}\left[\sum_{i\in\mathbb{I}^\nu_{t,\vartheta}} f(X^\nu_{t,x}(\tau^\nu_i+\delta^\nu_i-),\mathcal{E}^\nu_i)\right]
$$

where, by the flow property of X^ν,

$$
(4.11) \qquad \mathbb{E}\left[g(X^\nu_{t,x}(T)) + \sum_{i\in\mathbb{I}^\nu_{\vartheta,T}} f(X^\nu_{t,x}(\tau^\nu_i+\delta^\nu_i),\mathcal{E}^\nu_i-) \mid \mathcal{F}_\vartheta\right](\omega) = J(\vartheta(\omega),X^\nu_{t,x}(\vartheta)(\omega);\tilde{\nu}_\omega)
$$

with, for each $\omega\in\Omega$,

$$
\tilde{\nu}_\omega \ : \ \tilde{\omega}\in\Omega \mapsto \tilde{\nu}_\omega(\tilde{\omega}) = \nu(\omega^{\vartheta(\omega)} + \mathbf{T}_{\vartheta(\omega)}(\tilde{\omega}))
$$

which can be viewed, for each $\omega\in\Omega$, as a control independent of $\mathcal{F}_{\vartheta(\omega)}$. Since the dynamic of $X^{\tilde{\nu}_\omega}_{\vartheta(\omega),X^\nu_{t,x}(\vartheta)(\omega)}$ depends on $\tilde{\nu}_\omega$ only through its path after $\vartheta(\omega)$, this implies that, for each $\omega\in\Omega$,

$$
\begin{aligned}
J(\vartheta(\omega),X^\nu_{t,x}(\vartheta)(\omega);\tilde{\nu}_\omega) &\le \sup\left\{J(\vartheta(\omega),X^\nu_{t,x}(\vartheta)(\omega);\bar{\nu}),\ \bar{\nu}\in\mathcal{S}^a_{\vartheta(\omega),\Delta^\nu_\vartheta(\omega),\nu_\vartheta(\omega)}\right\}\\
&\le [V^*,g](\vartheta(\omega),X^\nu_{t,x}(\vartheta)(\omega),\Delta^\nu_\vartheta(\omega),\nu_\vartheta(\omega))\ ,
\end{aligned}
$$

and the result follows from (4.10) and (4.11).

2. We now prove the second inequality.

2.a. We first show that, for any $\varepsilon>0$, we can find two sequences $(t_n,x_n,\delta_n,e_n,A_n)_{n\ge1}$ in $D\times\mathscr{B}_D$ and $(\nu^n)_{n\ge1}$ in $\mathcal{S}$ such that $(A_n)_{n\ge1}$ forms a partition of D and, for each n,

$$
(4.12) \qquad
\begin{aligned}
&\nu^n\in\mathcal{S}^a_{t_n,\delta_n,e_n} \ \text{ and } \ J(t',x';\mathscr{P}^{e'}_{t'+\delta',t_n+\delta_n}(\nu^n)) \ \ge\ \varphi(t',x',\delta',e')-3\varepsilon \text{ for all } (t',x',\delta',e')\in A_n\\
&A_n\subset\mathcal{Q}_{r_n}(t_n,x_n,\delta_n,e_n)\cap D \ \text{ for some } r_n>0,
\end{aligned}
$$

where we use the notation

$$
\mathcal{Q}_{\mathring{r}}(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e}) \ :=\ \left\{(t',x',\delta',e')\in B_{\mathring{r}}(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e}) \ :\ t'\le\mathring{t}\ ,\ t'+\delta'\le\mathring{t}+\mathring{\delta}\right\}.
$$

By definition of $V\ge\varphi$, for each $(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e})\in D$ and $\varepsilon>0$, we can find $\mathring{\nu}=\mathring{\nu}^{(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e}),\varepsilon}$ in $\mathcal{S}$ such that

$$
(4.13) \qquad \mathring{\nu}\in\mathcal{S}^a_{\mathring{t},\mathring{\delta},\mathring{e}} \ \text{ and } \ J(\mathring{t},\mathring{x};\mathring{\nu})\ge V(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e})-\varepsilon\ge\varphi(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e})-\varepsilon\ .
$$

Moreover, it follows from Lemma 4.2 and the upper-semicontinuity of φ that we can find $\mathring{r}=\mathring{r}_{(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e})}$ in $(0,\infty)$ such that

$$
(4.14) \qquad J(t',x';\mathscr{P}^{e'}_{t'+\delta',\mathring{t}+\mathring{\delta}}(\mathring{\nu}))\ge J(\mathring{t},\mathring{x};\mathring{\nu})-\varepsilon \text{ and } \varphi(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e})\ge\varphi(t',x',\delta',e')-\varepsilon \quad \text{for all } (t',x',\delta',e')\in\mathcal{Q}_{\mathring{r}}(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e})\ .
$$

Clearly $\{\mathcal{Q}_r(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e})\ :\ (\mathring{t},\mathring{x},\mathring{\delta},\mathring{e})\in D, 0<r\le \mathring{r}_{(\mathring{t},\mathring{x},\mathring{\delta},\mathring{e})}\}$ forms a Vitali covering of D. It then follows from the Vitali's covering Theorem, see e.g. Lemma 1.9 p10 in [9], that we can find a countable sequence $(t_n,x_n,\delta_n,e_n,r_n)_{n\ge1}$ of elements of $D\times\mathbb{R}$, with $0<r_n<\mathring{r}_{(t_n,x_n,\delta_n,e_n)}$ for all $n\ge 1$, such that $D\subset\cup_{n\ge1}\mathcal{Q}_{r_n}(t_n,x_n,\delta_n,e_n)$. We finally construct the sequence $(A_n)_{n\ge1}$ by setting $A_1:=\mathcal{Q}_{r_1}(t_1,x_1,e_1,\delta_1)\cap D$, $C_0=\emptyset$ and

$$A_n := (\mathcal{Q}_{r_n}(t_n,x_n,\delta_n,e_n)\cap D)\backslash C_{n-1}\quad,\quad C_{n-1}:=C_{n-2}\cup A_{n-1}\qquad \text{for } n\ge 2\,.$$

The sequence $(\nu^n)_{n\ge1}$ is defined by $\nu^n := \mathring{\nu}^{(t_n,x_n,\delta_n,e_n),\varepsilon}$ for all n.

2.b. We are now in position to prove (4.2). Let ν be a arbitrary element of $\mathcal{S}^a_{t,\delta,e}$ and define

$$\hat\nu := \nu\mathbf{1}_{[0,\vartheta)}+\mathbf{1}_{[\vartheta,T]}\sum_{n\ge1}\mathscr{P}^{\nu_\vartheta}_{\vartheta+\Delta^\nu_\vartheta,t_n+\delta_n}(\nu^n)\mathbf{1}_{(\vartheta,X^\nu_{t,x}(\vartheta),\Delta^\nu_\vartheta,\nu_\vartheta)\in A_n}\,.$$

Since $\nu\in\mathcal{S}^a_{t,\delta,e}$, we have $(\vartheta,X^\nu_{t,x}(\vartheta),\Delta^\nu_\vartheta,\nu_\vartheta)\in D=\cup_{n\ge1}A_n$. Moreover, on $\{(\vartheta,X^\nu_{t,x}(\vartheta),\Delta^\nu_\vartheta,\nu_\vartheta)\in A_n\}$, we have $\vartheta+\Delta^\nu_\vartheta\le t_n+\delta_n$. It follows that $\hat\nu\in\mathcal{S}^a_{t,\delta,e}$, and therefore

$$\begin{aligned}V(t,x,\delta,e) &\ge J(t,x;\hat\nu)\\ &= \mathbb{E}\left[g(X^{\hat\nu}_{t,x}(T))+\sum_{i\in\mathbb{I}^{\hat\nu}_{\vartheta,T}}f(X^{\hat\nu}_{t,x}(\tau^{\hat\nu}_i+\delta^{\hat\nu}_i-),\mathcal{E}^{\hat\nu}_i)+\sum_{i\in\mathbb{I}^{\hat\nu}_{t,\vartheta}}f(X^{\hat\nu}_{t,x}(\tau^{\hat\nu}_i+\delta^{\hat\nu}_i-),\mathcal{E}^{\hat\nu}_i)\right],\end{aligned}$$

where, by the flow property of $X^{\hat\nu}$, the fact that ν^n is independent of $\mathcal{F}_{t_n}$ and that $\vartheta\le t_n$ on $\{(\vartheta,X^\nu_{t,x}(\vartheta),\Delta^\nu_\vartheta,\nu_\vartheta)\in A_n\}$,

$$\mathbb{E}\left[g(X^{\hat\nu}_{t,x}(T))\right]=\int\int g\left(X^{\tilde\nu_\omega(\tilde\omega)}_{\vartheta(\omega),X^\nu_{t,x}(\vartheta)(\omega)}(T)(\tilde\omega)\right)d\mathbb{P}(\tilde\omega)d\mathbb{P}(\omega)$$

and

$$\begin{aligned}&\mathbb{E}\left[\sum_{i\in\mathbb{I}^{\hat\nu}_{\vartheta,T}}f(X^{\hat\nu}_{t,x}(\tau^{\hat\nu}_i+\delta^{\hat\nu}_i-),\mathcal{E}^{\hat\nu}_i)\right]\\ &=\int\int\sum_{i\in\mathbb{I}^{\tilde\nu_\omega(\tilde\omega)}_{\vartheta(\omega),T}}f(X^{\tilde\nu_\omega(\tilde\omega)}_{\vartheta(\omega),X^\nu_{t,x}(\vartheta)(\omega)}((\tau^{\tilde\nu_\omega}_i+\delta^{\tilde\nu_\omega}_i-)\wedge T)(\tilde\omega),\mathcal{E}^{\tilde\nu_\omega}_i(\tilde\omega))d\mathbb{P}(\tilde\omega)d\mathbb{P}(\omega)\end{aligned}$$

where, for $\omega\in\Omega$,

$$\tilde\nu_\omega\ :\ \tilde\omega\in\Omega\mapsto\nu(\omega)\mathbf{1}_{[0,\vartheta(\omega))}+\mathbf{1}_{[\vartheta(\omega),T]}\sum_{n\ge1}\mathscr{P}^{\nu_\vartheta(\omega)}_{\vartheta(\omega)+\Delta^\nu_\vartheta(\omega),t_n+\delta_n}(\nu^n(\tilde\omega))\mathbf{1}_{(\vartheta(\omega),X^\nu_{t,x}(\vartheta)(\omega),\Delta^\nu_\vartheta(\omega),\nu_\vartheta(\omega))\in A_n}\,.$$

Hence, (4.12) implies that

$$
\begin{aligned}
V(t,x,\delta,e) &\geq \mathbb{E}\left[J(\vartheta, X^{\hat\nu}_{t,x}(\vartheta);\hat\nu) + \sum_{i\in\mathbb{I}^{\hat\nu}_{t,\vartheta}} f(X^{\hat\nu}_{t,x}(\tau^{\hat\nu}_i+\delta^{\hat\nu}_i-),\mathcal{E}^{\hat\nu}_i)\right] \\
&= \mathbb{E}\left[\sum_{n\geq1} J(\vartheta, X^{\nu}_{t,x}(\vartheta); \mathscr{P}^{\nu_\vartheta}_{\vartheta+\Delta^\nu_\vartheta, t_n+\delta_n}(\nu^n))\mathbf{1}_{(\vartheta, X^\nu_{t,x}(\vartheta),\Delta^\nu_\vartheta,\nu_\vartheta)\in A_n}\right] \\
&+ \mathbb{E}\left[\sum_{i\in\mathbb{I}^{\nu}_{t,\vartheta}} f(X^{\nu}_{t,x}(\tau^{\nu}_i+\delta^{\nu}_i-),\mathcal{E}^{\nu}_i)\right] \\
&\geq \mathbb{E}\left[\varphi(\vartheta, X^\nu_{t,x}(\vartheta),\Delta^\nu_\vartheta,\nu_\vartheta) + \sum_{i\in\mathbb{I}^{\nu}_{t,\vartheta}} f(X^{\nu}_{t,x}(\tau^{\nu}_i+\delta^{\nu}_i-),\mathcal{E}^{\nu}_i)\right] - 3\varepsilon.
\end{aligned}
$$

By arbitrariness of $\varepsilon > 0$ and $\nu \in \mathcal{S}^a_{t,\delta,e}$, this proves the required inequality. □

Remark 4.3. Note that, by replacing φ in (4.2) by a sequence $(\varphi_k)_{k\geq1}$ of upper semi-continuous functions satisfying

$$\varphi_k \leq V \quad \text{and} \quad \varphi_k \nearrow [V_*, g] \quad \text{on } D,$$

we can deduce a stronger version of (4.2):

$$\sup_{\nu\in\mathcal{S}^a_{t,\delta,e}} \mathbb{E}\left[[V_*,g](\vartheta^\nu, X^\nu_{t,x}(\vartheta^\nu),\Delta^\nu_{\vartheta^\nu},\nu_{\vartheta^\nu}) + \sum_{i\in\mathbb{I}^{\nu}_{t,\vartheta}} f(X^{\nu}_{t,x}(\tau^{\nu}_i+\delta^{\nu}_i-),\mathcal{E}^{\nu}_i)\right] \leq V(t,x,\delta,e)\ ,$$

where $[V_*,g](s,\cdot) := V_*(s,\cdot)\mathbf{1}_{s<T} + g\mathbf{1}_{s=T}$. In particular, if V is continuous, combining (4.1) and the previous inequality leads to the classical version of the dynamic programming principle (3.6).

4.2 Viscosity properties

Now we are in position to prove Theorem 3.2. We split the proof in different propositions.

4.2.1 Supersolution property

We start with the supersolution property in the domain $D = D_{E,>0} \cup D_\varpi$, recall the definitions (3.1)-(3.2).

Proposition 4.4. *The function V_* is a viscosity supersolution of (3.7) on D.*

Proof. The proof follows from standard arguments except that we use the non classical formulation of the dynamic programming principle (4.2). Fix $(t_0,x_0,\delta_0,e_0) \in D$ and let φ be a smooth function such that (t_0,x_0,δ_0,e_0) achieves a (global) minimum of $V_* - \varphi$ on D such that

$$0 = (V_* - \varphi)(t_0,x_0,\delta_0,e_0).$$

Let $(t_k, x_k, \delta_k, e_k)_{k\geq 1}$ be a sequence in D such that
(4.15)

$$(t_k, x_k, \delta_k, e_k) \longrightarrow (t_0, x_0, \delta_0, e_0) \text{ and } V(t_k, x_k, \delta_k, e_k) \longrightarrow V_*(t_0, x_0, \delta_0, e_0) \text{ as } k \longrightarrow \infty \,,$$

and observe that

$$(\varphi - V)(t_k, x_k, \delta_k, e_k) \longrightarrow 0 \text{ when } k \longrightarrow \infty \,. \tag{4.16}$$

Case 1. We first assume that $(t_0, x_0, \delta_0, e_0) \in D_{E,>0}$, recall the definition (3.1), and that

$$-\mathcal{L}^{e_0}\varphi(t_0, x_0, \delta_0, e_0) + \frac{\partial}{\partial \delta}\varphi(t_0, x_0, \delta_0, e_0) =: -2\varepsilon < 0 \,, \tag{4.17}$$

and work towards a contradiction. Define the function $\bar{\varphi}$ by

$$\bar{\varphi}(t, x, \delta, e) := \varphi(t, x, \delta, e) - |x - x_0|^4 - |t - t_0|^2 - |\delta - \delta_0|^2 \,, \tag{4.18}$$

so that $\bar{\varphi}$ also satisfies (4.17). By continuity of b and a, we can find $r > 0$ such that

$$\left(-\mathcal{L}^{e}\bar{\varphi} + \frac{\partial}{\partial \delta}\bar{\varphi}\right)(t, x, \delta, e) \leq 0 \text{ for } (t, x, \delta, e) \in B := B_r(t_0, x_0, \delta_0, e_0) \cap D_{E,>0} \,. \tag{4.19}$$

Given k large enough so that $(t_k, x_k, \delta_k, e_k) \in B$, let ν^k be any control in $\mathcal{S}^a_{t_k,\delta_k,e_k}$. Set $(X^k, \Delta^k) := (X^{\nu^k}_{t_k,x_k}, \Delta^{\nu^k})$ and define

$$\theta^k \;:=\; \inf\{s \geq t_k \;:\; (s, X^k(s), \Delta^k_s, \nu^k_s) \notin B\} \,.$$

For r small enough we have $\Delta^{\nu^k}_{\theta^k} > 0$ and therefore $\nu^k = e_k$ on $[t_k, \theta^k]$. Using Itô's Lemma, (4.19) and the definition of $\bar{\varphi}$, we thus obtain that

$$\bar{\varphi}(t_k, x_k, \delta_k, e_k) \leq \mathbb{E}\left[\bar{\varphi}(\theta^k, X^k(\theta^k), \Delta^k_{\theta^k}, e_k)\right] \leq \mathbb{E}\left[\varphi(\theta^k, X^k(\theta^k), \Delta^k_{\theta^k}, \nu^k_{\theta^k})\right] - \eta$$

where $\eta := \inf\{|x - x_0|^4 + |t - t_0|^2 + |\delta - \delta_0|^2,\ (t, x, \delta, e_k) \in \partial B_r(t_0, x_0, \delta_0, e_0),\ k \geq 1\} > 0$, observe that $|e_k - e_0| < r$. Since $I^{\nu^k}_{t_k,\theta^k} = \emptyset$, the above inequality combined with (4.16) and (4.18) contradict (4.2) for k large enough.

Case 2. We now assume that $(t_0, x_0, \delta_0, e_0) \in D_{\varpi}$, recall the definition (3.2). Since E is closed and $\varpi \notin E$, (4.15) implies that

$$(\delta_k, e_k) = (0, \varpi) \;\text{ for } k \text{ large enough.} \tag{4.20}$$

We now assume that

$$\min\left\{-\mathcal{L}^{\varpi}\varphi(t_0, x_0, 0, \varpi)\,,\; \varphi(t_0, x_0, 0, \varpi) - \mathcal{M}[\varphi](t_0, x_0)\right\} \;=: -2\varepsilon < 0 \,, \tag{4.21}$$

and work toward a contradiction. If

$$-\mathcal{L}^{\varpi}\varphi(t_0, x_0, 0, \varpi) = -2\varepsilon < 0 \,,$$

we can argue as above to obtain a contradiction to (4.2). If

$$\varphi(t_0, x_0, 0, \varpi) - \mathcal{M}[\varphi](t_0, x_0) =: -2\varepsilon < 0 \,,$$

we can find $(\hat{\delta}, \hat{e}) \in [\underline{\delta}, T - t_0) \times E$ and $r > 0$ such that

(4.22) $\varphi(t, x, 0, \varpi) - \varphi(t, x + \beta(x, \hat{e}, \hat{\delta}), \hat{\delta}, \hat{e}) \le -\varepsilon$, for $(t, x) \in B := B_r(t_0, x_0)$.

Let $\bar{\nu}$ denote the constant control that takes the value ϖ on $[0, T]$, set $\bar{X}^k := X^{\bar{\nu}}_{t_k, x_k}$ and

$$\theta^k \ := \ \inf\{s \ge t_k \ : \ (s, \bar{X}^k(s)) \notin B\} \wedge (t_k + k^{-1}) \ .$$

Note that for k large enough, we have $t_k + k^{-1} + \hat{\delta} \le T$. We can then define $\nu^k \in \mathcal{S}^a_{t_k,0,\varpi}$ by

$$\nu^k_t := \hat{e}\mathbf{1}_{t \in [\theta^k, \theta^k + \hat{\delta})} + \varpi \mathbf{1}_{t \notin [\theta^k, \theta^k + \hat{\delta})} \ , \ t \le T \ ,$$

and set $(X^k, \Delta^k) := (X^{\nu^k}_{t_k, x_k}, \Delta^{\nu^k})$. Using Itô's Lemma and (4.22), we obtain that

$$\varphi(t_k, x_k, 0, \varpi) \ \le \ \mathbb{E}\left[\varphi(\theta^k, X^k(\theta^k), \Delta^k_{\theta^k}, \nu^k_{\theta^k}) - \varepsilon\right] + C/k \ ,$$

for some $C > 0$ which does not depend on k. The above inequality combined with (4.16) contradict (4.2) for k large enough.

□

We now turn to the proof of the boundary conditions. We refer to (3.3)-(3.4)-(3.5) for the definitions of $D_{E,0}$, $D_{E,T}$ and D_T.

Proposition 4.5. *Fix $(t_0, x_0, \delta_0, e_0) \in \bar{D}$. Then,*

$$V_*(t_0, x_0, 0, e_0) \ \ge \ \begin{cases} V_*(t_0, x_0, 0, \varpi) + f(x_0, e_0) & \text{if} \quad (t_0, x_0, \delta_0, e_0) \in D_{E,0} \\ \mathcal{V}(t_0, x_0, e_0) & \text{if} \quad (t_0, x_0, \delta_0, e_0) \in D_{E,T} \\ g(x_0) + f(x_0, e_0) & \text{if} \quad (t_0, x_0, \delta_0, e_0) \in D_T \ . \end{cases}$$

Proof. We only prove the first inequality. The two other ones follow from similar arguments. Let $(t_k, x_k, \delta_k, e_k)_{k \ge 1}$ be a sequence in D such that

(4.23) $(t_k, x_k, \delta_k, e_k) \longrightarrow (t_0, x_0, 0, e_0)$ and $V(t_k, x_k, \delta_k, e_k) \longrightarrow V_*(t_0, x_0, 0, e_0)$ as $k \longrightarrow \infty$.

For each $k \ge 1$, define

$$\nu^k := \varpi \mathbf{1}_{[0, t_k + \delta_k)} + e_k \mathbf{1}_{[t_k + \delta_k, T]} \in \mathcal{S}^a_{t_k, \delta_k, e_k} \ ,$$

and set $X^k := X^{\nu^k}_{t_k, x_k}$. It follows from Remark 4.3 that

(4.24) $V(t_k, x_k, \delta_k, e_k) \ge \mathbb{E}\left[V_*(t_k + \delta_k, X^k(t_k + \delta_k), 0, \varpi) + f(X^k(t_k + \delta_k -), e_k)\right] \ , \ k \ge 1$.

Using standard estimates, see e.g. the proof of Lemma 4.2, one easily checks that $X^k(t_k + \delta_k -) \to x_0$ in L^p for all $p \ge 1$, and in particular $\mathbb{P}$-a.s., after possibly passing to a subsequence. It thus follows from the lower-semicontinuity of V_* and f that, up to a subsequence,

$$\liminf_{k \to \infty} V_*(t_k + \delta_k, X^k(t_k + \delta_k), 0, \varpi) + f(X^k(t_k + \delta_k -), e_k) \ge V_*(t_0, x_0, 0, \varpi) + f(x_0, e_0) \ \mathbb{P}\text{-a.s.}$$

The required result then follows from (4.23), (4.24), and the last inequality combined with polynomial growth property of f and V, see Remark 2.5, and Fatou's Lemma. □

4.2.2 Subsolution property

We start with the subsolution property in the domain $D = D_{E,>0} \cup D_{\varpi}$, recall the definitions in (3.1)-(3.2).

Proposition 4.6. *The function V^* is a viscosity subsolution of (3.7) on D.*

Proof. Fix $(t_0, x_0, \delta_0, e_0) \in D$ and let φ be a smooth function such that $(t_0, x_0, \delta_0, e_0)$ achieves a (global) maximum of $V^* - \varphi$ such that

$$0 = (V^* - \varphi)(t_0, x_0, \delta_0, e_0) .$$

In the following, we denote by $(t_k, x_k, \delta_k, e_k)_{k\geq 1}$ a sequence in D satisfying

$$(t_k, x_k, \delta_k, e_k) \longrightarrow (t_0, x_0, \delta_0, e_0) \text{ and } V(t_k, x_k, \delta_k, e_k) \longrightarrow V^*(t_0, x_0, \delta_0, e_0) \text{ as } k \longrightarrow \infty . \tag{4.25}$$

Case 1. We first assume that $(t_0, x_0, \delta_0, e_0) \in D_{E,>0}$ and that

$$-\mathcal{L}^{e_0}\varphi(t_0, x_0, \delta_0, e_0) + \frac{\partial}{\partial \delta}\varphi(t_0, x_0, \delta_0, e_0) =: 2\varepsilon > 0 ,$$

and work towards a contradiction. By continuity of b and a, we can find $r > 0$ such that

$$\left(-\mathcal{L}^{e}\varphi + \frac{\partial}{\partial \delta}\varphi\right)(t, x, \delta, e) \geq \varepsilon \quad \text{for } (t, x, \delta, e) \in B := B_r(t_0, x_0, \delta_0, e_0) \cap D_{E,>0} . \tag{4.26}$$

Moreover, we can always assume that $(t_0, x_0, \delta_0, e_0)$ achieves a strict local maximum, so that after possibly changing the value of ε, we have

$$\sup_{\partial_p B}(V^* - \varphi) =: -\varepsilon < 0 , \tag{4.27}$$

where $\partial_p B$ is the parabolic boundary of B. Fix $\nu^k \in \mathcal{S}_{t_k,\delta_k,e_k}$ and set

$$\theta^k := \inf\{s \geq t_k \;:\; (s, X^k(s), \Delta^k_s, \nu^k_s) \notin B\} ,$$

where $(X^k, \Delta^k) := (X^{\nu^k}_{t_k,x_k}, \Delta^{\nu^k})$. Observe that, for r small enough, $\Delta^k_{\theta^k} > 0$ and therefore $\nu^k = e_k$ on $[t_k, \theta^k]$. Applying Itô's Lemma to φ and using (4.26) and (4.27), we deduce that

$$\varphi(t_k, x_k, \delta_k, e_k) \geq \mathbb{E}\left[\varphi(\theta^k, X^k(\theta^k), \Delta^k_{\theta^k}, \nu_{\theta^k})\right] \geq \mathbb{E}\left[V^*(\theta^k, X^k(\theta^k), \Delta^k_{\theta^k}, \nu_{\theta^k})\right] + \varepsilon .$$

Since $\mathbb{I}^{\nu^k}_{t_k,\theta^k} = \emptyset$, this contradicts (4.1) for k large enough, recall (4.25) .
Case 2. We now assume that $(t_0, x_0, \delta_0, e_0) = (t_0, x_0, 0, \varpi) \in D_{\varpi}$ and

$$\min\{-\mathcal{L}^{\varpi}\varphi(t_0, x_0, 0, \varpi),\; \varphi(t_0, x_0, 0, \varpi) - \mathcal{M}[\varphi](t_0, x_0)\} =: 2\varepsilon > 0 ,$$

and work towards a contradiction. By continuity of b and a, we can find $r > 0$ such that

$$\min\{-\mathcal{L}^{\varpi}\varphi(\cdot, 0, \varpi),\; \varphi(\cdot, 0, \varpi) - \mathcal{M}[\varphi]\} \geq \varepsilon \;\text{ on } B := B_r(t_0, x_0) \cap ([0, T) \times \mathbb{R}^d) . \tag{4.28}$$

Moreover, without loss of generality we can assume that (t_0, x_0) achieves a strict local maximum, so that after possibly changing the value of ε

$$\sup_{\partial_p B}(V^*(\cdot, 0, \varpi) - \varphi(\cdot, 0, \varpi)) =: -\varepsilon < 0 , \tag{4.29}$$

where $\partial_p B$ is the parabolic boundary of B. Also observe that, since E is closed and $\varpi \notin E$, (4.25) implies that

$$(4.30) \qquad (\delta_k, e_k) = (0, \varpi) \;\text{ for } k \text{ large enough.}$$

Let $\nu^k \in \mathcal{S}^a_{t_k,0,\varpi} = \mathcal{S}^a_{t_k,\delta_k,e_k}$ be arbitrary, set $(X^k, \Delta^k, (\tau^k_i)_{i\geq 1}) := (X^{\nu^k}_{t_k,x_k}, \Delta^{\nu^k}, (\tau^{\nu^k}_i)_{i\geq 1})$ and define

$$\theta^k := \inf\{s \geq t_k \;:\; (s, X^k(s)) \notin B\} \;,\; \vartheta^k := \inf\{\tau^k_i,\; i \geq 1 \text{ s.t. } \tau^k_i \geq t_k\} \text{ and } \xi^k := \theta^k \wedge \vartheta^k \;.$$

Applying Itô's Lemma to φ, using (4.28), (4.29), and recalling (4.30) lead to

$$\begin{aligned} \varphi(t_k, x_k, \delta_k, e_k) \;&\geq\; \mathbb{E}\left[\varphi(\xi^k, X^k(\xi^k-), 0, \varpi)\right] \\ &\geq\; \mathbb{E}\left[\varphi(\xi^k, X^k(\xi^k), \Delta^k_{\xi^k}, \nu^k_{\xi^k}) + \varepsilon \mathbf{1}_{\xi^k=\vartheta^k}\right] \\ &\geq\; \mathbb{E}\left[V^*(\xi^k, X^k(\xi^k), \Delta^k_{\xi^k}, \nu^k_{\xi^k})\right] + \varepsilon \;. \end{aligned}$$

In view of (4.25) this leads to a contradiction with (4.1) for k large enough. □

We now turn to the boundary condition $\delta \to 0$. Recall the definition of $D_{E,0}$ in (3.3).

Proposition 4.7. *For all $(t_0, x_0, \delta_0, e_0) \in D_{E,0}$, we have*

$$V^*(t_0, x_0, 0, e_0) \leq V^*(t_0, x_0, 0, \varpi) + f(x_0, e_0) \;.$$

Proof. By following similar arguments as in the second step of the proof of Proposition 4.6 above, one easily checks that, for any smooth function $\bar\varphi$ such that $(t_0, x_0, 0, e_0)$ achieves a global maximum of $V^* - \bar\varphi$ satisfying $(V^* - \bar\varphi)(t_0, x_0, 0, e_0) = 0$, we have

$$\min\left\{-\mathcal{L}^{e_0}\bar\varphi(t_0, x_0, 0, e_0) + \frac{\partial}{\partial\delta}\bar\varphi(t_0, x_0, 0, e_0) \;,\; \bar\varphi(t_0, x_0, 0, e_0) - \bar\varphi(t_0, x_0, 0, \varpi) - f(x_0, e_0)\right\} \leq 0 \;.$$

Let now φ be a smooth function such that $(t_0, x_0, 0, e_0)$ achieves a global maximum of $V^* - \varphi$ satisfying $(V^* - \varphi)(t_0, x_0, 0, e_0) = 0$, and consider the function $\bar\varphi_\varepsilon$ defined as

$$\bar\varphi_\varepsilon(t, x, \delta, e) := \varphi(t, x, \delta, e) + \sqrt{\varepsilon + \delta} - \sqrt{\varepsilon}$$

for some $\varepsilon > 0$. Observe that $(t_0, x_0, 0, e_0)$ achieves a global maximum of $V^* - \bar\varphi_\varepsilon$. It thus follows that either

$$\bar\varphi_\varepsilon(t_0, x_0, 0, e_0) - \bar\varphi_\varepsilon(t_0, x_0, 0, \varpi) - f(x_0, e_0) \leq 0$$

or

$$-\mathcal{L}^{e_0}\bar\varphi_\varepsilon(t_0, x_0, 0, e_0) + \frac{\partial}{\partial\delta}\bar\varphi_\varepsilon(t_0, x_0, 0, e_0) + \varepsilon^{-\frac{1}{2}} \leq 0 \;.$$

Clearly, the second assertion can not hold for $\varepsilon > 0$ small enough. It follows that

$$\bar\varphi_\varepsilon(t_0, x_0, 0, e_0) - \bar\varphi_\varepsilon(t_0, x_0, 0, \varpi) - f(x_0, e_0) \leq 0$$

for all $\varepsilon > 0$ small enough, which provides the required result. □

We next consider the boundary conditions as $t + \delta \to T$ and $t = T$. Recall the definitions of $D_{E,T}$ and D_T in (3.4)-(3.5).

Proposition 4.8. *For all* $(t_0, x_0, \delta_0, e_0) \in D_{E,T}$*, we have*

$$V^*(t_0, x_0, \delta_0, e_0) \leq \mathcal{V}(t_0, x_0, e_0) \ .$$

For all $(T, x_0, \delta_0, e_0) \in D_T$*, we have*

$$V^*(T, x_0, \delta_0, e_0) \leq g(x_0) + f(x_0, e) \ .$$

Proof. We only prove the first assertion, the second one using the same kind of arguments. Let $(t_k, x_k, \delta_k, e_k)_{k\geq 1}$ be a sequence in D satisfying

$$(t_k, x_k, \delta_k, e_k) \longrightarrow (t_0, x_0, \delta_0, e_0) \text{ and } V(t_k, x_k, \delta_k, e_k) \longrightarrow V^*(t_0, x_0, \delta_0, e_0) \text{ as } k \longrightarrow \infty \ .$$

For k large enough, we have $t_k + \delta_k > T - \underline{\delta}$ so that, for any $\nu^k \in \mathcal{S}^a_{t_k,\delta_k,e_k}$, we have $\nu^k = e_k$ on $[t_k, t_k + \delta_k)$ and $\nu^k = \varpi$ on $[t_k + \delta_k, T]$, recall (2.1). It follows that

$$V(t_k, x_k, \delta_k, e_k) = \mathbb{E}\left[g(X^{\nu^k}_{t_k,x_k}(T)) + f(\mathcal{X}^{e_k}_{t_k,x_k}(t_k + \delta_k), e_k)\right] \ ,$$

recall the definition of $\mathcal{X}$ in (2.10). Moreover, since $t_k + \delta_k \to t_0 + \delta_0 = T$, standard estimates based on (2.5) and Remark 2.3 imply that $(X^{\nu^k}_{t_k,x_k}(T), \mathcal{X}^{e_k}_{t_k,x_k}(t_k+\delta_k)) \to (\mathcal{X}^{e_0}_{t_0,x_0}(T), \mathcal{X}^{e_0}_{t_0,x_0}(t_0+\delta_0))$ as $k \to \infty$ in L^p for all $p \geq 1$. The result then follows by taking the limsup as $k \to \infty$ in the above equality and using (2.7) as well as the dominated convergence theorem. □

4.3 A comparison result

In this section, we provide the proof of Theorem 3.3. We first show that (3.7) admits a classical strict supersolution in the following sense:

Proposition 4.9. *For any integer* $p \geq \gamma$*, there exists a function* Λ *on* $\mathbb{R}^d \times \mathbb{R}_+ \times \bar{E}$ *and* $\varrho > 0$ *satisfying*

(i) $\Lambda \in C^{2,1,0}(\mathbb{R}^d \times \mathbb{R}_+ \times \bar{E})$*,*

(ii) $\Lambda \geq g^+ + f^+ + \mathcal{V}^+$ *,*

(iii) $\inf_{(\delta,e)\in[0,T]\times\bar{E}} \Lambda(x, \delta, e)/|x|^p \to \infty$ *as* $|x| \to \infty$ *,*

(iv) $\langle b, D\Lambda\rangle + \frac{1}{2}\mathrm{Tr}[aa^* D^2\Lambda] \leq \varrho\Lambda$ *and* $-\partial\Lambda/\partial\delta + \langle b, D\Lambda\rangle + \frac{1}{2}\mathrm{Tr}[aa^*D^2\Lambda] \leq \varrho\Lambda$ *on* $\mathbb{R}^d \times \mathbb{R}_+ \times \bar{E}$*,*

(v) $\Lambda(x, 0, e) - f(x, e) - q(x) \geq \Lambda(x, 0, \varpi) \geq \Lambda(x + \beta(x, e, \delta), \delta, e) + \underline{\delta}$ *for all* $x \in \mathbb{R}^d$*,* $\delta \in [\underline{\delta}, T]$ *and* $e \in E$*, where* q *is a continuous and (strictly) positive function on* $\mathbb{R}^d$*.*

Proof. Let φ be a $C^1(\mathbb{R}_+)$ function with bounded first derivative such that $\varphi \geq 0$, $\varphi(0) = 1$ and $\varphi(\delta) = 0$ for $\delta \geq \underline{\delta}$ and let Λ be defined by:

$$\Lambda(x, \delta, e) \ = \ \mu\ (1 + |x|^{2p})(1 + \kappa\mathbf{1}_{e=\varpi}) + \big(2\kappa\mu\ (1 + |x|^{2p}) - \delta\big)\ \varphi(\delta)\mathbf{1}_{e\neq\varpi}$$

for some $\kappa > 0$ and $\mu > T$ such that $f(x, e)^+ + g(x)^+ + \mathcal{V}(t, x, e)^+ \leq \mu\ (1 + |x|^{2p})$, recall (2.7) and Remark 2.5. The first three assertions clearly hold for μ large enough, recall that $\varpi \notin E$ where E is closed. As for the fourth assertion, we recall that b and a are uniformly Lipschitz, so that the left-hand side is of order $(1 + |x|^{2p})$, which is dominated by $\varrho\Lambda$ for ϱ

large enough. The right-hand side inequality holds for ϱ large enough by similar arguments. Finally, recalling (2.5), we observe that, for μ and κ large enough,

$$\begin{aligned} \Lambda(x,0,e) - f(x,e) - \Lambda(x,0,\varpi) &= \kappa\mu\,(1+|x|^{2p}) - f(x,e) \geq \mu\,(1+|x|^{2p})\,, \\ \Lambda(x,0,\varpi) - \Lambda(x+\beta(x,e,\delta),\delta,e) &= \mu\,(1+|x|^{2p})(1+\kappa) - \mu\,(1+|x+\beta(x,e,\delta)|^{2p}) \\ &\geq \mu\kappa/2 \geq \underline{\delta} \end{aligned}$$

for all $x \in \mathbb{R}^d$, $\underline{\delta} \leq \delta \leq T$ and $e \in E$, which provides the last assertion for $q(x) := \mu\,(1+|x|^{2p})$. □

We can now provide the proof of Theorem 3.3.

Proof. [Theorem 3.3] Let u and v be as in Theorem 3.3 and let $p \geq \gamma$ be large enough so that

$$[v(t,x,\delta,e) - u(t,x,\delta,e)]^+ \leq C(1+|x|^p) \;\text{ on } \bar{D} \tag{4.31}$$

for some $C > 0$. We assume that

$$\sup_{\bar{D}}(v-u) \geq 2\eta \text{ for some } \eta > 0 \tag{4.32}$$

and work toward a contradiction.

1. Let $\varrho > 0$ and Λ be as in Proposition 4.9 for $p \geq \gamma$ satisfying (4.31). It follows from (4.32) that for $\lambda \in (0,1)$ small enough there is some $(t_\lambda,x_\lambda,\delta_\lambda,e_\lambda) \in \bar{D}$ such that

$$\max_{\bar{D}}(\tilde{v}-\tilde{w}) \;=\; (\tilde{v}-\tilde{w})(t_\lambda,x_\lambda,\delta_\lambda,e_\lambda) \;\geq\; \eta \;> 0\,, \tag{4.33}$$

where for a map φ on $\bar{D}$, we write $\tilde{\varphi}(t,x,\delta,e)$ for $e^{\varrho t}\varphi(t,x,\delta,e)$, and $w := (1-\lambda)u + \lambda\Lambda$. Observe that $\tilde{u}$, $\tilde{v}$ are super- and subsolution of

$$\tilde{\mathcal{H}}\varphi = 0 \tag{4.34}$$

on $\bar{D}$, where, for a smooth function φ defined on $\bar{D}$,

$$\tilde{\mathcal{H}}\varphi(t,x,\delta,e) := \begin{cases} \left(\varrho\frac{\partial}{\partial t} - \mathcal{L}^e + \frac{\partial}{\partial\delta}\right)\varphi(t,x,\delta,e) & \text{on } D_{E,>0}\,, \\ \varphi(t,x,\delta,e) - \varphi(t,x,0,\varpi) - \tilde{f}(x,e) & \text{on } D_{E,0}\,, \\ \varphi(t,x,\delta,e) - \tilde{\mathcal{V}}(t,x,e) & \text{on } D_{E,T}\,, \\ \min\left\{(\varrho\frac{\partial}{\partial t} - \mathcal{L}^\varpi)\varphi(t,x,\delta,e)\,;\;\varphi(t,x,\delta,e) - \mathcal{M}[\varphi](t,x)\right\} & \text{on } D_\varpi\,, \\ \varphi(t,x,\delta,e) - \tilde{g}(x) - \tilde{f}(x,e) & \text{on } D_T\,. \end{cases} \tag{4.35}$$

Also note that

$$(t_\lambda,x_\lambda,\delta_\lambda,e_\lambda) \notin D_{E,T} \cup D_T \tag{4.36}$$

since otherwise the super- and subsolution property of u and v would imply

$$(v-w)(t_\lambda,x_\lambda,\delta_\lambda,e_\lambda) \;\leq\; \lambda\left(\mathcal{V}(t_\lambda,x_\lambda,e_\lambda) \vee (g(x_\lambda)+f(x_\lambda,e_\lambda)) - \Lambda(t_\lambda,x_\lambda,\delta_\lambda,e_\lambda)\right)$$

which, in view of (ii) in Proposition 4.9, would contradict (4.33).

2. For $(t,x,\delta,e) \in \bar{D}$ and $n \geq 1$, we now set

$$\begin{aligned} \Gamma(t,x,y,\delta,e) &:= \tilde{v}(t,x,\delta,e) - \tilde{w}(t,y,\delta,e) \\ \Theta_n(t,x,y,\delta,e) &:= \Gamma(t,x,y,\delta,e) - \varphi_n(t,x,y,\delta,e)\,, \end{aligned}$$

where

$$\varphi_n(t,x,y,\delta,e) \;:=\; n|x-y|^{2p} + |x-x_\lambda|^{2p+2} + |t-t_\lambda|^2 + |\delta-\delta_\lambda|^2 + |e-e_\lambda| \,.$$

By the growth assumption on v, u and the fact that $\bar{E}$ is compact, there is $(t_n,x_n,y_n,\delta_n,e_n) \in \bar{D}^2$, with

$$\bar{D}^2 \;:=\; \{(t,x,y,\delta,e) \;:\; ((t,x,\delta,e),(t,y,\delta,e)) \in \bar{D}\times\bar{D}\} \,,$$

such that

$$\max_{\bar{D}^2} \Theta_n \;=\; \Theta_n(t_n,x_n,y_n,\delta_n,e_n) \,.$$

Since

$$\Gamma(t_n,x_n,y_n,\delta_n,e_n) \geq \Theta_n(t_n,x_n,y_n,\delta_n,e_n) \geq (\tilde{v}-\tilde{w})(t_\lambda,x_\lambda,\delta_\lambda,e_\lambda) \,,$$

it follows from the growth condition on v and u, (iii) of Proposition 4.9 and the upper-semicontinuity of Λ that, up to a subsequence,

$$(t_n,x_n,y_n,\delta_n,e_n) \;\longrightarrow\; (t_\lambda,x_\lambda,x_\lambda,\delta_\lambda,e_\lambda) \tag{4.37}$$

$$\varphi_n(t_n,x_n,y_n,\delta_n,e_n) \;\longrightarrow\; 0 \tag{4.38}$$

$$\Gamma(t_n,x_n,y_n,\delta_n,e_n) \;\longrightarrow\; \Gamma(t_\lambda,x_\lambda,x_\lambda,\delta_\lambda,e_\lambda) \,. \tag{4.39}$$

3. It follows from (4.36) and (4.37) that, after possibly passing to a subsequence,

$$(t_n,x_n,\delta_n,e_n) \notin D_{E,T} \cup D_T \quad \text{for all } n \geq 1 \,. \tag{4.40}$$

3.1. We now assume that, up to a subsequence, $e_n \neq \varpi$ for all $n \geq 1$.
If $|\{n : \delta_n = 0\}| = \infty$, then we can assume that, up to a subsequence, $\delta_n = 0$, i.e., $(t_n,x_n,\delta_n,e_n) \in D_{E,0}$ for all $n \geq 1$. It then follows from the super- and subsolution property of $\tilde{u}$ and $\tilde{v}$, and (v) of Proposition 4.9, that

$$\begin{aligned}
\tilde{v}(t_n,x_n,0,e_n) &\leq \tilde{v}(t_n,x_n,0,\varpi) + \tilde{f}(x_n,e_n)\\
\tilde{u}(t_n,y_n,0,e_n) &\geq \tilde{u}(t_n,y_n,0,\varpi) + \tilde{f}(y_n,e_n)\\
\tilde{\Lambda}(y_n,0,e_n) &\geq \tilde{\Lambda}(y_n,0,\varpi) + \tilde{f}(y_n,e_n) + e^{\varrho t_n}q(y_n) \,,
\end{aligned}$$

and therefore

$$\Gamma(t_n,x_n,y_n,0,e_n) \leq \Gamma(t_n,x_n,y_n,0,\varpi) + \Big(\tilde{f}(x_n,e_n) - \tilde{f}(y_n,e_n)\Big) - \lambda e^{\varrho t_n} q(y_n) \,. \tag{4.41}$$

Sending $n \to \infty$ and using (4.37) and (4.39) leads to

$$\Gamma(t_\lambda,x_\lambda,x_\lambda,\delta_\lambda,e_\lambda) \leq \Gamma(t_\lambda,x_\lambda,x_\lambda,0,\varpi) - \lambda e^{\varrho t_\lambda} q(y_\lambda)$$

which, recalling that $q > 0$ on $\mathbb{R}^d$, contradicts (4.33).
It follows from the above arguments that $|\{n : \delta_n = 0\}| < \infty$. In this case, $(t_n,x_n,\delta_n,e_n) \in D_{E,>0}$ for all $n \geq 1$ large enough, recall (4.40). Using the viscosity property of $\tilde{u}$ and $\tilde{v}$, (iv) in Proposition 4.9, and standard arguments based on Ishii's Lemma, see [8], together with (4.37) allows to deduce that

$$\varrho\Gamma(t_n,x_n,y_n,\delta_n,e_n) \leq O_\lambda(n^{-1}) \,,$$

where $O_\lambda(n^{-1}) \to 0$ as $n \to \infty$. Since $\varrho > 0$, combining the above inequality with (4.39) leads to a contradiction to (4.33).

3.2. We now assume that, up to a subsequence, $e_n = \varpi$, so that $(t_n, x_n, \delta_n, e_n) \in D_\varpi$ for all $n \geq 1$. Note that we can not have

$$\tilde{v}(t_n, x_n, 0, \varpi) - \sup_{(\delta,e)\in[\underline{\delta},T-t_n]\times E} \tilde{v}(t_n, x_n, \delta, e) \leq 0$$

along a subsequence, since otherwise the supersolution property of $\tilde{u}$ and (v) of Proposition 4.9 would imply

$$\Gamma(t_n, x_n, y_n, 0, \varpi) \leq \sup_{(\delta,e)\in[\underline{\delta},T-t_n]\times E} \Gamma(t_n, x_n, y_n, \delta, e) - \lambda e^{\varrho t_n} \underline{\delta} ,$$

which would contradict (4.33) for n large enough, recall (4.37), (4.39) and the fact that $\underline{\delta} > 0$. We can thus assume that $\tilde{v}(t_n, x_n, 0, \varpi) - \sup_{(\delta,e)\in[\underline{\delta},T-t_n]\times E} \tilde{v}(t_n, x_n, \delta, e) > 0$ for n large enough. Using again the viscosity properties of $\tilde{u}$ and $\tilde{v}$, (iv) of Proposition 4.9, standard arguments based on Ishii's Lemma, see [8], and (4.37) then leads

$$\varrho\Gamma(t_n, x_n, y_n, \delta_n, e_n) \leq O_\lambda(n^{-1}) ,$$

where $O_\lambda(n^{-1}) \to 0$ as $n \to \infty$. As above, this leads to a contradiction. □

5 Numerical approximation

In this section, we construct a finite difference scheme to solve the PDE (3.7) numerically, and prove the convergence of the numerical scheme.

5.1 Space discretization

Given a positive integer N, we discretize the set

$$\mathcal{T} := \{(t, \delta) \in [0, T] \times \mathbb{R}_+ \ : \ \underline{\delta} \leq t + \delta \leq T\}$$

in

$$\mathcal{T}_N := \{(ih_N \wedge T, (i + j)h_N), \ i = 0, \ldots, N \text{ and } j = 0, \ldots, N - i\}$$

where

$$h_N := T/N .$$

We next fix a positive integer M and $R > 0$ and approximate $\mathcal{R}^d_R := B_R(0) \subset \mathbb{R}^d$ by

$$\mathcal{R}^d_{MR} := \{-R + kh_{MR}, \ k = 0, \ldots, 2M\}^d$$

where

$$h_{MR} := R/M .$$

We finally consider an increasing sequence $(E_L)_{L\geq1}$ of finite subsets of E such that

$$\cup_{L\geq1} E_L = E .$$

For ease of notations, we set

$$\bar{D}^{RL}_{NM} := \left\{(t, x, \delta, e) \in \bar{D} \ : \ (t, \delta) \in \mathcal{T}_N, \ x \in \mathcal{R}^d_{MR} \ , \ e \in E_L \cup \{\varpi\}\right\} .$$

5.2 Finite difference approximation

From now on, we denote by x^i the i-th component of a vector $x \in \mathbb{R}^d$, and by A^{ij} the (i,j)-component of a matrix $A \in \mathbb{M}^d$. We use the notation l_i for the unit vector of $\mathbb{R}^d$ in the i^{th} coordinate direction.

We use the standard finite difference approximation, see [10] for a full description.

Space component:

$$\frac{\partial \psi}{\partial x^i}(t,x,\delta,e) \sim \begin{cases} \frac{\psi(t,x+h_{MR}l_i,\delta,e)-\psi(t,x,\delta,e)}{h_{MR}} & =: \ \Delta_{x^i}^{h_{MR}+}\psi(t,x,\delta,e) \text{ if } b^i(x,e)\geq 0 \\ \frac{\psi(t,x,\delta,e)-\psi(t,x-h_{MR}l_i,\delta,e)}{h_{MR}} & =: \ \Delta_{x^i}^{h_{MR}-}\psi(t,x,\delta,e) \text{ if } b^i(x,e) < 0 \end{cases},$$

$$\frac{\partial^2 \psi}{\partial x_i^2}(t,x,\delta,e) \sim \frac{\psi(t,x+h_{MR}l_i,\delta,e) - 2\psi(t,x,\delta,e) + \psi(t,x-h_{MR}l_i,\delta,e)}{h_{MR}^2} =: \Delta_{x^i x^i}^{h_{MR}}\psi(t,x,\delta,e) \ .$$

If $a^{ij}(x,e) \geq 0$, $i \neq j$, then

$$\begin{aligned}\frac{\partial^2 \psi}{\partial x^i x^j}(t,x,\delta,e) \quad &\sim \quad \frac{2\psi(t,x,\delta,e) + \psi(t,x+h_{MR}(l_i+l_j),\delta,e) + \psi(t,x-h_{MR}(l_i+l_j),\delta,e)}{h_{MR}^2} \\ &\quad -\frac{\psi(t,x+h_{MR}l_i,\delta,e)+\psi(t,x-h_{MR}l_i,\delta,e)}{2h_{MR}^2} \\ &\quad -\frac{\psi(t,x+h_{MR}l_j,\delta,e)+\psi(t,x-h_{MR}l_j,\delta,e)}{2h_{MR}^2} \\ &=: \ \Delta_{x^i x^j}^{h_{MR}+}\psi(t,x,\delta,e) \ .\end{aligned}$$

If $a^{ij}(x,e) < 0$, $i \neq j$, then

$$\begin{aligned}\frac{\partial^2 \psi}{\partial x^i x^j}(t,x,\delta,e) \quad &\sim \quad -\frac{[2\psi(t,x,\delta,e) + \psi(t,x+h_{MR}(l_i-l_j),\delta,e) + \psi(t,x-h_{MR}(l_i-l_j),\delta,e)]}{h_{MR}^2} \\ &\quad +\frac{\psi(t,x+h_{MR}l_i,\delta,e)+\psi(t,x-h_{MR}l_i,\delta,e)}{2h_{MR}^2} \\ &\quad +\frac{\psi(t,x+h_{MR}l_j,\delta,e)+\psi(t,x-h_{MR}l_j,\delta,e)}{2h_{MR}^2} \\ &=: \ \Delta_{x^i x^j}^{h_{MR}-}\psi(t,x,\delta,e) \ .\end{aligned}$$

For ease of notations we write:

$$\Delta_x^{h_{MR}}\psi(t,x,\delta,e) := \left(\Delta_{x^i}^{h_{MR}+}\psi(t,x,\delta,e)\mathbf{1}_{b^i(x,e)\geq 0} + \Delta_{x^i}^{h_{MR}-}\psi(t,x,\delta,e)\mathbf{1}_{b^i(x,e)<0}\right)_{i\leq d} \in \mathbb{R}^d$$

and

$$\Delta_{xx}^{h_{MR}}\psi := \left(\Delta_{x^i x^j}^{h_{MR}+}\psi(t,x,\delta,e)\mathbf{1}_{a^{ij}(x,e)\geq 0} + \Delta_{x^i x^j}^{h_{MR}-}\psi(t,x,\delta,e)\mathbf{1}_{a^{ij}(x,e)<0}\right)_{i,j\leq d} \in \mathbb{M}^d \ .$$

Time component:

$$\frac{\partial \psi}{\partial t}(t,x,\delta,e) \sim \frac{\psi(t+h_N,x,\delta,e)-\psi(t,x,\delta,e)}{h_N} \quad =: \quad \Delta_t^{h_N}\psi(t,x,\delta,e) \ .$$

Latency duration:

$$\frac{\partial \psi}{\partial \delta}(t,x,\delta,e) \sim \frac{\psi(t,x,\delta,e)-\psi(t,x,\delta-h_N,e)}{h_N} \quad =: \quad \Delta_\delta^{h_N}\psi(t,x,\delta,e) \ .$$

5.3 Approximation scheme of (3.7) and convergence

We now define $\tilde V^R_{NM}$ as the solution on $\bar D^{RL}_{NM}$ of:

$$\mathcal{H}^{RL}_{NM}\varphi(t,x,\delta,e)\mathbf{1}_{x\notin\partial\mathcal{R}^d_R} + (\varphi(t,x,\delta,e)-g(x))\,\mathbf{1}_{x\in\partial\mathcal{R}^d_R} = 0$$

where $\mathcal{H}^{RL}_{NM}\varphi(t,x,\delta,e)$ is given by

$$(5.1)\begin{cases} \left(-\mathcal{L}^{R,e}_{NM} + \Delta^{h_N}_{\delta}\right)\varphi(t,x,\delta,e) & \text{on } \bar D^{RL}_{NM}\cap D_{E,>0}\ ,\\ \varphi(t,x,\delta,e)-\varphi(t+h_N,x,0,\varpi)-f(x,e) & \text{on } \bar D^{RL}_{NM}\cap D_{E,0}\ ,\\ \varphi(t,x,\delta,e)-\mathcal{V}(t,x,e) & \text{on } \bar D^{RL}_{NM}\cap D_{E,T}\ ,\\ \min\left\{-\mathcal{L}^{\varpi}_{NM}\varphi(t,x,\delta,e)\ ,\ \varphi(t,x,\delta,e)-\mathcal{M}^R_{LM}[\varphi](t,x)\right\} & \text{on } \bar D^{RL}_{NM}\cap D_{\varpi}\ ,\\ \varphi(t,x,\delta,e)-g(x)-f(x,e) & \text{on } \bar D^{RL}_{NM}\cap D_T\ ,\end{cases}$$

with

$$\begin{aligned}\mathcal{L}^{R,e}_{MN}\varphi(t,x,\delta,e) &:= \Delta^{h_N}_t\varphi(t,x,\delta,e) + \langle b(x,e)\ ,\ \Delta^{h_{MR}}_x\varphi(t,x,\delta,e)\rangle + \frac{1}{2}\mathrm{Tr}\left[aa^*(x,e)\Delta^{h_{MR}}_{xx}\varphi(t,x)\right]\ .\\ \mathcal{M}^R_{LM}[\varphi](t,x) &:= \max_{\{\delta',e'\}\in([\underline{\delta},T-t]\cap\{ih_N\}_{1\le i\le N})\times E_L}\varphi(t,\Pi^R_M(x+\beta(x,e',\delta')),\delta',e')\ .\end{aligned}$$

Here, Π^R_M denotes the projection operator on $\mathcal{R}^d_{MR}$

$$\Pi^R_M(x) := \left([(-R\vee x^i\wedge R)/M]M\right)_{i\le d}\ ,$$

with $[\cdot]$ denoting the integer part.
From now on, we write $\tilde V_n$ for $\tilde V^{R_nL_n}_{N_nM_n}$ where $(N_n,M_n,R_n,L_n)_{n\ge1}$ is a sequence of positive integers such that

$$N_n, M_n, R_n, L_n\uparrow\infty \quad\text{as } n\to\infty\ ,$$

and we denote by $\overline{V}$ and $\underline{V}$ the relaxed semilimits of $\tilde V_n$:

$$\begin{aligned}\overline{V}(t,x,\delta,e) &:= \limsup_{\substack{(i_hh_N,\,i_xh_{MR},\,i_\delta h_N,\,e')\,\in\,\bar D\,\to\,(t,x,\delta,e)\\ n\,\to\,\infty}} \tilde V_n(i_hh_N,i_xh_{MR},i_\delta h_N,e')\\ \underline{V}(t,x,\delta,e) &:= \liminf_{\substack{(i_hh_N,\,i_xh_{MR},\,i_\delta h_N,\,e')\,\in\,\bar D\,\to\,(t,x,\delta,e)\\ n\,\to\,\infty}} \tilde V_n(i_hh_N,i_xh_{MR},i_\delta h_N,e')\ .\end{aligned}$$

One easily checks that the above scheme is monotone, in the terminology of [4]. Moreover, recalling (2.7) and (2.8), easy computations based on a induction argument also lead to the following uniform polynomial control on $\overline{V}$ and $\underline{V}$ under the additional classical condition:

$$(5.2)\qquad h^2_{MR} = h_N\ .$$

Proposition 5.1. *The above scheme is monotone. If the condition (5.2), then there exists a constant $C>0$, independent on N,M,L and R, such that*

$$|\overline{V}(t,x,\delta,e)| + |\underline{V}(t,x,\delta,e)| \le C(1+|x|^\gamma) \quad \text{on } \bar D\ .$$

Using the fact that $f(\cdot,\varpi)=0$ and $\mathcal{V}(T,\cdot)=g+f$, recall (2.7) and Remark 2.6, we now observe that, if a function φ satisfies

$$\max\{\varphi-\mathcal{V}\ ,\ \varphi-\varphi(\cdot,0,\varpi)-f\ ,\ \varphi-g-f\}\,\mathbf{1}_{e\neq\varpi} + (\varphi-g-f)\,\mathbf{1}_{e=\varpi}\ge 0 \ \text{ on } D_T\ ,$$

then it also satisfies

$$\varphi - g - f \geq 0 \;\text{ on } D_T\;.$$

Similarly, if it satisfies

$$\min\left\{\varphi - \mathcal{V}\;,\; \varphi - \varphi(\cdot, 0, \varpi) - f\;,\; \varphi - g - f\right\} \mathbf{1}_{e\neq\varpi} + (\varphi - g - f)\,\mathbf{1}_{e=\varpi} \leq 0 \;\text{ on } D_T\;,$$

then it also satisfies

$$\varphi - g - f \leq 0 \;\text{ on } D_T\;.$$

It then follows from the arguments of [4], and the continuity of f, g and $\mathcal{V}$, see Remark 2.6, that $\underline{V}$ is a supersolution on $\bar{D}$ of $\mathcal{H}^*\varphi = 0$ and that $\overline{V}$ is a subsolution on $\bar{D}$ of $\mathcal{H}_*\varphi = 0$ where

$$\mathcal{H}^*\varphi := \begin{cases} H_{E,>0}\varphi := \left(-\mathcal{L}^e + \frac{\partial}{\partial\delta}\right)\varphi & \text{on} \quad D_{E,>0}\;, \\ \max\left\{H_{E,>0}\varphi\;,\; \varphi - \varphi(\cdot, 0, \varpi) - f\right\} & \text{on} \quad D_{E,0}\;, \\ \max\left\{H_{E,>0}\varphi\;,\; \varphi - \mathcal{V}\right\} & \text{on} \quad D_{E,T}\;, \\ H_\varpi\varphi := \max\left\{\,-\mathcal{L}^\varpi\varphi\;;\; \varphi\,-\,\mathcal{M}[\varphi]\right\} & \text{on} \quad D_\varpi\;, \\ \max\left\{H_{E,>0}\varphi\mathbf{1}_E + H_\varpi\varphi\mathbf{1}_{\{\varpi\}}\;,\; \varphi - g - f\right\} & \text{on} \quad D_T\;. \end{cases} \tag{5.3}$$

and

$$\mathcal{H}_*\varphi := \begin{cases} H_{E,>0}\varphi := \left(-\mathcal{L}^e + \frac{\partial}{\partial\delta}\right)\varphi & \text{on} \quad D_{E,>0}\;, \\ \min\left\{H_{E,>0}\varphi\;,\; \varphi - \varphi(\cdot, 0, \varpi) - f\right\} & \text{on} \quad D_{E,0}\;, \\ \min\left\{H_{E,>0}\varphi\;,\; \varphi - \mathcal{V}\right\} & \text{on} \quad D_{E,T}\;, \\ H_\varpi\varphi := \min\left\{\,-\mathcal{L}^\varpi\varphi\;;\; \varphi\,-\,\mathcal{M}[\varphi]\right\} & \text{on} \quad D_\varpi\;, \\ \min\left\{H_{E,>0}\varphi\mathbf{1}_E + H_\varpi\varphi\mathbf{1}_{\{\varpi\}}\;,\; \varphi - g - f\right\} & \text{on} \quad D_T\;. \end{cases} \tag{5.4}$$

In order to conclude that $\underline{V} = \overline{V} = V$ on $\bar{D}$, it remains to prove the following result.

Proposition 5.2. *Let ψ be lower-semicontinuous function with polynomial growth. If ψ is a viscosity supersolution (resp. subsolution) of $\mathcal{H}^*\varphi = 0$ (resp. $\mathcal{H}_*\varphi = 0$) on $\bar{D}$, then ψ is a viscosity supersolution (resp. subsolution) of $\mathcal{H}\varphi = 0$.*

Proof. We only prove the supersolution property, the subsolution property being proved by similar arguments. Let ψ be a supersolution of $\mathcal{H}^*\varphi = 0$. Let $(t_0, x_0, \delta_0, e_0) \in \bar{D}$ and let φ be a smooth function such that $(t_0, x_0, \delta_0, e_0)$ achieves a (global) minimum of $\psi - \varphi$ satisfying $(\psi - \varphi)(t_0, x_0, \delta_0, e_0) = 0$. If $(t_0, x_0, \delta_0, e_0) \in D_{E,>0} \cup D_\varpi$ then $\mathcal{H}\varphi(t_0, x_0, \delta_0, e_0) \geq 0$. If $(t_0, x_0, \delta_0, e_0) \in D_{E,0} \cup D_T$, then similar arguments as in the proof of Proposition 4.7 shows that $\mathcal{H}\varphi(t_0, x_0, \delta_0, e_0) \geq 0$ too.
It remains to study the case where $(t_0, x_0, \delta_0, e_0) \in D_{E,T}$. We claim that, the map $(t, x) \in [0, T) \times \mathbb{R}^d \to \psi(t, x, T - t, e_0)$ is a supersolution of

$$\max\left\{-\mathcal{L}^{e_0}\varphi\;,\; \varphi - \mathcal{V}\right\}(\cdot, e_0) \geq 0 \;\text{ on } [0, T) \times \mathbb{R}^d$$

with the terminal condition

$$\max\left\{\varphi - g - f, \varphi - \mathcal{V}\right\}(T, \cdot, e_0) \geq 0 \;\text{ on } \mathbb{R}^d\;.$$

Since $\mathcal{V}$ is a subsolution of the same equation, recall Remark 2.6, applying a standard comparison principle (recall our Lipschitz continuity and growth assumptions and see e.g. [8]), will readily implies that $\psi \geq \mathcal{V}$.

We conclude this proof by proving the above claim. Fix $e_0 \in E$, and let $(t_0, x_0) \in [0,T] \times \mathbb{R}^d$ and φ be smooth function such that (t_0, x_0) achieves a global minimum (equal to 0) of $(t,x) \mapsto \psi(t,x,T-t,e_0) - \varphi(t,x)$. For $n \geq 1$, we define φ_n by $\varphi_n(t,x,\delta,e) := \varphi(t,x) - n(T-t-\delta) - |t-t_0|^{2p} - |x-x_0|^{2p} - |e-e_0|^{2p}$, for $p \geq 1$ such that $(t,x,\delta,e) \in \bar{D} \mapsto |\psi(t,x,\delta,e)|/(1+|x|^p)$ is bounded. Let $(t_n, x_n, \delta_n, e_n)_n$ be a global minimum point of $\psi - \varphi_n$. Writing that

$$
\begin{aligned}
\psi(t_0, x_0, T-t_0, e_0) - \varphi(t_0, x_0) &\geq (\psi - \varphi_n)(t_n, x_n, \delta_n, e_n) \\
&= (\psi - \varphi)(t_n, x_n, \delta_n, e_n) \\
&+ n(T - t_n - \delta_n) + |t_n - t_0|^{2p} + |x_n - x_0|^{2p} + |e_n - e_0|^{2p} \\
&\geq (\psi - \varphi)(t_n, x_n, \delta_n, e_n)
\end{aligned}
$$

one easily checks that

(5.5)

$$(t_n, x_n, e_n) \to (t_0, x_0, e_0),\ n(T - t_n - \delta_n) \to 0 \text{ and } \psi(t_n, x_n, \delta_n, e_n) \to \psi(t_0, x_0, T - t_0, e_0)\ .$$

Note that, since $e_0 \in E$, we have $e_n \neq \varpi$ for n large enough. Moreover, the supersolution property of ψ implies that $\mathcal{H}^* \varphi_n(t_n, x_n, \delta_n, e_n) \geq 0$. Since $-\partial\varphi_n/\partial t + \partial\varphi_n/\partial\delta = -\partial\varphi/\partial t + 2p(t_n - t_0)^{2p-1}$, it follows from (5.5) that, for n large enough,

$$
(5.6) \qquad
\begin{array}{rll}
-\mathcal{L}^{e_n}\varphi(t_n, x_n) \geq \varepsilon_n & \text{if} & (t_n, x_n, \delta_n, e_n) \in D_{E,>0}\ , \\
-\mathcal{L}^{e_n}\varphi \vee (\varphi - \mathcal{V})\,(t_n, x_n, e_n) \geq \varepsilon_n & \text{if} & (t_n, x_n, \delta_n, e_n) \in D_{E,T}\ , \\
-\mathcal{L}^{e_n}\varphi \vee (\varphi - g - f)\,(t_n, x_n, e_n) \geq \varepsilon_n & \text{if} & (t_n, x_n, \delta_n, e_n) \in D_T\ ,
\end{array}
$$

where $\varepsilon_n \to 0$ as $n \to \infty$. Taking the limit as $n \to \infty$ and using (5.5) then implies that

$$\max\left\{-\mathcal{L}^{e_0}\varphi\ ,\ \varphi - \mathcal{V}\right\}(t_0, x_0, e_0) \geq 0 \ \text{ if } t_0 < T$$

and

$$\max\left\{-\mathcal{L}^{e_0}\varphi\ ,\ \varphi - g - f\ ,\ \varphi - \mathcal{V}\right\}(t_0, x_0, e_0) \geq 0 \ \text{ if } t_0 = T$$

which, by similar arguments as in the proof of Proposition 4.7, implies that

$$\max\left\{\varphi - g - f\ ,\ \varphi - \mathcal{V}\right\}(t_0, x_0, e_0) \geq 0 \ \text{ if } t_0 = T\ .$$

□

We can now conclude by using the comparison principle of Theorem 3.3, recall Proposition 5.2, Proposition 5.1, Theorem 3.2 and Remark 2.5.

Theorem 5.3. *We have:*

$$\underline{V} = \overline{V} = V \ \text{ on } \bar{D}\ .$$

6 Numerical illustration

As a numerical illustration, we consider the algorithm presented in Example 2.8 above. Clearly, it is too simplistic for practical purposes. Our aim is not to demonstrate its superiority with respect to other well-known algorithms, but only to show how the control adapts automatically to the market conditions each time a new slice is launched, on a simple case where its behavior is predictable.

We consider the following set of parameters. The trading period corresponds to a period of 3 hours. The price process is assumed to follow a Black and Scholes dynamics with zero drift

$S_t = S_0 e^{-\frac{1}{2}\sigma^2 t+\sigma W_t}$, where $S_0 := 13$ and the annualized volatility is 25%. Adding a drift would only change the optimal strategy in an obvious manner, depending on its sign. We assume a deterministic evolution of the instantaneous volume traded on the market $(V_t)_{t\le T}$ as given below. It corresponds to an intensity in minutes. The impact function η is given by $\eta(e,v) = 0.4(e/v)^{1.1}$. This coincides with plausible calibrated data. We take $\underline{\delta} = 5$ minutes. For this numerical test, we restrict to values of the latency time in the set $5, 10, \ldots, 60$ minutes. The different values of the buying rate are 50, 100, 150, ..., 500. It correspond to numbers of bought stocks per minute.
The final cost is given by $c(Q,v) = \eta(Q/(0.417v))$, which implicitly means that the trader has 25 seconds to finalize the operation, i.e. he must buy Q in 0.417 minutes at a rate $Q/0.417$. We consider two different types of functions ℓ: either ℓ is the identity, $\ell(r) = r$, or ℓ is of exponential type, $\ell(r) = e^{10^{-5}r} \wedge 100$. The value 100 corresponds to more than four times the cost evaluated with the exponential function $r \mapsto e^{10^{-5}r}$ of the operation which consists in buying 15000 stocks with a constant rate, assuming that the market volume takes the minimal value corresponding to the U-shaped path defined below, and for a constant stock price equal to 26. This is an extreme scenario. However, truncating the exponential function is needed in order to ensure that the value function is finite, since a log normal distribution does not admit exponential moments.

We first consider the case where ℓ is linear. In this case, the controller is risk neutral so that he has no incentive to buy the stock quickly because of a risk of increase of the price (recall that here the price is a martingale). In Figure 1 and Figure 2, we compare the case where the market volume is constant $V_t = 50000$, on the left, to the case where the market volume is strongly U-shaped: $V_t = 50000(1.1 - 0.9\sin(\pi t/T))$ with $T = 180$ minutes. This volume is also given per minutes. Both figures provide the optimal buying rate in terms of the remaining time $T - t$ and the remaining quantity to buy Q_t, for $S_t = S_0$. A typical path has to be read from north-west to south-east, since Q decreases as time goes buy. As expected, when the path of the market volume is U-shaped the optimal rate strongly decreases in the middle of the period, when the market volume is low and the impact on the price of the stock is high. This is compensated by a higher rate at the beginning of the period.

In Figure 5 and Figure 6, we provide the maximal value of the optimal latency time δ. Recall that the existence of multiple optima is possible since there is no additional cost related to the launching of a new slice, see Remark 2.1. In most cases, the maximal value is above the minimal threshold of 5 minutes. This support our choice of considering the latency time as a control, even in the absence of a cost associated to a change of parameters: when a latency time of, e.g., 15 minutes is optimal, the trader can let the algorithm run for this time period without having to launch it again every 5 minutes, and thus take care of his other positions. In the case where the market volume is strongly U-shaped, it is smaller at the beginning of the trading period, in comparison to the constant volume case. This is due to the fact that the algorithm knows that the market volume is going to decrease strongly (since it is deterministic) and that he will need to reduce the buying rate. It is small near the terminal time because of the constraint $t + \delta \le T$. When looking at the picture backward in time, i.e. as $T - t$ increases, we see that the maximal value first increases and then drops down very quickly. The first phenomena is due to the fact that the buying rate is essentially kept constant at its maximal value near the terminal time. Then, this rate decreases as the algorithm has more time to buy the shares. The period during which the algorithm reduces the buying rate naturally coincides with a lower latency time.

We next consider the case where ℓ is of exponential type. The optimal buying rates are reported in Figures 3-4. Because the controller is now risk adverse, he has an incentive to buy the stocks more quickly in order to avoid an increase of the price. This can be seen by comparing Figures 3-4 with Figures 1-2. However, we do not observe significant changes in the maximal optimal latency times.

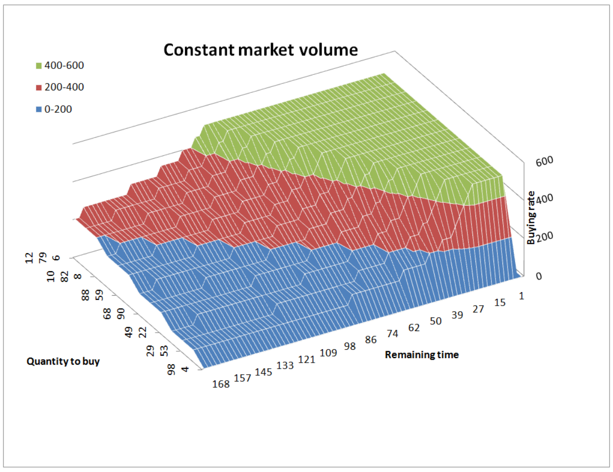


Figure 1: Buying rate - flat volume - Linear cost

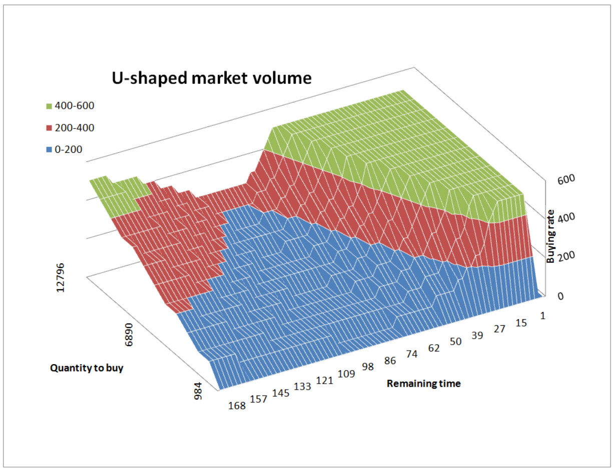


Figure 2: Buying rate - U-shaped volume - Linear cost

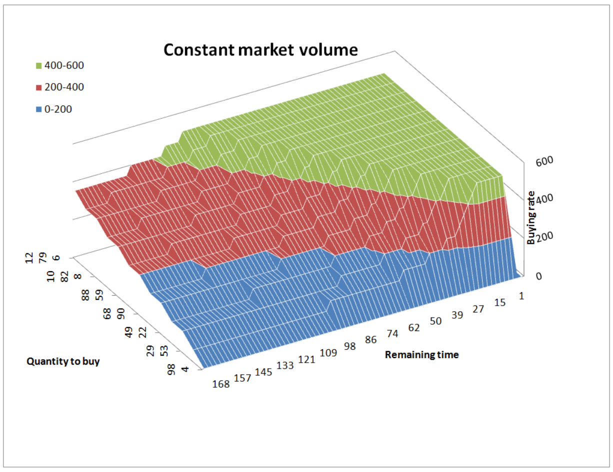


Figure 3: Buying rate - flat volume - Exponential cost

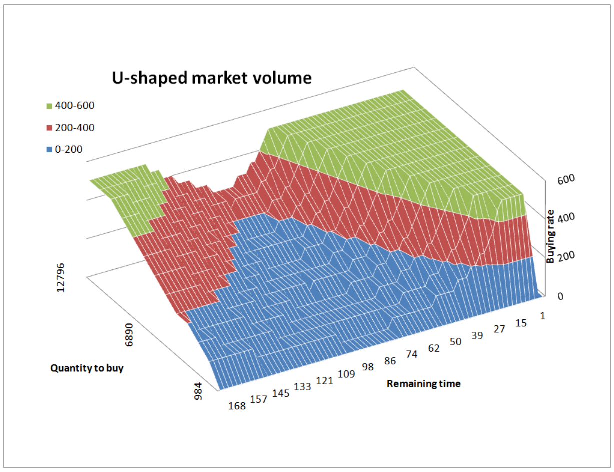


Figure 4: Buying rate - U-shaped volume - Exponential cost

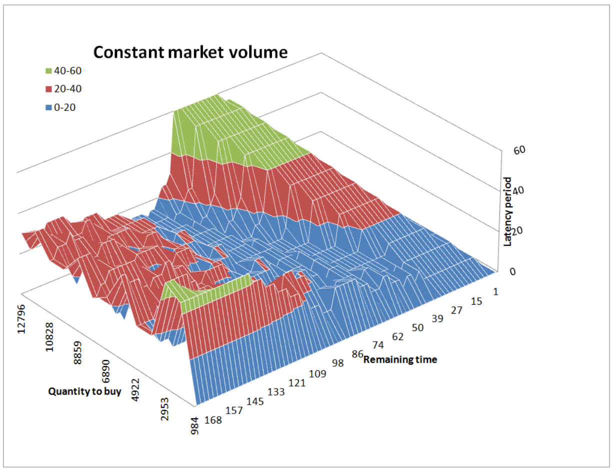


Figure 5: Latency period - flat volume - Linear cost

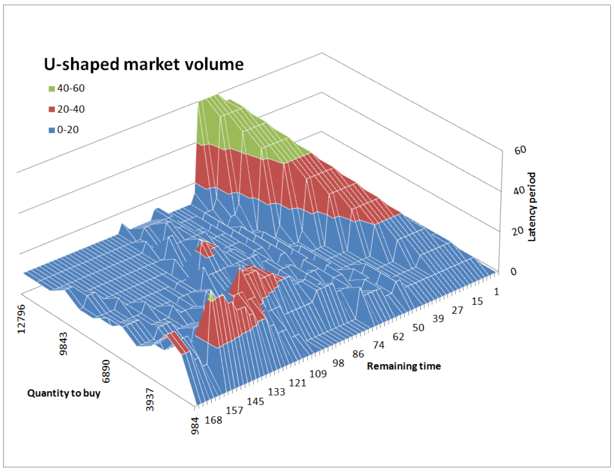


Figure 6: Latency period - U-shaped volume - Linear cost

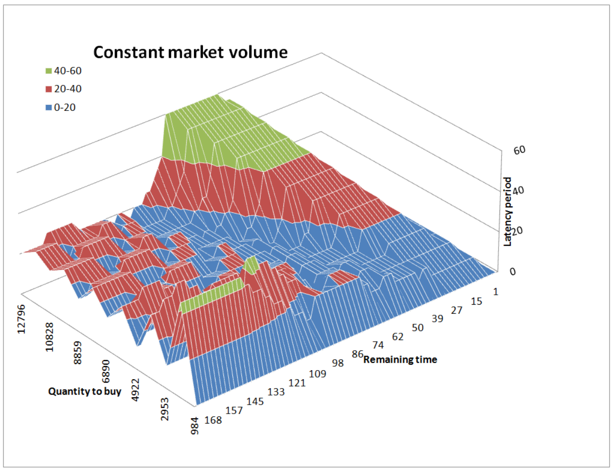


Figure 7: Latency period - flat volume - Exponential cost

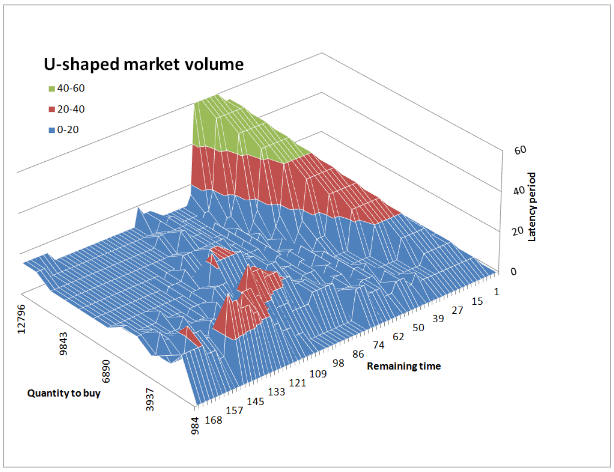


Figure 8: Latency period - U-shaped volume - Exponential cost

References

[1] R. Almgren. Optimal trading in a dynamic market. Technical Report 2, 2009.

[2] R. Almgren and B. Harts. A dynamic algorithm for smart order routing. Technical report, StreamBase, 2008.

[3] R. F. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3(2):5–39, 2000.

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

[5] D. Bertsimas, A. W. Lo, and P. Hummel. Optimal control of execution costs for portfolios. *Computing in Science and Engg.*, 1(6):40–53, 1999.

[6] A. Bensoussan and P.-L. Lions. *Impulse control and quasi-variational inequalities.* Gauthier-Villars Paris, 1984.

[7] B. Bouchard and N. Touzi. Weak dynamic programming principle for viscosity solutions. Technical report, CEREMADE, 2009.

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

[9] K. J. Falconer. *The Geometry of Fractal Sets (Cambridge Tracts in Mathematics).* Cambridge University Press, July 1986.

[10] B. J. Lapeyre, A. Sulem, and D. Talay. *Understanding Numerical Analysis for Option Pricing.* Cambridge University Press, February 2009.

[11] G. Pagès, S. Laruelle, and C.-A. Lehalle. Optimal split of orders across liquidity pools: a stochastic algorithm approach. Technical report, 2009.