



**HAL**  
open science

# A Stochastic Target Problem for Branching Diffusions

Idris Kharroubi, Antonio Ocello

► **To cite this version:**

| Idris Kharroubi, Antonio Ocello. A Stochastic Target Problem for Branching Diffusions. 2022. hal-03695036v2

**HAL Id: hal-03695036**

**<https://hal.science/hal-03695036v2>**

Preprint submitted on 16 Jun 2022 (v2), last revised 23 Jun 2022 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Stochastic Target Problem for Branching Diffusions

Idris Kharroubi\*

LPSM, UMR CNRS 8001,  
Sorbonne Université and Université Paris Cité,  
idris.kharroubi @ sorbonne-universite.fr

Antonio Ocello

LPSM, UMR CNRS 8001,  
Sorbonne Université and Université Paris Cité,  
antonio.ocello @ sorbonne-universite.fr

June 16, 2022

## Abstract

We consider an optimal stochastic target problem for branching diffusion processes. This problem consists in finding the minimal condition for which a control allows the underlying branching process to reach a target set at a finite terminal time for each of its branches. This problem is motivated by an example from fintech where we look for the super-replication price of options on blockchain based cryptocurrencies. We first state a dynamic programming principle for the value function of the stochastic target problem. We then show that the value function can be reduced to a new function with a finite dimensional argument by a so called branching property. Under wide conditions, this last function is shown to be the unique viscosity solution to an HJB variational inequality.

**MSC Classification-** 35K10, 49L20, 49L25, 60J80, 91G20

**Keywords**— Stochastic control, fintech, cryptocurrencies options, super-replication branching diffusion process, dynamic programming principle, Hamilton-Jacobi-Bellman equation, viscosity solution.

## 1 Introduction

The theory of optimal stochastic control has been extensively developed since the pioneering works in the 1950 decade. One reason for the growing attraction of this theory is the variety of its applications, such as physics, biology, economics or finance.

In the last field, stochastic control theory appears to be a very natural tool as it provides solutions to the optimal portfolio choice issue. The need to control the risks related to financial

---

\*Research of the author partially supported by ANR grant RELISCOPE.

investments leads to new kind of stochastic optimization problems. Here, one looks for the minimal initial endowment needed to find a financial strategy whose final position satisfies some given constraints. Such optimization problems are called optimal stochastic target problem and have been widely studied (see e.g. [24, 25, 2, 4, 3]).

The classical stochastic control theory has also been developed for other kind of stochastic processes such as branching diffusions. Those processes describe the evolution of a population of individuals with similar features concerning their dynamics and their reproduction. Branching processes have been first studied by Skorohod [23] and Ikeda et al. [13, 14, 15], who provided Feynmann-Kac presentation of solution to parabolic semi-linear PDEs. Since those pioneering works, branching processes have been extensively studied in particular their scaling limits and the link with superprocesses (see [8]). Recently, they were also used by Henry-Labordère et al. [12] for Monte Carlo based numerical approximation of solutions to semilinear parabolic PDEs.

In the case where the branching processes are controlled, Üstünel [27] considers a finite horizon optimization problem. He restricted to Markov controls acting only on the drift coefficient. Following a martingale problem approach, he proved existence of optimal controls under wide conditions. Nisio [21] considers the case where both the drift and diffusion coefficients are controlled. She characterizes the related value function as a viscosity solution to a nonlinear parabolic PDE of HJB type. Then, Claisse [5] extends the previous results by allowing controls that may not preserve independence of the particles and considering the lifespan and the progeny coefficients to depend on the position and the control. Following the approach of Fleming and Soner [11] which relies on a result due to Krylov [19], the value function is approximated by a sequence of smooth value functions corresponding to small perturbations of the initial problem. This allows to prove a dynamic programming principle and to derive a related dynamic programming equation.

In this paper, we investigate the stochastic target problem in the case where the underlying controlled process is a branching diffusion. The problem consists in finding a minimal initial condition for a given target branching diffusion such that it dominates a function of another controlled branching diffusion for each alive particle.

As the starting point of the target branching process can be seen as a measure and may contain several atoms, the previous problem is not well posed. We therefore give an extended equivalent formulation of the problem which consists in finding the minimal value for which one can find an initial measure condition whose atoms are dominated by this value and such that the related branching processes satisfy the terminal constraint.

Such a problem finds an application in mathematical finance, when dealing with the optimal investment on crypto-currencies. For these assets, branching may appear due to their structure, leading to new assets (see e.g. [10]). In this framework, the super-replication issue remains unsolved and our framework provides a possible solution. A detailed example is given in this article.

To characterize the value function of our branching stochastic target problem, we adopt a dynamic programming principle approach. Contrary to [5], our argument do not rely on the existence of regular solution to approximated PDEs. Instead, our proof of the dynamic programming principle relies on probabilistic arguments. We use a measurable selection theorem similar to that of [24] and combine it with a conditioning property for the law of the controlled process getting the dynamic programming principle.

We use the dynamic programming programming principle to characterize the value function as a solution to a dynamic programming partial differential equation. We first show, as in [5], a branching property on the value function. This branching property relates the value function at a given starting measure to the values taken on its atoms. This allows to see the value function as a sequence of function from  $[0, T] \times \mathbb{R}^d$  to  $\mathbb{R}$  indexed by the (countable) set  $\mathcal{I}$  of particle labels. Contrary to the classical branching property, ours writes the value function at the starting measure

as the maximum of values at the atoms of the measure. Hence, it entails irregularity in the value function brings us out of the range of regular solutions.

We therefore adopt the framework of viscosity solutions. Using the dynamic programming principle, the value function is characterized by a viscosity solution to a partial differential inequality of two terms. The first term is the classical nonlinear second order operator for classical diffusion processes, written as a supremum of a linear operator over controls that kill the diffusive part (see [4]). The restriction to those controls is due to the terminal constraint imposed with probability one. The second term expresses a monotonicity with respect to the label. More precisely, the value function taken at some label must be greater than its value on any other offspring label. Surprisingly, our PDE do not contain any polynomial of the value function as we classically have in PDEs related to branching processes. This is actually due to the particular structure of the control problem we consider. We complete this parabolic PDE property by a terminal condition.

To get a complete characterization of our value function, we finally consider the uniqueness to the PDE. Under an additional assumption on the asymptotic behaviour of the value function in the space variable and the label variable, we prove a comparison theorem using the classical approach of doubling variable and applying Ishii's lemma. This characterizes the value function as the unique viscosity solution to the PDE and shows its continuity on the parabolic interior of the domain.

The remainder of the paper is organized as follows. In Section 2, we present the branching stochastic target problem and provide an example of application inspired from fintech. In Section 3, we set the dynamic programming principle. We finally show in Section 4 viscosity properties of the value function and provide a uniqueness result to the related PDE. Finally we relegate some technical results needed in the proof of the conditioning property to the appendix.

## 2 The problem

### 2.1 Branching diffusions

We start by a description of the underlying controlled processes. As those processes are of branching type, we first introduce the label set.

**Label set** For  $n \geq 1$ , a multi-integer  $i = (i_1, \dots, i_n) \in \mathbb{N}^n$  is simply denoted by  $i = i_1 \dots i_n$ . For  $n, m \geq 1$  and two multi-integers  $i = i_1 \dots i_n \in \mathbb{N}^n$  and  $j = j_1 \dots j_m \in \mathbb{N}^m$ , we define their concatenation  $ij \in \mathbb{N}^{n+m}$  by

$$ij = i_1 \dots i_n j_1 \dots j_m . \tag{2.1}$$

To describe the evolution of the particle population, we introduce the set of labels  $\mathcal{I}$  defined by

$$\mathcal{I} = \{\emptyset\} \cup \bigcup_{n=1}^{+\infty} \mathbb{N}^n .$$

The label  $\emptyset$  corresponds to the mother particle. We extend the concatenation (2.1) to the whole set  $\mathcal{I}$  by

$$\emptyset i = i \emptyset = i$$

for all  $i \in \mathcal{I}$ . When the particle labelled  $i = i_1 \dots i_n \in \mathbb{N}^n$  gives birth to  $k$  particles, the off-springs are labelled  $i0, \dots, i(k-1)$ . We also define the partial ordering relation  $\preceq$  (resp.  $\prec$ ) by

$$\begin{aligned} j \preceq i &\Leftrightarrow \exists \ell \in \mathcal{I} : i = j\ell \\ (\text{resp. } j \prec i &\Leftrightarrow \exists \ell \in \mathcal{I} \setminus \{\emptyset\} : i = j\ell) \end{aligned}$$

for all  $i, j \in \mathcal{I}$ . We introduce the distance  $d_{\mathcal{I}}$  on  $\mathcal{I}$  defined by

$$d^{\mathcal{I}}(i, j) = \sum_{\ell=p+1}^n (i_{\ell} + 1) + \sum_{\ell'=p+1}^m (j_{\ell'} + 1),$$

for  $i = i_1 \dots i_n \in \mathbb{N}^n, j = j_1 \dots j_m \in \mathbb{N}^m$ , with

$$p = \max\{\ell \geq 1 : i_{\ell} = j_{\ell}\}.$$

We next write  $|i| := d^{\mathcal{I}}(i, \emptyset)$  for  $i \in \mathcal{I}$ .

**Set of finite measures** In the sequel we shall consider finite measure on  $\mathcal{I} \times \mathbb{R}^{\ell}$  for  $\ell \geq 1$ . For that, we endow the set  $\mathcal{I} \times \mathbb{R}^{\ell}$  with the metric  $d$  defined by

$$d((i, x), (j, y)) = d_{\mathcal{I}}(i, j) + |x - y|, \quad i, j \in \mathcal{I}, x, y \in \mathbb{R}^{\ell}.$$

$\mathcal{I} \times \mathbb{R}^{\ell}$  is then separable and complete. We denote by  $\mathcal{M}_F(\mathcal{I} \times \mathbb{R}^{\ell})$  set of the set of finite measures on  $\mathcal{I} \times \mathbb{R}^{\ell}$ . From Lemma 4.5 [17],  $\mathcal{M}_F(\mathcal{I} \times \mathbb{R}^{\ell})$  endowed with the topology of the weak convergence is Polish. We recall that we say that a sequence  $(\nu_n)_{n \geq 0}$  weakly converges to  $\nu$  in  $\mathcal{M}_F(\mathcal{I} \times \mathbb{R}^{\ell})$  if  $\int f d\nu_n \rightarrow \int f d\nu$  as  $n \rightarrow +\infty$  for any continuous and bounded function  $f$  from  $\mathcal{I} \times \mathbb{R}^{\ell}$  to  $\mathbb{R}$ . A possible metric associated to the weak topology on  $\mathcal{M}_F(\mathcal{I} \times \mathbb{R}^{\ell})$  is the Prokhorov metric (see Lemma 4.3 in [17]). We next define the subset  $E_{\ell}$  of  $\mathcal{M}_F(\mathcal{I} \times \mathbb{R}^{\ell})$  by

$$E_{\ell} = \left\{ \sum_{i \in V} \delta_{(i, x)} ; V \subseteq \mathcal{I}, V \text{ finite}, x^i \in \mathbb{R}^{\ell} \text{ and } i \not\sim j \text{ for } i, j \in V \right\}. \quad (2.2)$$

By Proposition A.6,  $E_d$  is Polish as well.

**Probabilistic setting** We fix a deterministic terminal time  $T > 0$  and a filtered probability space  $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$  satisfying the usual conditions. Suppose that this probability space is endowed with a family of processes  $(B^i, Q^i)_{i \in \mathcal{I}}$  such that

- $(B_t^i)_{t \in [0, T]}$  is an  $\mathbb{F}$ -standard Brownian motion in  $\mathbb{R}^m$  for all  $i \in \mathcal{I}$ ;
- $Q^i(dt, dk)$  is an  $\mathbb{F}$ -Poisson random measure on  $[0, T] \times \mathbb{N}$  with intensity measure  $dt \gamma p_k \delta_k$  for all  $i \in \mathcal{I}$ , with  $\gamma > 0, p_k \geq 0$  for  $k \geq 0$  and  $\sum_{k \geq 0} p_k = 1, \delta_k$  being the Dirac measure at  $k$ ;
- $\{B^i, Q^j, i, j \in \mathcal{I}\}$  forms a family of mutually independent processes.

Having in mind these processes, we precise a better probability space.

- Let  $\Omega^0$  be the space of continuous functions from  $[0, T]$  that are  $\mathbb{R}^m$ -valued starting at 0. Let  $\mathbb{F}^0 := (\mathcal{F}_t^0)_{t \in [0, T]}$  the filtration generated by the canonical process  $B(\omega^0) := \omega^0, \omega^0 \in \Omega^0$ . We endow  $(\Omega^0, \mathcal{F}_T^0)$  with the Wiener measure  $\mathbb{P}^0$ .

- Let  $\Omega^1$  be the set of measures  $\omega^1$  on  $\mathbb{R}_+ \times \mathbb{N}$  of the form  $\omega^1 = \sum_{k \geq 0} \delta_{(t_k, n_k)}$ . Let  $\mathbb{F}^1 := (\mathcal{F}_t^1)_{t \in [0, T]}$  be the filtration generated by the canonical process  $Q(\omega^1) = \omega^1$ :

$$\mathcal{F}_t^1 := \sigma(Q([0, s] \times \{k\}) : s \in [0, t], k \in \mathbb{N}), \quad t \in [0, T].$$

We endow  $(\Omega^1, \mathcal{F}_T^1)$  with the Poisson measure  $\mathbb{P}^1$  of intensity  $dt \gamma \sum_{k \geq 0} p_k \delta_k$ , that is the probability measure such that  $Q$  is a Poisson point process with intensity  $dt \gamma \sum_{k \geq 0} p_k \delta_k$ .

Following the structure we expect for  $\{B^i, Q^j, i, j \in \mathcal{I}\}$ , we define the filtered space  $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ , where  $\Omega = (\Omega^0 \times \Omega^1)^{\mathcal{I}}$ ,  $\mathbb{P} = (\mathbb{P}^0 \otimes \mathbb{P}^1)^{\otimes \mathcal{I}}$ ,  $\mathcal{F}$  is the  $\mathbb{P}$ -augmentation of  $(\mathcal{F}_T^0 \otimes \mathcal{F}_T^1)^{\otimes \mathcal{I}}$  and  $\mathbb{F} = (\mathcal{F}_t)_{t \in [0, T]}$  is the  $\mathbb{P}$ -augmentation of the filtration  $((\mathcal{F}_t^0 \otimes \mathcal{F}_t^1)^{\otimes \mathcal{I}})_{t \in [0, T]}$ . On this space we extend the definition of the processes  $B^i$  and  $Q^i$  for  $i \in \mathcal{I}$  as the previously described processes  $B$  and  $Q$  composed with the projections on each component, *i.e.*

$$B^i(\omega) := \omega^{0,i}, \quad Q^j(\omega) := \omega^{1,j}, \quad \omega = (\omega^{0,i}, \omega^{1,i})_{i \in \mathcal{I}} \in \Omega.$$

We also define the process  $\xi$  valued in  $\mathcal{M}_F(\mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$  by

$$\xi_t = \sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \delta_{(i, n, B_t^i, Q^i([0, t] \times \{n\}))} \quad (2.3)$$

for  $t \in [0, T]$ . We then notice that the filtration  $\mathbb{F}$  is the completed filtration generated by the process  $\xi$ .

To stress the dependence in time, we will use the following notations. For  $t \in [0, T]$  and  $\omega = (\omega^0, \omega^1) \in \Omega$ , we define the stopped path at time  $t$  by  $\omega_{\cdot \wedge t} = (\omega_{\cdot \wedge t}^0, \omega_{\cdot \wedge t}^1)$  where

$$\omega_{\cdot \wedge t}^0 = (\omega_{s \wedge t}^0)_{s \geq 0} \quad \text{and} \quad \omega_{\cdot \wedge t}^1 = \omega^1(\cdot \cap [0, t] \times \mathbb{N}).$$

For a process  $(X_t)_{t \in [0, T]}$  and a random time  $\tau : \Omega \rightarrow [0, T]$ , we denote by  $(X_{t \wedge \tau})_{t \in [0, T]}$  the process defined by

$$X_{t \wedge \tau}(\omega) = X_t(\omega_{\cdot \wedge \tau(\omega)}), \quad t \in [0, T], \omega \in \Omega.$$

For  $\omega, \tilde{\omega} \in \Omega$  and a random time  $\tau : \Omega \rightarrow [0, T]$ , we define the concatenation path  $\omega \oplus_{\tau} \tilde{\omega} = (\omega^{0,i} \oplus_{\tau} \tilde{\omega}^{0,i}, \omega^{1,i} \oplus_{\tau} \tilde{\omega}^{1,i})_{i \in \mathcal{I}}$  by

$$(\omega^{0,i} \oplus_{\tau} \tilde{\omega}^{0,i})_s = \omega_s^{0,i} \mathbf{1}_{s < \tau(\omega)} + (\tilde{\omega}_s^{0,i} - \tilde{\omega}_{\tau(\omega)}^{0,i} + \omega_{\tau(\omega)}^{0,i}) \mathbf{1}_{s \geq \tau(\omega)}, \quad s \in [0, T],$$

and

$$\omega^{1,i} \oplus_{\tau(\omega)} \tilde{\omega}^{1,i} = \omega^{1,i}(\cdot \cap [0, \tau(\omega)] \times \mathbb{N}) + \tilde{\omega}^{1,i}(\cdot \cap (\tau(\omega), T] \times \mathbb{N}).$$

for  $i \in \mathcal{I}$ . For a random variable  $S$  valued in some Polish space, we also define the shifted random variable  $S^{\tau, \omega}$  by

$$S^{\tau, \omega}(\tilde{\omega}) = S(\omega \oplus_{\tau} \tilde{\omega}), \quad \tilde{\omega} \in \Omega. \quad (2.4)$$

**Alive particles** We define the set  $\mathcal{V}_t$  of alive particles at time  $t$  as follows.

- At time  $t = 0$ , the set is reduced to the mother particle :  $\mathcal{V}_0 = \{\emptyset\}$ .
- For a time  $t \geq 0$ , a particle  $i \in \mathcal{V}_t$  dies at the first time  $\tau_i > t$  the related Poisson measure  $Q^i$  jumps after  $t$ , *i.e.*

$$\tau_i = \inf\{s > t : Q^i((t, s] \times \mathbb{N}) = 1\}.$$

- At time  $\tau_i$ , this particle gives birth to  $k$  particles  $i0, \dots, i(k-1)$ , with  $k$  such that  $Q^i(\{\tau_i\} \times \{k\}) = 1$ :

$$\mathcal{V}_{\tau_i} = (\mathcal{V}_{\tau_i-} \setminus \{i\}) \cup \{i0, \dots, i(k-1)\}.$$

**Controlled population** Take  $A$  a Polish space with metric  $d_A$ . We assume  $d_A$  to be bounded (if not so, we replace  $d_A$  by  $d_A \wedge 1$  and still have a Polish space). We define a control  $\alpha$  as a family  $(\alpha^i)_{i \in \mathcal{I}}$  of  $\mathbb{F}$ -progressively measurable processes valued in  $A$ . We denote by  $\mathcal{A}$  the set of such controls.

Let  $\lambda : \mathbb{R}^d \times A \rightarrow \mathbb{R}^d$  and  $\sigma : \mathbb{R}^d \times A \rightarrow \mathbb{R}^{d \times m}$  be measurable functions. For a given control  $\alpha \in \mathcal{A}$ , each particle  $i \in \mathcal{I}$  of the controlled population is born, evolves and dies to give birth to off-springs according to the set  $\mathcal{V}$  defined above. We denote by  $X_s^i$  the position at time  $s$  of a particle  $i \in \mathcal{V}_s$ . For  $i \in \mathcal{I}$  alive at time  $t$ , let  $\tau_i \geq t$  be the random time of its death, giving birth to  $k$  off-springs. The position at a time  $s \geq \tau_i$  of the off-springs  $i_0, \dots, i_{k-1}$  are given by

$$X_{\tau_i}^{il} = X_{\tau_i}^i \quad (2.5)$$

$$dX_s^{il} = \lambda(X_s^{il}, \alpha_s^{il}) ds + \sigma(X_s^{il}, \alpha_s^{il}) dB_s^{il}, \quad (2.6)$$

for  $\ell = 0, \dots, k-1$ , such that  $i\ell$  is alive at time  $s$ . We represent the population of alive particles by the following measure valued process

$$Z_s = \sum_{i \in \mathcal{V}_s} \delta_{(i, X_s^i)}, \quad s \geq 0.$$

The process  $Z$  takes values in the Polish space  $E_d$  defined by (2.2).

For a function  $f : \mathcal{I} \times \mathbb{R}^d \rightarrow \mathbb{R}$ , and a measure  $\mu = \sum_{i \in \mathcal{V}} \delta_{(i, x_i)} \in E_d$ , we set

$$f(\mu) = \int_{\mathcal{I} \times \mathbb{R}^d} f d\mu = \sum_{i \in \mathcal{V}} f_i(x_i).$$

We introduce the second order local operators  $L^a$ ,  $a \in A$  defined by

$$L^a \varphi(x) = \lambda(x, a)^\top D\varphi(x) + \frac{1}{2} \text{Tr} \left( \sigma \sigma^\top(x, a) D^2 \varphi(x) \right), \quad x \in \mathbb{R}^d,$$

for  $\varphi \in C^2(\mathbb{R}^d)$ , where  $D\varphi$  and  $D^2\varphi$  denote respectively, the gradient and the Hessian matrix of  $\varphi$ .

For a control  $\alpha \in \mathcal{A}$  and a function  $f : [0, T] \times \mathcal{I} \times \mathbb{R}^d \rightarrow \mathbb{R}$  such that  $f_i(\cdot) \in C^{1,2}([0, T] \times \mathbb{R}^d)$  for all  $i \in \mathcal{I}$ , the following SDE characterises the behaviour of  $Z$ :

$$\begin{aligned} f(t, Z_t) &= f(s, Z_s) + \int_s^t \sum_{i \in \mathcal{V}_u} Df_i(u, X_u^i)^\top \sigma(X_u^i, \alpha_u^i) dB_u^i \\ &\quad + \int_s^t \sum_{i \in \mathcal{V}_u} (\partial_t + L^{\alpha_u^i}) f_i(u, X_u^i) du \\ &\quad + \int_{(s,t] \times \mathbb{N}} \sum_{i \in \mathcal{V}_{u-}} \sum_{k \geq 0} \left( \sum_{\ell=0}^{k-1} f_{i\ell} - f_i \right) (u, X_u^i) Q^i(du dk) \end{aligned} \quad (2.7)$$

for all  $s, t \in [0, T]$  such that  $s \leq t$ .

**Target branching diffusion** To each alive particle  $i \in \mathcal{V}_s$ , we associate a target position at time  $s$  denoted by  $Y_s^i$ . Let  $\lambda_Y : \mathbb{R}^d \times \mathbb{R} \times A \rightarrow \mathbb{R}$  and  $\sigma_Y : \mathbb{R}^d \times A \rightarrow \mathbb{R}^{1 \times m}$  be measurable functions. Let  $\tau_i \geq t$  be the random time of death of  $i \in \mathcal{I}$ , the target position at time  $s \geq \tau_i$  is given by

$$Y_{\tau_i}^{il} = Y_{\tau_i}^i \quad (2.8)$$

$$dY_s^{il} = \lambda_Y(X_s^{il, \alpha}, Y_s^{il, \alpha}, \alpha_s^{il}) ds + \sigma_Y(X_s^{il, \alpha}, \alpha_s^{il}) dB_s^{il}, \quad (2.9)$$

for  $\ell = 0, \dots, k-1$ , such that particle  $i\ell$  is alive at time  $s$ .

We use the notation  $\hat{\cdot}$  to define the quantities associated to the pair  $\begin{pmatrix} X_s^i \\ Y_s^i \end{pmatrix}$ , considering the previous problem but on  $\mathbb{R}^{d+1}$ . Therefore, we have  $\hat{X}_s^i := \begin{pmatrix} X_s^i \\ Y_s^i \end{pmatrix}$ ,  $\hat{\lambda}(\hat{X}_s^i, \alpha_s^i) := \begin{pmatrix} \lambda(X_s^i, \alpha_s^i) \\ \lambda_Y(X_s^i, Y_s^i, \alpha_s^i) \end{pmatrix}$  and  $\hat{\sigma}(\hat{X}_s^i, \alpha_s^i) := \begin{pmatrix} \sigma(X_s^i, \alpha_s^i) \\ \sigma_Y(X_s^i, \alpha_s^i) \end{pmatrix}$ . Under those hypotheses, assuming  $i$  is alive, its position  $\hat{X}^i$  evolves according to

$$d\hat{X}_s^i = \hat{\lambda}(\hat{X}_s^i, \alpha_s^i)ds + \hat{\sigma}(\hat{X}_s^i, \alpha_s^i) dB_s^i. \quad (2.10)$$

The resulting population process valued in  $E_{d+1}$  is

$$\hat{Z}_t = \sum_{i \in \mathcal{V}_s} \delta_{(i, X_s^i, Y_s^i)}, \quad s \geq 0.$$

As before, we define the related second order local operators  $\hat{L}^a$ ,  $a \in A$  by

$$\hat{L}^a \hat{\varphi}(\hat{x}) = \hat{\lambda}(\hat{x}, a)^\top D \hat{\varphi}(\hat{x}) + \frac{1}{2} \text{Tr}(\hat{\sigma} \hat{\sigma}^\top(\hat{x}, a) D^2 \hat{\varphi}(\hat{x})), \quad \hat{x} \in \mathbb{R}^{d+1},$$

for  $\hat{\varphi} \in C^2(\mathbb{R}^{d+1})$ , where  $D \hat{\varphi}$  and  $D^2 \hat{\varphi}$  denote respectively, the gradient and the Hessian matrix of  $\hat{\varphi}$ .

For a control  $\alpha \in \mathcal{A}$  and a function  $\hat{f} : [0, T] \times \mathcal{I} \times \mathbb{R}^{d+1} \rightarrow \mathbb{R}$  such that  $\hat{f}_i(\cdot) \in C^{1,2}([0, T] \times \mathbb{R}^{d+1})$  for all  $i \in \mathcal{I}$ , the SDE related to  $\hat{Z}$  takes the following form:

$$\begin{aligned} \hat{f}(t, \hat{Z}_t) &= \hat{f}(s, \hat{Z}_s) + \int_s^t \sum_{i \in \mathcal{V}_u} D \hat{f}_i(u, \hat{X}_u^i)^\top \hat{\sigma}(\hat{X}_u^i, \alpha_u^i) dB_u^i \\ &+ \int_s^t \sum_{i \in \mathcal{V}_u} (\partial_t + \hat{L}^{\alpha_u^i}) \hat{f}_i(u, \hat{X}_u^i) du \\ &+ \int_{(s,t] \times \mathbb{N}} \sum_{i \in \mathcal{V}_{u-}} \sum_{k \geq 0} \left( \sum_{\ell=0}^{k-1} \hat{f}_{i\ell} - \hat{f}_i \right) (u, \hat{X}_u^i) Q^i(du dk) \end{aligned} \quad (2.11)$$

for all  $s, t \in [0, T]$  such that  $s \leq t$ .

**Well posedness** To ensure the well definition of the presented controlled processes, we make the following assumption.

**Assumption A1.** (i) The coefficients  $p_k$ ,  $k \geq 0$ , satisfy

$$\sum_{k \geq 0} k p_k = M < +\infty.$$

(ii) The functions  $\lambda$ ,  $\sigma$ ,  $\lambda_Y$  and  $\sigma_Y$  satisfy

$$\sup_{a \in A} |\lambda(0, a)| + |\sigma(0, a)| + |\lambda_Y(0, 0, a)| + |\sigma_Y(0, a)| < +\infty.$$

(iii) There exists a constant  $L > 0$  such that

$$\begin{aligned} &|\lambda(x, a) - \lambda(x', a)| + |\sigma(x, a) - \sigma(x', a)| \\ &+ |\lambda_Y(x, y, a) - \lambda_Y(x', y', a')| + |\sigma_Y(x, a) - \sigma_Y(x', a)| \leq L (|x - x'| + |y - y'|) \end{aligned}$$

for all  $x, x' \in \mathbb{R}^d$ ,  $y, y' \in \mathbb{R}$  and  $a \in A$ .



(iv) There exists a nondecreasing function  $w : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that  $w(r) \xrightarrow{r \rightarrow 0} 0$  and

$$\begin{aligned} & |\lambda(x, a) - \lambda(x, a')| + |\sigma(x, a) - \sigma(x, a')| \\ & + |\lambda_Y(x, y, a) - \lambda_Y(x, y, a')| + |\sigma_Y(x, a) - \sigma_Y(x, a')| \leq w(d_A(a, a')) \end{aligned}$$

for all  $x \in \mathbb{R}^d$ ,  $y \in \mathbb{R}$  and  $a, a' \in A$ .

For any initial condition  $t \in [0, T]$ ,  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$  and  $y_i \in \mathbb{R}$  for  $i \in V$ , we extend the controlled branching processes  $(X, Y)$ . For that the set of alive particles  $\mathcal{V}^{t, \mu}$  is defined as follows.

- For  $s \in [0, t]$ ,  $\mathcal{V}_s^{t, \mu} = V$ .
- For  $s \geq t$ , a particle  $i \in \mathcal{V}_s$  dies at the first time  $\tau_i > s$  the related Poisson measure  $Q^i$  jumps after  $s$ :

$$\tau_i = \inf\{r > s : Q^i([s, r] \times \mathbb{N}) = 1\}.$$

- At time  $\tau_i$ , the particle  $i$  gives birth to  $k$  particles  $i0, \dots, i(k-1)$ , with  $k$  such that  $Q^i(\{\tau_i\} \times \{k\}) = 1$ :

$$\mathcal{V}_{\tau_i}^{t, \mu} = \left( \mathcal{V}_{\tau_i^-}^{t, \mu} \setminus \{i\} \right) \cup \{i0, \dots, i(k-1)\}.$$

Then, the controlled branching population process  $X^{t, \mu, \alpha} = (X_s^{t, \mu, \alpha, i}, i \in \mathcal{V}_s^{t, \mu})_{s \in [0, T]}$  is defined by the initial condition

$$X_s^{t, \mu, \alpha} = (x_i, i \in V), \quad s \in [0, t],$$

together with dynamics (2.5)-(2.6). We also denote by  $\hat{\mu} \in E_{d+1}$  the extended measure as

$$\hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)},$$

and  $Y^{t, \hat{\mu}, \alpha} = (Y_s^{t, \hat{\mu}, \alpha, i}, i \in \mathcal{V}_s^{t, \mu})_{s \in [0, T]}$  the controlled branching target process with initial condition

$$Y_s^{t, \hat{\mu}, \alpha, i} = y_i, \quad s \in [0, t],$$

for all  $i \in V$ , together with dynamics (2.8)-(2.9). Let  $Z^{t, \mu, \alpha}$  and  $\hat{Z}^{t, \hat{\mu}, \alpha}$  be

$$Z_s^{t, \mu, \alpha} = \sum_{i \in \mathcal{V}_s^{t, \mu}} \delta_{(i, X_s^{t, \mu, \alpha, i})} \quad \text{and} \quad \hat{Z}_s^{t, \hat{\mu}, \alpha} = \sum_{i \in \mathcal{V}_s^{t, \mu}} \delta_{(i, X_s^{t, \mu, \alpha, i}, Y_s^{t, \hat{\mu}, \alpha, i})}$$

for  $s \in [0, T]$ .

In this setting, we have the following non-explosion result.

**Proposition 2.1.** *Suppose that Assumptions A1 (i)-(ii)-(iii) hold. Fix  $t \in [0, T]$ ,  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$ ,  $\hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)} \in E_{d+1}$  and  $\alpha \in \mathcal{A}$ .*

(i) *The set of alive particles  $\mathcal{V}_s^{t, \mu}$  is uniquely defined and is finite for all  $s \in [0, T]$ . More precisely, we have*

$$\mathbb{E} \left[ \sup_{s \in [0, T]} |\mathcal{V}_s^{t, \mu}| \right] \leq |V| e^{\gamma M(T-t)}$$

where  $|\mathcal{V}|$  stands for the cardinal of a subset  $\mathcal{V}$  of  $\mathcal{I}$ .

(ii) *There exists a unique  $\mathbb{F}$ -adapted process  $(Z^{t, \mu, \alpha})$  (resp.  $(\hat{Z}^{t, \hat{\mu}, \alpha})$ ) valued in  $E_d$  (resp.  $E_{d+1}$ ). Moreover, the process  $Z^{t, \mu, \alpha}$  (resp.  $(\hat{Z}^{t, \hat{\mu}, \alpha})$ ) satisfies (2.7) (resp. (2.11)).*

We refer to [5, Proposition 2.1] for the proof of this proposition.

**Remark 2.1.** For any  $i \in \mathcal{I}$  the processes  $X^{t,\mu,\alpha,i}$  and  $Y^{t,\hat{\mu},\alpha,i}$  are defined on times  $s \in [t, T]$  such that  $i \in \mathcal{V}_s^{t,\mu}$ . However, we can extend their definition to the whole interval  $[t, T]$ . Suppose first that  $i$  has no ancestor in  $\mathcal{V}_t^{t,\mu}$ :

$$j \not\leq i \text{ for all } j \in \mathcal{V}_t^{t,\mu}.$$

Then we define processes  $X^{t,\mu,\alpha,i}$  and  $Y^{t,\hat{\mu},\alpha,i}$  as the unique solutions to

$$\begin{aligned} dX_s^{t,\mu,\alpha,i} &= \lambda(X_s^{t,\mu,\alpha,i}, \alpha_s^i) ds + \sigma(X_s^{t,\mu,\alpha,i}, \alpha_s^i) dB_s^i \\ dY_s^{t,\hat{\mu},\alpha,i} &= \lambda_Y(X_s^{t,\mu,\alpha,i}, Y_s^{t,\hat{\mu},\alpha,i}, \alpha_s^i) ds + \sigma_Y(X_s^{t,\mu,\alpha,i}, \alpha_s^i) dB_s^i \end{aligned}$$

for  $s \in [t, T]$ , with initial condition  $X_t^{t,\mu,\alpha,i} = 0$  and  $Y_t^{t,\hat{\mu},\alpha,i} = 0$ . On the complementary case, it exists  $j \in \mathcal{V}_t^{t,\mu}$  such that  $j \leq i$ . Then there exists  $k \geq 1$  and  $\ell_1, \dots, \ell_k$  such that

$$i = j\ell_1 \dots \ell_k.$$

We denote the associated branching times by  $(S_0, \dots, S_k)$ :

$$S_m = \inf \left\{ s > S_{m-1} : Q^{j\ell_1 \dots \ell_m}((S_{m-1}, s] \times \{n_m\}) = 1 \right\}$$

where  $n_m \geq \ell_{m+1} + 1$  for  $m = 0, \dots, k$  with  $S_{-1} = t$ . Then we define the extended processes  $X^{t,\mu,\alpha,i}$  and  $Y^{t,\hat{\mu},\alpha,i}$  by

$$\begin{aligned} X_s^{t,\mu,\alpha,i} &= \mathbb{1}_{[t, S_0)}(s) X_s^{t,\mu,\alpha,j} + \sum_{m=1}^{k-1} \mathbb{1}_{[S_{m-1}, S_m)}(s) X_s^{t,\mu,\alpha, j\ell_1 \dots \ell_m} + \mathbb{1}_{[S_{k-1}, +\infty)}(s) X_s^{t,\mu,\alpha,i} \\ Y_s^{t,\hat{\mu},\alpha,i} &= \mathbb{1}_{[t, S_0)}(s) Y_s^{t,\hat{\mu},\alpha,j} + \sum_{m=1}^{k-1} \mathbb{1}_{[S_{m-1}, S_m)}(s) Y_s^{t,\hat{\mu},\alpha, j\ell_1 \dots \ell_m} + \mathbb{1}_{[S_{k-1}, +\infty)}(s) Y_s^{t,\hat{\mu},\alpha,i} \end{aligned}$$

for  $s \in [t, T]$ .

These extended processes can be seen as solution to a Brownian stochastic differential equation with Lipschitz coefficients. Obvious in the first case, to show it in the second one, we consider the ancestor Brownian motion  $\bar{B}^i$  defined by

$$\begin{aligned} \bar{B}_s^i &= B_s^j \mathbb{1}_{[t, S_0)} + \sum_{m=1}^{k-1} \mathbb{1}_{[S_{m-1}, S_m)}(s) \left( B_s^{j\ell_1 \dots \ell_m} - B_{S_{m-1}}^{j\ell_1 \dots \ell_m} + B_{S_{m-1}}^{j\ell_1 \dots \ell_{m-1}} \right) \\ &\quad + \mathbb{1}_{[S_{k-1}, +\infty)}(s) \left( B_s^i - B_{S_{k-1}}^i + B_{S_{k-1}}^{j\ell_1 \dots \ell_{k-1}} \right), \end{aligned}$$

for  $s \in [t, T]$ . This process is continuous, centered, with independent increments and variance equal to  $t$ , therefore a Brownian motion by Lévy's characterisation. Then the extended processes  $X^{t,\mu,\alpha,i}$  and  $Y^{t,\hat{\mu},\alpha,i}$  are the unique solutions to the SDE

$$dX_s^{t,\mu,\alpha,i} = \bar{\lambda}(s, X_s^{t,\mu,\alpha,i}) ds + \bar{\sigma}(s, X_s^{t,\mu,\alpha,i}) d\bar{B}_s^i \quad (2.12)$$

$$dY_s^{t,\hat{\mu},\alpha,i} = \bar{\lambda}_Y(s, X_s^{t,\mu,\alpha,i}, Y_s^{t,\hat{\mu},\alpha,i}) ds + \bar{\sigma}_Y(s, X_s^{t,\mu,\alpha,i}) d\bar{B}_s^i \quad (2.13)$$

for  $s \in [t, T]$ , with initial condition  $X_t^{t, \mu, \alpha, i} = x_i$  and  $Y_t^{t, \mu, \alpha, i} = y_i$ . The coefficients being given by

$$\begin{aligned}\bar{\lambda}(s, x) &= \mathbb{1}_{[t, S_0)} \lambda(x, \alpha_s^j) + \sum_{m=1}^{k-1} \mathbb{1}_{[S_{m-1}, S_m)}(s) \lambda(x, \alpha_s^{j \ell_1 \dots \ell_m}) + \mathbb{1}_{[S_{k-1}, +\infty)}(s) \lambda(x, \alpha_s^i) \\ \bar{\sigma}(s, x) &= \mathbb{1}_{[t, S_0)} \sigma(x, \alpha_s^j) + \sum_{m=1}^{k-1} \mathbb{1}_{[S_{m-1}, S_m)}(s) \sigma(x, \alpha_s^{j \ell_1 \dots \ell_m}) + \mathbb{1}_{[S_{k-1}, +\infty)}(s) \sigma(x, \alpha_s^i) \\ \bar{\lambda}_Y(s, x, y) &= \mathbb{1}_{[t, S_0)} \lambda_Y(x, y, \alpha_s^j) + \sum_{m=1}^{k-1} \mathbb{1}_{[S_{m-1}, S_m)}(s) \lambda_Y(x, y, \alpha_s^{j \ell_1 \dots \ell_m}) + \mathbb{1}_{[S_{k-1}, +\infty)}(s) \lambda_Y(x, y, \alpha_s^i) \\ \bar{\sigma}_Y(s, x) &= \mathbb{1}_{[t, S_0)} \sigma_Y(x, \alpha_s^j) + \sum_{m=1}^{k-1} \mathbb{1}_{[S_{m-1}, S_m)}(s) \sigma_Y(x, \alpha_s^{j \ell_1 \dots \ell_m}) + \mathbb{1}_{[S_{k-1}, +\infty)}(s) \sigma_Y(x, \alpha_s^i)\end{aligned}$$

for  $(s, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$ . Under Assumption A1, those coefficients satisfy classical Lipschitz and boundedness assumption to have uniqueness and stability of solutions. In the sequel, we shall refer by  $X^{t, \mu, \alpha, i}$  and  $Y^{t, \hat{\mu}, \alpha, i}$  either to the processes themselves or to their extended definitions if the processes are considered outside their living interval.

Under the additional regularity assumption on the coefficients with respect to the control, we have a stability result for the branching system.

**Proposition 2.2.** *Suppose that Assumptions A1 holds and fix  $t \in [0, T]$ ,  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$ ,  $\hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)} \in E_{d+1}$ ,  $\alpha \in \mathcal{A}$ . Let  $(t_n)_{n \geq 1}$ ,  $(\hat{\mu}_n = \sum_{i \in V_n} \delta_{(i, x_i^n, y_i^n)})_{n \geq 1}$  and  $(\alpha^n)_{n \geq 1}$  be sequences of  $\mathbb{R}_+$ ,  $E_{d+1}$  and  $\mathcal{A}$  such that*

$$(t_n, \hat{\mu}_n) \xrightarrow[n \rightarrow +\infty]{} (t, \hat{\mu}),$$

and

$$\mathbb{E} \int_0^T d_A(\alpha_s^i, \alpha_s^{n, i}) ds \xrightarrow[n \rightarrow +\infty]{} 0$$

for all  $i \in \mathcal{I}$ . Then,

$$\mathbb{E} \left[ \left( |X_s^{t_n, \mu_n, \alpha_n, i} \mathbb{1}_{i \in \mathcal{V}_s^{t_n, \mu_n}} - X_s^{t, \mu, \alpha, i} \mathbb{1}_{i \in \mathcal{V}_s^{t, \mu}}|^2 + |Y_s^{t_n, \hat{\mu}_n, \alpha_n, i} \mathbb{1}_{i \in \mathcal{V}_s^{t_n, \mu_n}} - Y_s^{t, \hat{\mu}, \alpha, i} \mathbb{1}_{i \in \mathcal{V}_s^{t, \mu}}|^2 \right) \right] \xrightarrow[n \rightarrow +\infty]{} 0$$

for all  $s \in [t, T]$ , where  $\mu_n = \sum_{i \in V_n} \delta_{(i, x_i^n)} \in E_d$  for any  $n \geq 1$ .

*Proof.* We proceed in three steps.

*Step 1.* We first prove that

$$\mathbb{1}_{\mathcal{V}_s^{t_n, \mu_n}}(i) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}\text{-a.s.}} \mathbb{1}_{\mathcal{V}_s^{t, \mu}}(i)$$

for all  $i \in \mathcal{I}$ . For that, we distinguish two cases.

*Case 1.* Suppose that  $\mathbb{1}_{\mathcal{V}_s^{t, \mu}}(i) = 1$ . Then, there exist  $j \in \mathcal{V}_t^{t, \mu}$  and  $\ell_1, \dots, \ell_k$  such that  $i = j \ell_1 \dots \ell_k$  and

$$t < S_1 < \dots < S_{k-1} \leq s < S_k$$

where  $S_1, \dots, S_k$  are the successive branching times:

$$S_m = \inf \left\{ r > S_{m-1} : Q^{j\ell_1 \dots \ell_m} ((S_{m-1}, r] \times \{n_m\}) = 1 \right\} \quad (2.14)$$

with  $n_m \geq \ell_{m+1} + 1$  for  $m = 1, \dots, k$ . Since  $\hat{\mu}_n \rightarrow \hat{\mu}$  and  $j \in \mathcal{V}_t^{t, \mu}$ , there exists  $N \geq 1$  such that

$$j \in \mathcal{V}_{t_n}^{t_n, \mu_n} \text{ for all } n \geq N. \quad (2.15)$$

We then get from (2.14) and (2.15) that

$$i \in \mathcal{V}_s^{t_n, \mu_n}.$$

for  $n$  large enough.

*Case 2.* Suppose that  $\mathbb{1}_{\mathcal{V}_s^{t, \mu}}(i) = 0$ . We then have two subcases.

*Subcase 2.1.* There exist  $j \in \mathcal{V}_t^{t, \mu}$  and  $\ell_1, \dots, \ell_k$  such that  $i = j\ell_1 \dots \ell_k$ . We then have

$$s > S_k \quad \text{or} \quad s < S_{k-1} \quad (2.16)$$

where  $S_1, \dots, S_k$  are defined by (2.14). Since  $\hat{\mu}_n \rightarrow \hat{\mu}$ , we have  $i \in \mathcal{V}_{t_n}^{t_n, \mu_n}$  for  $n$  large enough and we get from (2.16) that  $\mathbb{1}_{\mathcal{V}_s^{t_n, \mu_n}}(i) = 0$  large enough.

*Subcase 2.1.*  $j \notin \mathcal{V}_t^{t, \mu}$  for any  $j \preceq i$ . Since the set of ancestor of  $i$  is finite and  $\hat{\mu}_n \rightarrow \hat{\mu}$ , we get  $j \notin \mathcal{V}_{t_n}^{t_n, \mu_n}$  for any  $j \preceq i$  for  $n$  large enough. Therefore, we have  $\mathbb{1}_{\mathcal{V}_s^{t_n, \mu_n}}(i) = 0$  for  $n$  large enough.

*Step 2.* We prove that

$$\mathbb{E} \left[ \left( |X_s^{t_n, \mu_n, \alpha_n, i} - X_s^{t, \mu, \alpha, i}|^2 + |Y_s^{t_n, \hat{\mu}_n, \alpha_n, i} - Y_s^{t, \hat{\mu}, \alpha, i}|^2 \right) \right] \xrightarrow{n \rightarrow +\infty} 0$$

for  $s \in [0, T]$  and  $i \in \mathcal{I}$ . Since  $\hat{\mu}_n \rightarrow \hat{\mu}$  as  $n \rightarrow +\infty$ , we have  $X_{t_n}^{t_n, \mu_n, \alpha_n, i} \rightarrow X_t^{t, \mu, \alpha, i}$  and  $Y_{t_n}^{t_n, \hat{\mu}_n, \alpha_n, i} \rightarrow Y_t^{t, \hat{\mu}, \alpha, i}$  as  $n \rightarrow +\infty$ . Using Assumption A1 (iii), we can apply Theorem 8.1 in [18] and we get the result.

*Step 3.* We then write

$$\begin{aligned} \mathbb{E} \left[ \left( |X_s^{t_n, \mu_n, \alpha_n, i} \mathbb{1}_{i \in \mathcal{V}_s^{t_n, \mu_n}} - X_s^{t, \mu, \alpha, i} \mathbb{1}_{i \in \mathcal{V}_s^{t, \mu}}|^2 + |Y_s^{t_n, \hat{\mu}_n, \alpha_n, i} \mathbb{1}_{i \in \mathcal{V}_s^{t_n, \mu_n}} - Y_s^{t, \hat{\mu}, \alpha, i} \mathbb{1}_{i \in \mathcal{V}_s^{t, \mu}}|^2 \right) \right] &\leq \\ &2\mathbb{E} \left[ \left( |X_s^{t_n, \mu_n, \alpha_n, i} - X_s^{t, \mu, \alpha, i}|^2 + |Y_s^{t_n, \hat{\mu}_n, \alpha_n, i} - Y_s^{t, \hat{\mu}, \alpha, i}|^2 \right) \right] \\ &+ 2\mathbb{E} \left[ \left( |X_s^{t, \mu, \alpha, i}|^2 + |Y_s^{t, \hat{\mu}, \alpha, i}|^2 \right) \left( \mathbb{1}_{i \in \mathcal{V}_s^{t, \mu}} - \mathbb{1}_{i \in \mathcal{V}_s^{t_n, \mu_n}} \right)^2 \right]. \end{aligned}$$

Using the dominated convergence theorem we get from Step 1

$$\mathbb{E} \left[ \left( |X_s^{t, \mu, \alpha, i}|^2 + |Y_s^{t, \hat{\mu}, \alpha, i}|^2 \right) \left( \mathbb{1}_{i \in \mathcal{V}_s^{t, \mu}} - \mathbb{1}_{i \in \mathcal{V}_s^{t_n, \mu_n}} \right)^2 \right] \xrightarrow{n \rightarrow +\infty} 0.$$

This last convergence and Step 2 give the result.  $\square$

Focusing on conditional laws of the controlled processes, we have a representation result. Define  $\mathbb{D}([0, T], \mathcal{M}_F(I \times \mathbb{R}^{m+1}))$  as the set of *càdlàg* functions from  $[0, T]$  to  $\mathcal{M}_F(I \times \mathbb{R}^{m+1})$ . We endow this set with the Skorokhod metric related to the Prokhorov distance and the related Borel  $\sigma$ -algebra. From Doob's functional representation Theorem (see e.g. Lemma 1.13 in [16]) for any control  $\alpha$ , there exists a  $\mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{D}([0, T], \mathcal{M}_F(I \times \mathbb{R}^{m+1})))$ -measurable function  $\tilde{\alpha} : [0, T] \times$

$\mathbb{D}([0, T], \mathcal{M}_F(I \times \mathbb{R}^{m+1})) \rightarrow A^{\mathcal{I}}$  such that  $\alpha^i(\omega) = \tilde{\alpha}^i(s, \xi(\omega_{\cdot \wedge s})) = \tilde{\alpha}^i(s, \xi(\omega))$  for any  $s \in [0, T]$ ,  $\omega \in \Omega$  and  $i \in \mathcal{I}$ . In the sequel, we identify the control  $\alpha$  with its related function  $\tilde{\alpha}$  and we still denote by  $\mathcal{A}$  the set of those controls.

For  $\alpha \in \mathcal{A}$ , an  $\mathbb{F}$ -stopping time  $\tau$  and  $\omega \in \Omega$ , we define the control  $\alpha^{\tau(\omega), \omega}$  by

$$\left(\alpha^{\tau(\omega), \omega}\right)^i(s, \xi(\tilde{\omega})) = \alpha^i\left(s, \xi^{\tau(\omega), \omega}(\tilde{\omega})\right)$$

for  $i \in \mathcal{I}$ ,  $s \geq 0$  and  $\tilde{\omega} \in \Omega$ , where  $\xi^{\omega, \tau(\omega)}$  is given by (2.4).

**Theorem 2.1** (Conditioning property). *Suppose that Assumption A1 holds and fix  $t \in [0, T]$ ,  $\hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)} \in E_{d+1}$  and  $\alpha \in \mathcal{A}$ . Then, for any bounded measurable function  $f : \mathbb{D}([0, T], E_{d+1}) \rightarrow \mathbb{R}$  and any  $\mathbb{F}$ -stopping time  $\tau$ , we have*

$$\mathbb{E} \left[ f \left( \hat{X}^{t, \hat{\mu}, \alpha} \right) \middle| \mathcal{F}_\tau \right] (\omega) = F \left( \tau(\omega), \hat{X}_{\cdot \wedge \tau}^{t, \hat{\mu}, \alpha}(\omega), \alpha^{\tau(\omega), \omega} \right), \quad \mathbb{P}(d\omega) - a.s.$$

where

$$F(s, \hat{x}, \beta) = \mathbb{E} \left[ f \left( (\hat{x}_t \mathbf{1}_{t < s} + \hat{X}_t^{s, \hat{x}, \beta} \mathbf{1}_{t \geq s})_{t \in [0, T]} \right) \right]$$

for all  $s \in [0, T]$ ,  $\hat{x} \in \mathbb{D}([0, T], E_{d+1})$  and  $\beta \in \mathcal{A}$ .

The proof of this result is postponed to Appendix A.3. It follows the same lines as the proof of Theorem 2 in [6], and relies on a uniqueness property for the related branching martingale controlled problem which is studied in Appendix A.2.

## 2.2 The stochastic target problem

To define the stochastic target problem, let  $g : \mathcal{I} \times \mathbb{R}^d \rightarrow \mathbb{R}$  be a function satisfying the following assumption.

**Assumption A2.** *The function  $g_i$  is continuous on  $\mathbb{R}^d$  for all  $i \in \mathcal{I}$ .*

Fix an initial time  $t \in [0, T]$  and an initial population  $\mu = \sum_{i \in V} \delta_{(i, x_i)}$ . We look for an initial position  $y$  for the target process and a control  $\alpha \in \mathcal{A}$  such that

$$Y_t^{t, \hat{\mu}, \alpha, i} = y, \quad i \in V,$$

and  $Y^{t, \hat{\mu}, \alpha}$  and  $X^{t, \mu, \alpha}$  satisfies the terminal constraints

$$Y_T^{t, \hat{\mu}, \alpha, i} \geq g_i(X_T^{t, \mu, \alpha, i}), \quad i \in \mathcal{V}_T^{t, \mu}.$$

More precisely, we look for the *reachability set*

$$\mathcal{R}(t, \mu) = \left\{ y \in \mathbb{R}, \quad : \exists \alpha \in \mathcal{A} : Y_T^{t, \hat{\mu}, \alpha, i} \geq g_i(X_T^{t, \mu, \alpha, i}), \quad i \in \mathcal{V}_T^{t, \mu} \right. \\ \left. \text{with } \hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y)} \right\}.$$

for  $t \in [0, T]$  and  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$ . Since the target processes  $Y^i$  has an explicit impact only on its drift  $\lambda_Y$  and not on its diffusion coefficient  $\sigma_Y$ , the reachability set satisfies the following *monotonicity property*.

**Proposition 2.3.** *Suppose that Assumptions A1 holds. For  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$  and  $y \in \mathcal{R}(t, \mu)$  we have  $[y, \infty[ \subseteq \mathcal{R}(t, \mu)$ .*

*Proof.* Fix a control  $\alpha = (\alpha^i)_{i \in \mathcal{I}}$  and a starting point  $(t, \mu)$ . We take  $y \in \mathcal{R}(t, \mu)$ ,  $y' \geq y$  and write  $\hat{\mu}$  (resp.  $\hat{\mu}'$ ) for  $\sum_{i \in V} \delta_{(i, x_i, y)}$  (resp.  $\sum_{i \in V} \delta_{(i, x_i, y')}$ ),  $Y^i$  (resp.  $Y'^i$ ) for  $Y^{t, \hat{\mu}, \alpha, i}$  (resp.  $Y^{t, \hat{\mu}', \alpha, i}$ ) and  $\delta Y^i$  for  $Y'^i - Y^i$ . We then have

$$\delta Y_s^i = (y' - y) + \int_t^s \chi_u \delta Y_u^i du$$

for  $s \geq t$ , where  $\chi$  is given by

$$\chi_u := \frac{\bar{\lambda}_Y(u, X_u^i, Y_u^i) - \bar{\lambda}_Y(u, X_u^i, Y_u'^i)}{\delta Y_u^i}, \quad u \geq 0,$$

with  $\bar{\lambda}_Y$  defined in Remark 2.1. From the Lipschitz property of  $\lambda_Y$  in Assumption A1,  $\chi$  is bounded and

$$\delta \bar{Y}_T^i = (y' - y) \exp\left(\int_t^T \chi_u du\right) \geq 0, \quad \mathbb{P} - a.s.$$

Since  $y \in \mathcal{Y}(t, \mu)$ , we get

$$Y_T^{t, \mu, \alpha, y', i} \geq Y_T^{t, \mu, \alpha, y, i} \geq g_i(X_T^{t, \mu, \alpha, i}), \quad \mathbb{P} - a.s.$$

This is true for all  $i \in \mathcal{V}_T^{t, \mu}$ , therefore  $y' \in \mathcal{R}(t, \mu)$ .  $\square$

From Proposition 2.3, the closure  $\overline{\mathcal{R}(t, \mu)}$  of the reachability set is a half line interval characterized by its lower bound. We then define the value function  $v$  as the infimum of  $\mathcal{R}$ :

$$\begin{aligned} v(t, \mu) &:= \inf \mathcal{R}(t, \mu) \\ &= \inf \left\{ y \in \mathbb{R} : \exists \alpha \in \mathcal{A}, Y_t^{t, \hat{\mu}, \alpha, i} = y \forall i \in V, \right. \\ &\quad \left. \text{and } Y_T^{t, \hat{\mu}, \alpha, i} \geq g_i(X_T^{t, \mu, \alpha, i}) \forall i \in \mathcal{V}_T^{t, \mu} \text{ a.s.} \right\} \end{aligned} \quad (2.17)$$

for all  $t \in [0, T]$  and  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$ , with the usual convention that  $\inf(\emptyset) = +\infty$ . Our aim is to provide an analytical characterisation of the value function  $v$ .

**Remark 2.2.** *The value function  $v$  or the reachability set  $\mathcal{R}$  might not be well defined in the case where an extinction of the alive population of particle happens before  $T$ . In this case we take the convention that the terminal condition is always satisfied if  $\mathcal{V}_T^{t, \mu} = \emptyset$ . In the sequel, we keep this convention for other constraints on  $(X_\theta^{t, \mu, \alpha, i}, Y_\theta^{t, \hat{\mu}, \alpha, i})$  with  $\mathcal{V}_\theta^{t, \mu} = \emptyset$  and  $\theta$  a stopping time.*

We next provide a new formulation of the function  $v$ .

**Proposition 2.4.** *Under Assumptions A1, the value function  $v$  satisfies the following identity*

$$\begin{aligned} v(t, \mu) &= \inf \left\{ y \in \mathbb{R} : \exists \alpha \in \mathcal{A}, \exists \hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)} \in E_{d+1} \text{ such that} \right. \\ &\quad \left. y_i \leq y \forall i \in V, \text{ and } Y_T^{t, \hat{\mu}, \alpha, i} \geq g_i(X_T^{t, \mu, \alpha, i}) \forall i \in \mathcal{V}_T^{t, \mu} \text{ a.s.} \right\} \end{aligned} \quad (2.18)$$

for all  $t \in [0, T]$  and  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$ .

*Proof.* Denote by  $\tilde{v}(t, \mu)$  the right hand side of (2.18). Since the set whose infimum is  $v(t, \mu)$  is included in the one whose infimum is  $\tilde{v}(t, \mu)$ , we obviously have

$$\tilde{v}(t, \mu) \leq v(t, \mu).$$

Fix now  $y \in \mathbb{R}$  for which there exist  $\alpha \in \mathcal{A}$  and  $\hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)} \in E_{d+1}$  such that

$$y_i \leq y, \quad i \in V,$$

and

$$Y_T^{t, \hat{\mu}, \alpha, i} \geq g_i(X_T^{t, \mu, \alpha, i}), \quad i \in \mathcal{V}_T^{t, \mu}$$

Set  $\bar{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)} \in E_{d+1}$ . By the comparison argument used in the proof of Proposition 2.3, we have

$$Y_T^{t, \bar{\mu}, \alpha, i} \geq Y_T^{t, \hat{\mu}, \alpha, i} \geq g_i(X_T^{t, \mu, \alpha, i}), \quad i \in \mathcal{V}_T^{t, \mu}$$

Therefore  $y \geq v(t, \mu)$  and  $\tilde{v}(t, \mu) \geq v(t, \mu)$ . □

### 2.3 An example of application from fintech

Fintech is the contraction of the words finance and technology. It refers to recent technologies that allows for the improvement and the automation of the delivery and use of financial services. The field has emerged at the beginning of the 21-st century and covered technologies used by established financial institutions. Since that time, the field has evolved to also include crypto-currencies which are decentralised financial assets. Those assets are based on the block-chain technology. The main idea of that structure is to keep any new transaction registered in a chain by adding new blocks and sharing the extension of the original chain over the network, so that every user keeps in mind the transaction and can certify it. We refer to [20] for a description of how a block-chain base crypto-currency works in the case of the Bitcoin.

Due to the structure of this kind of assets, a fork can appear in the chain (see [10]). In this case, the original asset is transformed into several assets. A natural question that arises is how to evaluate an option on crypto-currencies in this case. We present here the example of the super-replication of options on asset that may fork and show that it is a particular case of the branching stochastic target presented above.

We consider a financial market on which is defined a crypto-currency with price process  $(S_t)_{t \in [0, T]}$ . We suppose that the process  $S$  is a branching diffusion and describe its dynamics. We first define the set  $\mathcal{V}_t$  of alive particles at time  $t \in [0, T]$  as previously done in Section 2.1. The initial condition for the process  $S$  is a constant ( $S_0 > 0$ ). Assume the version  $i \in \mathcal{I}$  of the crypto-currency is alive at time  $t \in [0, T]$ , dies at some random time  $\tau_i \geq t$  and gives birth to  $k$  new versions  $i_0, \dots, i_{k-1}$ . The position at a time  $s \geq \tau_i$  of the new the crypto-currencies are given by

$$S_{\tau_i}^{i\ell} = S_{\tau_i}^i \tag{2.19}$$

$$dS_s^{i\ell} = S_s^{i\ell} \left( bds + cdB_s^{i\ell} \right), \tag{2.20}$$

for  $\ell = 0, \dots, k-1$  and  $s \geq \tau_i$  such that version  $i\ell$  is alive at time  $s$ . Here  $b$  and  $c$  are two positive constants.

In addition to that asset, we assume that there exists on the market a non-risky asset  $S^0$  with deterministic interest rate  $r > 0$  and with initial condition  $S_0^0 = 1$ , that is  $S_t^0 = e^{rt}$  for  $t \in [0, T]$ .

An investment strategy consists in a process  $\pi = (\pi_t^i)_{t \in [0, T], i \in \mathcal{I}}$  of  $\mathbb{F}$ -progressive processes valued in  $[0, 1]$ , where  $\pi_t^i$  represents the proportion of the wealth invested in the version  $S^i$  of the crypto-currency. We denote by  $\mathcal{A}$  the set of such strategies. For  $\pi \in \mathcal{A}$ , we also denote by  $V^{V_0, \pi}$  the self financing wealth process related to the initial capital  $V_0$  and strategy  $\pi$ . According to (2.19)-(2.20) it is given by

$$V_{\tau_i}^{V_0, \pi, i\ell} = V_{\tau_i}^{V_0, \pi, i} \quad (2.21)$$

$$dV_s^{V_0, \pi, i\ell} = V_s^{V_0, \pi, i\ell} \left( ((b-r)\pi_s^{i\ell} + r)ds + c\pi_s^{i\ell} dB_s^{i\ell} \right), \quad (2.22)$$

for  $\ell = 0, \dots, k-1$  and  $s \geq \tau_i$  such that version  $i\ell$  is alive at time  $s$ .

We then consider a financial derivative on the asset  $S$  that consists in a Put Option but with a strike  $K_i$  depending on the version of the crypto-currency  $S$ . Such a product can express the need to hedge against a decrease of the value of the asset  $S$  that depends on the branch.

The computation of the super-replication problem leads to solve the following stochastic target problem

$$w_0 = \inf \left\{ \nu \in \mathbb{R}_+ : \exists \pi \in \mathcal{A}, V_T^{\nu, \pi, i} \geq (K_i - S_T^i)_+ + \kappa \forall i \in \mathcal{V}_T \text{ a.s.} \right\},$$

where  $\kappa$  is a positive constant representing some friction. We next modify this problem to satisfy our assumptions. For that, we first define the processes

$$\begin{aligned} Y_t^{y, \pi, i} &= \log \left( V_t^{e^y, \pi, i} \right) \\ X_t^i &= \log \left( S_t^i \right) \end{aligned}$$

for  $t \in [0, T]$  and  $i \in \mathcal{V}_t$ . From (2.19)-(2.20) and (2.21)-(2.21), we get

$$X_{\tau_i}^{i\ell} = X_{\tau_i}^i, \quad dX_s^{i\ell} = \left( b - \frac{c^2}{2} \right) ds + c dB_s^{i\ell}, \quad (2.23)$$

$$Y_{\tau_i}^{y, \pi, i\ell} = Y_{\tau_i}^{y, \pi, i}, \quad dY_s^{y, \pi, i\ell} = \left( (b-r)\pi_s^{i\ell} - \frac{1}{2}c^2(\pi_s^{i\ell})^2 + r \right) ds + c\pi_s^{i\ell} dB_s^{i\ell}, \quad (2.24)$$

for  $\ell = 0, \dots, k-1$  and  $s \geq \tau_i$  such that version  $i\ell$  is alive at time  $s$ . We observe that the dynamics of the processes  $Y$  and  $X$  satisfy Assumption A1. We also define the functions  $g$  as

$$g_i(x) = \log \left( (K_i - e^x)_+ + \kappa \right), \quad (x, i) \in \mathbb{R} \times \mathcal{I},$$

which satisfies Assumption A2. Finally, we define the optimal value

$$v_0 = \inf \left\{ y \in \mathbb{R} : \exists \pi \in \mathcal{A}, Y_T^{y, \pi, i} \geq g_i(X_T^i) \forall i \in \mathcal{V}_T \text{ a.s.} \right\},$$

a special case of (2.17). We notice that the optimal value  $w_0$  is related to  $v_0$  by

$$w_0 = \exp(v_0).$$

We suppose that  $\bar{K} := \sup_{i \in \mathcal{I}} K_i < +\infty$ . The value function  $v$  related to  $v_0$  is then bounded. Indeed, by taking the initial condition  $t \in [0, T]$  and  $y = -r(T-t) + \log(\bar{K} + \kappa)$  and the control  $\pi_t^i = 0$  for  $i \in \mathcal{I}$  and  $t \in [0, T]$ , we get from (2.24)

$$Y_T^{t, \hat{\mu}, \pi, i} \geq g_i(X_T^{t, \mu, i}), \quad i \in \mathcal{V}_T^{t, \mu}$$



for  $\mu = \sum_{i \in V} \delta_{(i, x^i)} \in E_d$  and  $\mu = \sum_{i \in V} \delta_{(i, x^i, y)} \in E_{d+1}$ . Therefore

$$v(t, \mu) \leq -r(T-t) + \log(\bar{K} + \kappa), \quad (t, \mu) \in [0, T] \times E_d.$$

Moreover, for any  $y \in \mathcal{R}(t, \mu)$  and  $\pi$  the related admissible control, we have

$$\left( (b-r)\pi_s^{i\ell} - \frac{1}{2}c^2(\pi_s^{i\ell})^2 + r \right) \leq \left( \frac{b-r}{c} \right)^2 + r$$

Therefore we get

$$y + \left( \left( \frac{b-r}{c} \right)^2 + r \right) (T-t) \geq \mathbb{E} \left[ Y_T^{t, \hat{\mu}, \pi, i} \right] \geq \mathbb{E} \left[ g_i(X_T^{t, \mu, \pi, i}) \right] \geq \log(\kappa).$$

Therefore,

$$v(t, \mu) \geq - \left( \left( \frac{b-r}{c} \right)^2 + r \right) (T-t) + \log(\bar{K} + \kappa), \quad (t, \mu) \in [0, T] \times E_d.$$

In particular,  $v$  satisfies the growth condition (4.66) of the comparison Theorem 4.5. If we suppose also that  $r = 0$  and  $g_i = 0$  for  $i \in \mathcal{I}$  of the form  $i = i_1 \cdots i_n$  with  $i_\ell \geq I$  for some  $\ell$  where  $I$  is a given bound, then  $v$  also satisfies condition (4.65) of Theorem 4.5.

## 3 Dynamic programming

### 3.1 Measurable selection

In establishing a dynamic programming principle, we need an admissible control as concatenation of admissible controls depending on the position of the branching processes at an intermediary time. For this end, we use a *measurable selection approach*.

Let  $\mathcal{U}$  be the *target set* defined by

$$\mathcal{U}(t, \hat{\mu}) = \left\{ \alpha \in \mathcal{A} : Y_T^{t, \hat{\mu}, \alpha, i} \geq g_i(X_T^{t, \mu, \alpha, i}) \forall i \in \mathcal{V}_T^{t, \mu, \alpha} \text{ a.s.} \right\},$$

for  $(t, \hat{\mu}) \in [0, T] \times E_{d+1}$  with  $\hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)}$  and  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$ . Let  $S := [0, T] \times E_{d+1}$  and

$$D := \{(t, \hat{\mu}) \in S : \mathcal{U}(t, \hat{\mu}) \neq \emptyset\}.$$

Our aim is to exhibit a function that associates to each  $(t, \hat{\mu}) \in D$  a control  $\alpha \in \mathcal{U}(t, \hat{\mu})$  in a measurable way.

We denote by  $\mathcal{P}(S)$  the set of probability measures on  $(S, \mathcal{B}([0, T]) \otimes \mathcal{B}(E_{d+1}))$  and we endow  $\mathcal{A}$  with the Borel  $\sigma$ -algebra  $\mathcal{B}(\mathcal{A})$  related to the distance

$$(\alpha, \alpha') \mapsto \sum_{i \in \mathcal{I}} \frac{1}{2^{|i|}} \wedge \mathbb{E} \int_0^T |\alpha_s^i - \alpha'_s{}^i| ds$$

where  $|i| = i_1 + \cdots + i_n$  for  $i = (i_1, \dots, i_n) \in \mathbb{N}^n$  and  $n \geq 1$ . We then have the following measurable selection result.

**Lemma 3.1.** *Suppose that Assumptions A1 and A2 hold. For each  $\nu \in \mathcal{P}(S)$ , there exists a measurable function  $\phi_\nu : (D, \mathcal{B}(D)) \rightarrow (\mathcal{A}, \mathcal{B}(\mathcal{A}))$  such that*

$$\phi_\nu(t, \hat{\mu}) \in \mathcal{U}(t, \hat{\mu}) \text{ for } \nu\text{-a.e. } (t, \hat{\mu}) \in D .$$

*Proof.*  $S$  being endowed with the product  $\sigma$ -algebra  $\mathcal{B}([0, T]) \otimes \mathcal{B}(E_{d+1})$  is a Borel space as product of Borel spaces. Also  $\mathcal{A}$  endowed with  $\mathcal{B}(\mathcal{A})$  is a Borel space. Let  $C$  be the following set

$$C := \{(t, \hat{\mu}) \in S \times \mathcal{A} : \alpha \in \mathcal{U}(t, \hat{\mu})\} .$$

From Proposition 2.2 and Assumption A2,  $C$  is closed and a fortiori a Borel subset of  $S \times \mathcal{A}$ .

• *Step 1: Measurable selector.*

Since  $C$  is a Borel set, it is analytic by [1, Proposition 7.36]. From the Jankov-von Neumann measurable selection theorem (see e.g. [1, Proposition 7.49]), there exists an analytically measurable function  $\phi : D \rightarrow \mathcal{A}$  such that

$$\{(t, \hat{\mu}, \phi(t, \hat{\mu})) : (t, \hat{\mu}) \in S\} \subset C .$$

• *Step 2: Construction of a Borel measurable  $\phi_\nu$  such that  $\phi_\nu = \phi$   $\nu$ -almost everywhere.*

Fix  $\nu \in \mathcal{P}(S)$  and denote by  $\mathcal{B}_\nu(S)$  the completion of the Borel  $\sigma$ -algebra  $\mathcal{B}(S)$  under  $\nu$ . From [1, Corollary 7.42.1] any analytic set is universally measurable. Therefore  $\phi$  is universally measurable, and, from the definition of the universal  $\sigma$ -algebra,  $\phi$  is  $\mathcal{B}_\nu(S)$ -measurable. Since  $\mathcal{B}_\nu(S)$  is the completion of  $\mathcal{B}(S)$  under  $\nu$ , there exists a Borel measurable map  $\phi_\nu$  such that  $\phi_\nu(t, \hat{\mu}) = \phi(t, \hat{\mu})$  for  $\nu$ -almost every  $(t, \hat{\mu}) \in S$ .  $\square$

## 3.2 Dynamic programming principle

For  $t \in [0, T]$ , we denote by  $\mathcal{T}_{[t, T]}$  the set of  $\mathbb{F}$  stopping times valued in  $[t, T]$ . The dynamic programming principle may be stated as follows.

**Theorem 3.2.** *Under Assumptions A1 and A2, the value function satisfies*

$$v(t, \mu) = \inf \left\{ y \in \mathbb{R} : \exists \alpha \in \mathcal{A}, \exists \hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)} \in E_{d+1} \text{ such that} \right. \\ \left. y_i \leq y \ \forall i \in V, \text{ and } Y_\theta^{t, \hat{\mu}, \alpha, i} \geq v \left( \theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i})} \right) \ \forall i \in \mathcal{V}_\theta^{t, \mu} \text{ a.s.} \right\} \quad (3.25)$$

for any  $(t, \mu) \in [0, T] \times E_d$  and  $\theta \in \mathcal{T}_{[t, T]}$ .

*Proof.* We first define the *reachability sets* by

$$\mathcal{Y}(t, \mu) := \left\{ (y_i)_{i \in V} \in \mathbb{R}^V : \mathcal{U}(t, \hat{\mu}) \neq \emptyset \text{ with } \hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)} \right\} .$$

and

$$\mathcal{Y}^\theta(t, \mu) = \left\{ (y_i)_{i \in V} \in \mathbb{R}^V : \exists \alpha \in \mathcal{A} \text{ such that} \right. \\ \left. Y_\theta^{t, \hat{\mu}, y, \alpha, i} \geq v \left( \theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i})} \right) \ \forall i \in \mathcal{V}_\theta^{t, \mu} \text{ a.s. with } \hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)} \in E_{d+1} \right\} .$$

for  $t \in [0, T]$  and  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$  and  $\theta \in \mathcal{T}_{[t, T]}$ . Fix now  $t \in [0, T]$  and  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$ . Denote by  $v_\theta(t, \mu)$  the right hand side of (3.25).

To prove  $v(t, \mu) \geq v_\theta(t, \mu)$ , we show that  $\mathcal{Y}(t, \mu) \subset \mathcal{Y}^\theta(t, \mu)$ . Let  $(y_i)_{i \in V} \in \mathcal{Y}(t, \mu)$ . By definition there exists  $\alpha \in \mathcal{A}$  such that

$$Y_T^{t, \hat{\mu}, \alpha, j} \geq g_j \left( X_T^{t, \mu, \alpha, j} \right) \quad \forall j \in \mathcal{V}_T^{t, \mu}.$$

From the uniqueness of solutions to (2.5)-(2.6) and (2.8)-(2.9) (or equivalently (2.12)-(2.12)) we get the following flow property

$$\begin{aligned} X_T^{t, \mu, \alpha, j} &= X_T^{\theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i})}, \alpha, j}, \\ Y_T^{t, \hat{\mu}, \alpha, j} &= Y_T^{\theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i}, Y_\theta^{t, \hat{\mu}, y, \alpha, i})}, \alpha, j}, \end{aligned}$$

for all  $i \in \mathcal{V}_\theta^{t, \mu}$  and  $j \in \mathcal{V}_T^{t, \mu}$  such that  $i \preceq j$ .

We therefore get

$$Y_T^{\theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i}, Y_\theta^{t, \hat{\mu}, y, \alpha, i})}, \alpha, j} \geq g_j \left( X_T^{\theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i})}, \alpha, j} \right), \quad j \in \mathcal{V}_T^{\theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i})}},$$

for all  $i \in \mathcal{V}_\theta^{t, \mu}$ . Given the definition of the value function  $v$ , we get  $Y_\theta^{t, \hat{\mu}, \alpha, i} \geq v \left( \theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i})} \right)$  for all  $i \in \mathcal{V}_\theta^{t, \mu}$  a.s. and  $(y_i)_{i \in V} \in \mathcal{Y}^\theta(t, \mu)$ .

We now turn to the reverse inequality  $v_\theta(t, \mu) \geq v(t, \mu)$ . To this end, we prove that  $\mathcal{Y}_\varepsilon^\theta(t, \mu) \subset \mathcal{Y}(t, \mu)$  for any  $\varepsilon > 0$ , where

$$\mathcal{Y}_\varepsilon^\theta(t, \mu) = \left\{ (y_i + \varepsilon)_{i \in V} : (y_i)_{i \in V} \in \mathcal{Y}^\theta(t, \mu) \right\}.$$

Let  $(y_i)_{i \in V} \in \mathcal{Y}^\theta(t, \mu)$  and  $\alpha \in \mathcal{A}$  such that  $Y_\theta^{t, \hat{\mu}, \alpha, i} \geq v \left( \theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i})} \right)$  for all  $i \in \mathcal{V}_\theta^{t, \mu}$  a.s. where  $\hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)}$ . Fix now  $\varepsilon > 0$  and set  $\hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y_i)}$  and  $\hat{\mu}_\varepsilon = \sum_{i \in V} \delta_{(i, x_i, y_i + \varepsilon)}$ . From the definition of the value function and the strict monotonicity of the flow w.r.t. the initial value, we get  $Y_\theta^{t, \hat{\mu}, \alpha, i}(\omega) < Y_\theta^{t, \hat{\mu}_\varepsilon, \alpha, i}(\omega) \in \mathcal{Y} \left( \theta, \delta_{(i, X_\theta^{t, \mu, \alpha, i})} \right)(\omega)$  for all  $i \in \mathcal{V}_\theta^{t, \mu}$  for  $\mathbb{P}$ -a.e.  $\omega \in \Omega$ . Consider the probability measure  $\nu$  induced on  $S$  by

$$\omega \mapsto \left( \theta, \hat{Z}_\theta^{t, \hat{\mu}_\varepsilon, \alpha} \right)(\omega),$$

and  $\phi_\nu$  the measurable map defined in Lemma 3.1. We have

$$Y_T^{\tilde{t}, \tilde{\mu}, \tilde{y}, \phi_\nu(\tilde{t}, \tilde{\mu}, \tilde{y}), i} \geq g_i \left( X_T^{\tilde{t}, \tilde{\mu}, \phi_\nu(\tilde{t}, \tilde{\mu}, \tilde{y}), i} \right) \quad \forall i \in \mathcal{V}_T \text{ } \mathbb{P}\text{-a.s. for } \nu\text{-a.e. } (\tilde{t}, \tilde{\mu}, \tilde{y}) \in D. \quad (3.26)$$

We define  $\hat{\Xi} := \left( \theta, \hat{Z}_\theta^{t, \hat{\mu}_\varepsilon, \alpha} \right)$  and  $\Xi := \left( \theta, Z_\theta^{t, \mu, \alpha} \right)$ . For  $i_T \in \mathcal{V}_T^{t, \mu}$ , if  $i_\theta \in \mathcal{V}_\theta^{t, \mu}$  such that  $i_\theta \preceq i_T$ , the initial conditions at time  $\theta$  for  $Y_T^{i_T}$  and  $X_T^{i_T}$  are respectively  $\Xi$  and  $\hat{\Xi}$ .

Binding flow properties with the measurable selector, we can find a negligible set  $N_1$  of  $\mathcal{F}$  such that there exists negligible set  $N_{2, \omega_1}$  of  $\mathcal{F}$  for each  $\omega_1 \in N_1^c$  such that

$$Y_T^{\hat{\Xi}(\omega_1), \phi_\nu(\hat{\Xi}(\omega_1)), i}(\omega_2) \geq g_{i_T} \left( X_T^{\Xi(\omega_1), \phi_\nu(\hat{\Xi}(\omega_1)), i}(\omega_2) \right) \quad \forall i \in \mathcal{V}_T^{\Xi(\omega_1)}(\omega_2)$$

for all  $\omega_1 \in N_1^c$  and  $\omega_2 \in N_{2, \omega_1}^c$ .

We now define the set  $\bar{N} := \{\omega : \omega \in N_1^c, \omega \in N_{2,\omega}\}$  and we prove that  $\bar{N}$  is negligible. We first have  $\bar{N} \subset N_1^c \cap \bar{N}_2$  where

$$\bar{N}_2 = \left\{ \omega \in \Omega : \exists i \in \mathcal{V}_T^{\Xi(\omega)}(\omega), Y_T^{\hat{\Xi}(\omega), \phi_\nu(\hat{\Xi}(\omega)), i}(\omega) < g_i \left( X_T^{\Xi(\omega), \phi_\nu(\hat{\Xi}(\omega)), i}(\omega) \right) \right\}.$$

The set  $\bar{N}_2$  can be rewritten as

$$\bar{N}_2 = \left\{ \omega \in \Omega : \prod_{i \in \mathcal{V}_T^{\Xi(\omega)}} \mathbb{1}_{Y_T^{\hat{\Xi}(\omega), \phi_\nu(\hat{\Xi}(\omega)), i}(\omega) \geq g_i \left( X_T^{\Xi(\omega), \phi_\nu(\hat{\Xi}(\omega)), i}(\omega) \right)} = 0 \right\}.$$

Taking the conditional expectation w.r.t.  $\mathcal{F}_\theta$ , we have up to a negligible set

$$\bar{N}_2 = \left\{ \omega \in \Omega : \mathbb{E} \left[ \prod_{i \in \mathcal{V}_T^{\Xi(\omega)}} \mathbb{1}_{Y_T^{\hat{\Xi}(\omega), \phi_\nu(\hat{\Xi}(\omega)), i}(\omega) \geq g_i \left( X_T^{\Xi(\omega), \phi_\nu(\hat{\Xi}(\omega)), i}(\omega) \right)} \middle| \mathcal{F}_\theta \right] (\omega) = 0 \right\}.$$

Using Theorem 2.1 we get

$$\begin{aligned} \bar{N}_2 &= \left\{ \omega \in \Omega : \int_{\Omega} \prod_{i \in \mathcal{V}_T^{\Xi(\omega)}(\omega \oplus_\theta \omega')} \mathbb{1}_{Y_T^{\hat{\Xi}(\omega), \phi_\nu(\hat{\Xi}(\omega)), i}(\omega \oplus_\theta \omega') \geq g_i \left( X_T^{\Xi(\omega), \phi_\nu(\hat{\Xi}(\omega)), i}(\omega \oplus_\theta \omega') \right)} d\mathbb{P}(\omega') = 0 \right\} \\ &= \{ \omega \in \Omega : \mathbb{P}(N_{2,\omega}) = 0 \}. \end{aligned}$$

Therefore we get, up to a negligible set,  $\bar{N}_2 \subset N_1$  and  $\mathbb{P}(\bar{N}_2) = 0$

We now fix  $\alpha \in A$  and define the control  $\bar{\alpha} = (\bar{\alpha}^i)_{i \in \mathcal{I}}$  by

$$\bar{\alpha}^i(\omega) := \begin{cases} \alpha^i(\omega) \mathbb{1}_{[0, \theta(\omega))} + \phi_\nu^i(\hat{\Xi}(\omega))(\omega) \mathbb{1}_{[\theta(\omega), T]} & \text{if } \omega \in \Omega \setminus \bar{N} \\ a & \text{if } \omega \in \bar{N} \end{cases}$$

for all  $i \in \mathcal{I}$  with  $a \in A$ . Since  $\left( Y_T^{\hat{\Xi}(\omega), \phi_\nu(\hat{\Xi}(\omega)), i}, X_T^{\Xi(\omega), \phi_\nu(\hat{\Xi}(\omega)), i} \right) = \left( Y_T^{t, \hat{\mu}_\varepsilon, \bar{\alpha}, i}, X_T^{t, \mu, \bar{\alpha}, i} \right)$  for each  $i \in \mathcal{V}_T^{t, \mu}$  a.s. and  $\bar{N}_2$  is negligible, we get  $(y_i + \varepsilon)_{i \in V} \in \mathcal{Y}(t, \mu)$ .  $\square$

## 4 PDE characterisation

### 4.1 Branching property

Conditionally to their birth, the alive particles, and consequently their branches, are independent in the uncontrolled case. In our case, this *branching property* is passed down to the value function in the following way.

**Proposition 4.5** (Branching property). *Let Assumption A1 holds. The value function  $v$  satisfies*

$$v(t, \mu) = \max_{i \in V} v(t, \delta_{(i, x^i)}) \quad (4.27)$$

for any  $(t, \mu = \sum_{i \in V} \delta_{(i, x^i)}) \in [0, T] \times E_d$ .

*Proof.* For  $\mu = \sum_{i \in V} \delta_{(i, x_i)} \in E_d$ , we define

$$K^\mu := \left\{ y \in \mathbb{R} : \exists \alpha \in \mathcal{A}, Y_T^{t, \hat{\mu}, \alpha, i} \geq g_i(X_T^{t, \mu, \alpha, i}) \forall i \in V \text{ a.s. with } \hat{\mu} = \sum_{i \in V} \delta_{(i, x_i, y)} \right\}.$$

Proving  $v(t, \mu) \geq \max_{i \in V} v(t, \delta_{(i, x_i)})$  comes to verify that  $K^\mu \subseteq \bigcap_{j \in V} K^{\delta_{(j, x^j)}}$ , i.e.  $K^\mu \subseteq K^{\delta_{(j, x^j)}}$  for each  $j \in V$ . If  $y \in K^\mu$ , there exists  $\alpha$  satisfying the constraints in  $T$  a.s. With this same  $\alpha$ , zooming in on the sub-population generated by each  $j \in V$ , we must satisfy the condition of  $K^{\delta_{(j, x^j)}}$ . Therefore,  $y \in K^{\delta_{(j, x^j)}}$ .

Let  $j$  be the index that realises the maximum in the righthand side of (4.27). The monotonicity property given by Proposition 2.3 implies  $K^{\delta_{(j, x^j)}} \subseteq K^{\delta_{(i, x^i)}}$  for all  $i \in V$ . Then, if  $y \in K^{\delta_{(j, x^j)}}$ , let  $\alpha^i$  be a control for  $i \in V$  that meets the demand of  $K^{\delta_{(i, x^i)}}$ . To prove  $y \in K^\mu$  we must exhibit a control that satisfies the requirements of such a set. Having a control  $\alpha$  taken as  $\alpha^i$  on the branches generated by each  $i \in V$ , we meet the conditions of  $K^\mu$ . Therefore,  $\max_{i \in V} v(t, \delta_{(i, x^i)}) = v(t, \delta_{(j, x^j)}) \leq v(t, \mu)$   $\square$

From this result, we can focus on the function  $\bar{v}$  defined on  $\mathcal{I} \times [0, T] \times \mathbb{R}^d$  by

$$\bar{v}_i(t, x) = v(t, \delta_{(i, x)})$$

for  $(i, t, x) \in \mathcal{I} \times [0, T] \times \mathbb{R}^d$ . We provide in the next sections a PDE characterisation of the function  $\bar{v}$ .

## 4.2 Dynamic programming equation

### 4.2.1 The equation on the parabolic interior

In a stochastic target problem, wishing to hit a given target with probability one, we must de-generate along certain directions. Moreover, we also need to control the uncertainty related the possible branching. This property enables the characterisation of the value function  $\bar{v}$  as a solution the following PDE

$$\min \left\{ -\partial_t \bar{v}_i(t, x) + F(x, \bar{v}_i(t, x), D\bar{v}_i(t, x), D_x^2 \bar{v}_i(t, x)) ; \bar{v}_i(t, x) - \sup_{0 \leq k < \bar{K}} \bar{v}_{ik}(t, x) \right\} = 0 \quad (4.28)$$

for  $(t, x) \in [0, T) \times \mathbb{R}^d$ , where

$$\begin{aligned} \bar{K} &= \sup \{k + 1 \in \mathbb{N} : p_k > 0\}, \\ F(\Theta) &= \sup \left\{ \lambda_Y(x, y, a) - \lambda(x, a)^\top p - \frac{1}{2} \text{Tr} \left( \sigma \sigma^\top(x, a) M \right) : a \in \mathcal{N}(x, p) \right\} \end{aligned}$$

for  $\Theta = (x, y, p, M) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}^d$ , and

$$\mathcal{N}(x, p) = \{a \in A : N^a(x, p) = 0\} \quad \text{and} \quad N^a(x, p) = \sigma_Y(x, a) - \sigma(x, a)^\top p$$

for  $x, p \in \mathbb{R}^d$ .

Since the control set  $A$  is not necessarily compact, the operator associated to this PDE may not be continuous. We therefore need to define a weak formulation of (4.28). For that, we introduce the relaxed semilimits of  $F$  given by

$$F^*(\Theta) = \limsup_{\varepsilon \rightarrow 0, \Theta' \rightarrow \Theta} F_\varepsilon(\Theta') \quad \text{and} \quad F_*(\Theta) = \liminf_{\varepsilon \rightarrow 0, \Theta' \rightarrow \Theta} F_\varepsilon(\Theta')$$

where

$$F_\varepsilon(\Theta) = \sup \left\{ \lambda_Y(x, y, a) - \lambda(x, a)^\top p - \frac{1}{2} \text{Tr} \left( \sigma \sigma^\top(x, a) M \right) : a \in \mathcal{N}_\varepsilon(x, p) \right\}$$

for  $\Theta = (x, y, p, M) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}^d$  and  $\varepsilon \geq 0$ , and

$$\mathcal{N}_\varepsilon(x, p) = \{a \in A : |N^a(x, p)| \leq \varepsilon\} \quad \text{and} \quad N^a(x, p) = \sigma_Y(x, a) - \sigma(x, a)^\top p$$

for  $x, p \in \mathbb{R}^d$ . Observe that  $(\mathcal{N}_\varepsilon)_{\varepsilon \geq 0}$  is non-decreasing so that

$$F_*(\Theta) = \liminf_{\Theta' \rightarrow \Theta} F_0(\Theta') \quad (4.29)$$

Since some  $\mathcal{N}_\varepsilon(x, p)$  may be empty, we shall use the standard convention  $\sup \emptyset = -\infty$  all over this paper. For ease of notations, we also write  $F\varphi(t, x)$  in place of  $F(x, \varphi(t, x), D\varphi(t, x), D_x^2\varphi(t, x))$  for a regular function  $\varphi$ . We similarly use the notations  $F^*\varphi$  and  $F_*\varphi$ .

As the value function may not be regular, we use the framework of discontinuous viscosity solutions. To this end, we define the lower- and upper-semicontinuous envelopes  $f_*$  and  $f^*$  of a locally bounded function  $f : [0, T] \times \mathbb{R}^d \times \mathcal{I} \rightarrow \mathbb{R}$  by

$$f_i^*(t, x) = \limsup_{\substack{(t', x') \rightarrow (t, x) \\ t' < T}} f_i(t', x') \quad \text{and} \quad f_{i,*}(t, x) = \liminf_{\substack{(t', x') \rightarrow (t, x) \\ t' < T}} f_i(t', x') \quad (4.30)$$

for  $(t, x, i) \in [0, T] \times \mathbb{R}^d \times \mathcal{I}$ . We are now able to provide the definition of a viscosity solution to (4.28).

**Definition 4.1.** *Let  $u : [0, T] \times \mathbb{R}^d \times \mathcal{I} \rightarrow \mathbb{R}$  be a locally bounded function.*

*(i)  $u$  is a viscosity supersolution to (4.28) if for any  $(t_0, x_0, i_0) \in [0, T] \times \mathbb{R}^d \times \mathcal{I}$  and any  $\varphi_i \in C^{1,2}([0, T] \times \mathbb{R}^d)$  for  $i \in \mathcal{I}$  and  $\bar{\varphi} \in C^0([0, T] \times \mathbb{R}^d)$  such that*

$$\begin{aligned} \sup_{i \in \mathcal{I}} |\varphi_i(t, x)| &\leq \bar{\varphi}(t, x), \quad \forall (t, x) \in [0, T] \times \mathbb{R}^d, \\ 0 &= (u_{i_0,*} - \varphi_{i_0})(t_0, x_0) = \min_{\mathcal{I} \times [0, T] \times \mathbb{R}^d} (u_{*,*} - \varphi). \end{aligned}$$

*we have*

$$\min \left\{ -\partial_t \varphi_{i_0}(t_0, x_0) + F^* \varphi_{i_0}(t_0, x_0); \left( \varphi_{i_0} - \sup_{0 \leq k < \bar{K}} \varphi_{i_0 k} \right)(t_0, x_0) \right\} \geq 0.$$

*(ii)  $u$  is a viscosity subsolution to (4.28) if for any  $(t_0, x_0, i_0) \in [0, T] \times \mathbb{R}^d \times \mathcal{I}$  and any  $\varphi_i \in C^{1,2}([0, T] \times \mathbb{R}^d)$  for  $i \in \mathcal{I}$  and  $\bar{\varphi} \in C^0([0, T] \times \mathbb{R}^d)$  such that*

$$\begin{aligned} \sup_{i \in \mathcal{I}} |\varphi_i(t, x)| &\leq \bar{\varphi}(t, x), \quad \forall (t, x) \in [0, T] \times \mathbb{R}^d, \\ 0 &= (u_{i_0}^* - \varphi_{i_0})(t_0, x_0) = \max_{\mathcal{I} \times [0, T] \times \mathbb{R}^d} (u_{*,*}^* - \varphi). \end{aligned}$$

*we have*

$$\min \left\{ -\partial_t \varphi_{i_0}(t_0, x_0) + F_* \varphi_{i_0}(t_0, x_0); \left( \varphi_{i_0} - \sup_{0 \leq k < \bar{K}} \varphi_{i_0 k} \right)(t_0, x_0) \right\} \leq 0.$$

*(iii)  $u$  is a viscosity solution to (4.28) if it is both a viscosity sub and supersolution to (4.28).*

We notice that the definition of viscosity solution is slightly different from the classical one as we impose a bound in the label particle  $i$  for test functions.

Following [4], we introduce the a continuity assumption on the kernel that is used to prove the subsolution property.

**Assumption A3.** *Let  $B$  be a subset of  $\mathbb{R}^d \times \mathbb{R}^d$  such that  $\mathcal{N}_0 \neq \emptyset$  on  $B$ . Then, for every  $\varepsilon > 0$ ,  $(x_0, p_0) \in \text{int}(B)$ , and  $a_0 \in \mathcal{N}_0(x_0, p_0)$ , there exists an open neighborhood  $B'$  of  $(x_0, p_0)$  and a locally Lipschitz map  $\hat{a}$  defined on  $B'$  such that  $|\hat{a}(x_0, p_0) - a_0| \leq \varepsilon$  and*

$$\hat{a}(x, p) \in \mathcal{N}_0(x, p) \quad \text{for all } (x, p) \in B' .$$

We are now able to state our result.

**Theorem 4.3.** *Suppose that  $\bar{v}$  is locally bounded on  $[0, T] \times \mathbb{R}^d \times \mathcal{I}$ .*

- (i) *Under Assumptions A1, the value function  $\bar{v}$  is a viscosity supersolution to (4.28)*
- (ii) *If in addition Assumption A3 holds,  $\bar{v}$  is a viscosity subsolution to (4.28)*

#### 4.2.2 Terminal condition

To get a complete characterisation of the function  $\bar{v}$ , we need to add a terminal equation to (4.28). By the definition of the stochastic target problem, we have

$$\bar{v}_i(T, x) = g_i(x) \tag{4.31}$$

for every  $(x, i) \in \mathbb{R}^d \times \mathcal{I}$ . The possible discontinuities of  $\bar{v}$  might imply that  $\bar{v}_*$  and  $\bar{v}^*$  do not agree with the boundary condition (4.31). To get the proper terminal condition, we introduce the set-valued map

$$\mathbf{N}(x, p) = \{r \in \mathbb{R}^m : r = N^a(x, p) \text{ for some } a \in A\}$$

together with the signed distance function from its complement set  $\mathbf{N}^c$  to the origin

$$\delta = \text{dist}(0, \mathbf{N}^c) - \text{dist}(0, \mathbf{N}) ,$$

where  $\text{dist}$  stands for the Euclidean distance. Then,

$$0 \in \text{int}\mathbf{N}(x, p) \Leftrightarrow \delta(x, p) > 0 . \tag{4.32}$$

For simplicity of notations, we will write  $\delta\varphi(x)$  for  $\delta(x, D\varphi(x))$  for a regular function  $\varphi$ . Then, the terminal condition takes the following form

$$\min \left\{ \bar{v}_i(T, x) - g_i(x) ; \delta\bar{v}_i(T, x) ; \left( \bar{v}_i - \sup_{0 \leq k < \bar{K}} \bar{v}_{ik} \right) (T, x) \right\} = 0 \tag{4.33}$$

for  $(x, i) \in \mathbb{R}^d \times \mathcal{I}$ .

We give the definition of a viscosity solution to (4.33). We recall that the definitions of the envelopes  $u^*$  and  $u_*$  of a locally bounded function  $u$  are given by (4.30).

**Definition 4.2.** Let  $u : [0, T] \times \mathbb{R}^d \times \mathcal{I} \rightarrow \mathbb{R}$  be a locally bounded function.

(i)  $u$  is a viscosity supersolution to (4.33) if for any  $(x_0, i_0) \in \mathbb{R}^d \times \mathcal{I}$  and any  $\varphi_i \in C^2(\mathbb{R}^d)$  for  $i \in \mathcal{I}$  and  $\bar{\varphi} \in C^0(\mathbb{R}^d)$  such that

$$\begin{aligned} \sup_{i \in \mathcal{I}} |\varphi_i(x)| &\leq \bar{\varphi}(x), \quad \forall x \in \mathbb{R}^d, \\ 0 &= u_{i_0}^*(T, x_0) - \varphi_{i_0}(x_0) = \min_{\mathcal{I} \times \mathbb{R}^d} (u^*(T, \cdot) - \varphi) \end{aligned}$$

we have

$$\min \left\{ \varphi_{i_0}(x) - g_{i_0}(x_0); \delta_* \varphi_{i_0}(x_0); \varphi_{i_0}(T, x_0) - \sup_{0 \leq k < \bar{K}} \varphi_{i_0 k}(T, x_0) \right\} \geq 0.$$

(ii)  $u$  is a viscosity subsolution solution to (4.33) if for any  $(x_0, i_0) \in \mathbb{R}^d \times \mathcal{I}$  and any  $\varphi_i \in C^2(\mathbb{R}^d)$  for  $i \in \mathcal{I}$  and  $\bar{\varphi} \in C^0(\mathbb{R}^d)$  such that

$$\begin{aligned} \sup_{i \in \mathcal{I}} |\varphi_i(x)| &\leq \bar{\varphi}(x), \quad \forall x \in \mathbb{R}^d, \\ 0 &= u_{i_0, *}(T, x_0) - \varphi_{i_0}(x_0) = \max_{\mathcal{I} \times \mathbb{R}^d} (u_{*, *}(T, \cdot) - \varphi) \end{aligned}$$

we have

$$\min \left\{ (\varphi_{i_0}(x) - g_i(x)) \mathbf{1}_{F^* \varphi_{i_0}(x) < \infty}; \delta^* \varphi_{i_0}(x); \varphi_{i_0}(T, x) - \sup_{0 \leq k < \bar{K}} \varphi_{i_0 k}(T, x) \right\} \leq 0.$$

(iii)  $u$  is a viscosity solution to (4.33) if it is both a viscosity sub and supersolution to (4.33).

The terminal viscosity property is stated as follows.

**Theorem 4.4.** Suppose that  $\bar{v}$  is locally bounded on  $[0, T] \times \mathbb{R}^d \times \mathcal{I}$ .

(i) Under Assumptions A1 and A2,  $\bar{v}$  is a viscosity supersolution to (4.33).

(ii) If in addition Assumption A3 holds,  $\bar{v}$  is a viscosity subsolution to (4.33).

## 4.3 Viscosity properties on $[0, T] \times \mathbb{R}^d \times \mathcal{I}$

### 4.3.1 Viscosity supersolution property

Fix  $(i_0, t_0, x_0) \in \mathcal{I} \times [0, T] \times \mathbb{R}^d$  and let  $\varphi \in C^0([0, T] \times \mathbb{R}^d)$  and  $\varphi_i \in C^{1,2}([0, T] \times \mathbb{R}^d)$  for  $i \in \mathcal{I}$  be such that

$$\sup_i |\varphi_i| \leq \varphi \tag{4.34}$$

and

$$0 = (\bar{v}_{i_0, *} - \varphi_{i_0})(t_0, x_0) = \min_{(i, t, x) \in \mathcal{I} \times [0, T] \times \mathbb{R}^d} (\bar{v}_{i, *} - \varphi_i)(t, x). \tag{4.35}$$

Without loss of generality we can assume this minimum to be strict in  $(t, x)$  once fixed  $i_0$ .



**Step 1.** We first prove that  $\varphi_{i_0}(t_0, x_0) - \sup_{0 \leq \ell \leq k-1} \varphi_{i_0 \ell}(t_0, x_0) \geq 0$  for any  $k$  such that  $p_k > 0$ . Let  $(t_n, x_n)$  be a sequence in  $[0, T] \times \mathbb{R}^d$  such that

$$(t_n, x_n) \rightarrow (t_0, x_0) \quad \text{and} \quad \bar{v}_{i_0}(t_n, x_n) \rightarrow \bar{v}_{i_0, *}(t_0, x_0) \quad \text{as } n \rightarrow \infty.$$

Set  $y_0 := \varphi_{i_0}(t_0, x_0)$ ,  $\hat{x}_0 := (x_0, y_0)$ ,  $y_n := \bar{v}_{i_0}(t_n, x_n) + 1/n$  and  $\hat{x}_n := (x_n, y_n)$ . Define the stopping time  $\theta_n = \inf\{s \geq t_n : Q^{i_0}((t_n, s] \times \mathbb{N}) \geq 1\}$  and the random variable  $k_n$  such that  $Q^{i_0}((t_n, \theta_n] \times \{k_n\}) = 1$ . From Theorem 3.2, the continuity of the trajectories and since  $y_n > \bar{v}_{i_0}(t_n, x_n)$  there exists  $\alpha^n \in \mathcal{A}$  such that

$$Y_{\theta_n^-}^{t_n, \delta_{(i_0, \hat{x}_n)}, \alpha^n, i_0} \geq \max_{0 \leq \ell \leq k_n - 1} \bar{v}_{i_0 \ell} \left( \theta_n, X_{\theta_n^-}^{t_n, \delta_{(i_0, x_n)}, \alpha^n, i_0} \right) \geq \max_{0 \leq \ell \leq k_n - 1} \varphi_{i_0 \ell} \left( \theta_n, X_{\theta_n}^{t_n, \delta_{(i_0, x_n)}, \alpha^n, i_0} \right).$$

on  $\{\theta_n \leq T\}$ . To alleviate the notation, we shall denote  $X_t^{n, i} := X_t^{t_n, \delta_{(i_0, x_n)}, \alpha^n, i}$  and  $Y_t^{n, i} := Y_t^{t_n, \delta_{(i_0, x_n)}, y_n, \alpha^n, i}$  for  $n \geq 1$  and  $t \in [t_n, T]$ . Therefore, we get

$$\gamma \sum_{k=0}^{+\infty} p_k \int_0^T \mathbb{E} \left[ \mathbb{1}_{s \leq \theta_n \leq T} \mathbb{1}_{Y_s^{n, i_0} < \max_{0 \leq \ell \leq k-1} \varphi_{i_0 \ell}(s, X_s^{n, i_0})} \right] ds = 0,$$

which means

$$\int_0^T \mathbb{E} \left[ \mathbb{1}_{s \leq \theta_n \leq T} \mathbb{1}_{Y_s^{n, i_0} < \max_{0 \leq \ell \leq k-1} \varphi_{i_0 \ell}(s, X_s^{n, i_0})} \right] ds = 0$$

for all  $k \geq 1$  such that  $p_k > 0$ . We therefore get

$$\mathbb{E} \left[ \mathbb{1}_{s \leq \theta_n \leq T} \mathbb{1}_{Y_s^{n, i_0} < \max_{0 \leq \ell \leq k-1} \varphi_{i_0 \ell}(s, X_s^{n, i_0})} \right] = 0 \quad (4.36)$$

for Lebesgue almost all  $s \in [t_n, T]$ . Since the process  $Y^{n, i_0} - \max_{0 \leq \ell \leq k-1} \varphi_{i_0 \ell}(\cdot, X^{n, i_0})$  is continuous and  $\mathbb{P}(\theta_n \in [t_n, T]) > 0$ , Fatou's Lemma applied to a sequence  $(s_k)_k$  converging to  $t_n$  and satisfying (4.36) gives

$$y_n \geq \max_{0 \leq \ell \leq k-1} \varphi_{i_0 \ell}(t_n, x_n)$$

for all  $k \geq 1$  such that  $p_k > 0$ . Sending  $n$  to infinity gives the result.

**Step 2.** We now prove that

$$-\frac{\partial \varphi_{i_0}}{\partial t}(t_0, x_0) + F^* \varphi_{i_0}(t_0, x_0) \geq 0$$

Assume to the contrary that  $(-\partial_t \varphi_{i_0} + F^* \varphi_{i_0})(t_0, x_0) = -2\eta$  for some  $\eta > 0$ , and let us work towards a contradiction. By definition of  $F^*$ , we may find  $\varepsilon \in (0, T - t_0)$ , such that

$$-\partial_t \varphi_{i_0}(t, x) + \lambda_Y(x, y, a) - L^a \varphi_{i_0}(t, x) \leq -\eta \quad \text{for all } a \in \mathcal{N}_\varepsilon(x, D\varphi_{i_0}(t, x)) \quad (4.37)$$

and  $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$  such that  $(t, x) \in B_\varepsilon(t_0, x_0)$  and  $|y - \varphi_{i_0}(t, x)| \leq \varepsilon$ ,

where  $B_\varepsilon(t_0, x_0)$  denotes the ball of radius  $\varepsilon$  around  $(t_0, x_0)$ . Let  $\partial_p B_\varepsilon(t_0, x_0) = \{t_0 + \varepsilon\} \times \text{cl}(B_\varepsilon(t_0, x_0)) \cup [t_0, t_0 + \varepsilon) \times \partial B_\varepsilon(x_0)$  denote the parabolic boundary of  $B_\varepsilon(t_0, x_0)$  and observe that

$$\zeta = \min_{\partial_p B_\varepsilon(t_0, x_0)} (\bar{v}_{i_0, *} - \varphi_{i_0}) > 0 \quad (4.38)$$

since  $(t_0, x_0)$  is a strict minimizer of  $\bar{v}_{i_0, *} - \varphi_{i_0}$  on  $[0, T] \times \mathbb{R}^d$ .

**Step 3.** We now show that (4.37) and (4.38) lead to a contradiction to (3.25). Let  $(t_n, x_n)$  in  $[0, T] \times \mathbb{R}^d$  such that

$$(t_n, x_n) \rightarrow (t_0, x_0) \text{ and } \bar{v}_{i_0}(t_n, x_n) \rightarrow \bar{v}_{i_0, *}(t_0, x_0) \text{ as } n \rightarrow \infty.$$

We then set  $y_0 := \varphi_{i_0}(t_0, x_0)$ ,  $\hat{x}_0 := (x_0, y_0)$ ,  $y_n := \bar{v}_{i_0}(t_n, x_n) + 1/n$ ,  $\hat{x}_n := (x_n, y_n)$ ,  $\beta_n := y_n - \varphi_{i_0}(t_n, x_n)$  and notice that

$$\beta_n \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (4.39)$$

From the definition of the value function and the fact that  $y_n > \bar{v}_{i_0}(t_n, x_n)$  for each  $n \geq 1$ , there exists some  $\alpha^n$  in  $\mathcal{A}$  such that  $Y_T^{t_n, \delta(i_0, x_n), y_n, \alpha^n, i} \geq g_i \left( X_T^{t_n, \delta(i_0, x_n), \alpha^n, i} \right)$  for all  $i \in \mathcal{V}_T^{t_n, \delta(i_0, x_n), \alpha^n}$ . To alleviate the notation, we shall denote

$$X_t^{n, i} := X_t^{t_n, \delta(i_0, x_n), \alpha^n, i}, \quad Y_t^{n, i} := Y_t^{t_n, \delta(i_0, x_n), y_n, \alpha^n, i} \text{ and } \mathcal{V}_t^n := \mathcal{V}_t^{t_n, \delta(i_0, x_n), \alpha^n}$$

for  $n \geq 1$  and  $t \in [t_n, T]$ . Define the following stopping times

$$\begin{aligned} \tau_n &:= \inf\{s \geq t_n : \exists i \in \mathcal{V}_s^n, (s, X_s^{n, i}) \notin B_\varepsilon(t_0, x_0)\}, \\ \tau_n^\varepsilon &:= \inf\{s \geq t_n : \exists i \in \mathcal{V}_s^n, |Y_s^{n, i} - \varphi_i(s, X_s^{n, i})| \geq \varepsilon\}, \\ \tau_n^r &:= \inf\{s \geq t_n : Q^{i_0}((t_n, s] \times \mathbb{N}) = 1\}, \\ \theta_n &:= \tau_n \wedge \tau_n^\varepsilon \wedge \tau_n^r. \end{aligned}$$

We also set

$$\begin{aligned} A_n &= \left\{s \in [t_n, \theta_n) : -\partial_t \varphi_{i_0}(s, X_s^{n, i_0}) + \lambda_Y(X_s^{n, i_0}, Y_s^{n, i_0}, \alpha_{i_0}^n) - L^{\alpha_{i_0}^n} \varphi_{i_0}(s, X_s^{n, i_0}) > -\eta\right\}, \quad (4.40) \\ \psi_s^n &= N^{\alpha_{i_0}^n}(X_s^{n, i_0}, D\varphi_{i_0}(s, X_s^{n, i_0})). \end{aligned}$$

We notice that (4.37) implies

$$|\psi_s^n| > \varepsilon \text{ for } s \in A_n. \quad (4.41)$$

It follows from Theorem 3.2 that

$$Y_{t \wedge \theta_n}^{n, i} \geq \bar{v}_i(t \wedge \theta_n, X_{t \wedge \theta_n}^{n, i}) \quad \forall i \in \mathcal{V}_{t \wedge \theta_n}^n, t \in [t_n, T].$$

and since  $\bar{v}_i \geq \bar{v}_{i, *}$   $\geq \varphi_i$

$$Y_{\theta_n \wedge t}^{n, i} \geq \varphi_i(\theta_n \wedge t, X_{\theta_n \wedge t}^{n, i}) \quad \forall i \in \mathcal{V}_{\theta_n}^n. \quad (4.42)$$

Using the definition of  $\zeta$  in (4.38) and  $\theta_n$ , and the continuity of the trajectories, we get

$$\begin{aligned} Y_{t \wedge \theta_n}^{n, i_0} &\geq \varphi_{i_0}(t \wedge \theta_n, X_{t \wedge \theta_n}^{n, i_0}) + (\zeta \mathbb{1}_{\{\theta_n = \tau_n\}} + \varepsilon \mathbb{1}_{\{\tau_n^\varepsilon = \theta_n\}} \cap \{\theta_n < \tau_n\}) \mathbb{1}_{\{\theta_n \leq t\}} \cap \{\theta_n < \tau_n^r\} \\ &\geq \varphi_{i_0}(t \wedge \theta_n, X_{t \wedge \theta_n}^{n, i_0}) + \zeta \wedge \varepsilon \mathbb{1}_{\{\theta_n \leq t\}} \cap \{\theta_n < \tau_n^r\}. \end{aligned}$$

Therefore, from (4.42) and the previous inequality, we have

$$-\zeta \wedge \varepsilon \mathbb{1}_{\{\theta_n > t\} \cup \{\theta_n = \tau_n^r\}} \leq -\zeta \wedge \varepsilon + Y_{t \wedge \theta_n}^{n, i} - \varphi_i(t \wedge \theta_n, X_{t \wedge \theta_n}^{n, i})$$

Applying the dynamics (2.11) of  $\hat{Z}^{t_n, \delta(i_0, \hat{x}_n), \alpha^n}$  to the function  $(t, x, y, i) \mapsto y - \varphi_{i_0}(t, x)$ , it follows from the definition of  $\psi_n$  and  $\theta_n$ , and (4.40) that

$$\begin{aligned}
-\zeta \wedge \varepsilon \mathbb{1}_{\{\theta_n > t\} \cup \{\theta_n = \tau_n^r\}} &\leq \beta_n - \zeta \wedge \varepsilon + \int_{t_n}^{t \wedge \theta_n} \psi_s^{n \top} dB_u^{i_0} \\
&+ \int_{t_n}^{t \wedge \theta_n} \left[ -\partial_t \varphi_{i_0}(u, X_u^{n, i_0}) + \lambda_Y(X_u^{n, i_0}, Y_u^{n, i_0}, \alpha_u^{n, i_0}) - L^{\alpha_u^{n, i_0}} \varphi_{i_0}(u, X_u^{n, i_0}) \right] du \\
&+ \int_{(t_n, \theta_n \wedge t]} \sum_{k \geq 0} \left( (k-1) Y_u^{n, i_0} - \left( \sum_{\ell=0}^{k-1} \varphi_{i_0 \ell} - \varphi_{i_0} \right) (u, X_u^{n, i_0}) \right) Q^{i_0}(dudk) \\
&\leq \beta_n - \zeta \wedge \varepsilon + \int_{t_n}^{t \wedge \theta_n} \psi_s^{n \top} dB_u^{i_0} \\
&+ \int_{t_n}^{t \wedge \theta_n} \left[ -\partial_t \varphi_{i_0}(u, X_u^{n, i_0}) + \lambda_Y(X_u^{n, i_0}, Y_u^{n, i_0}, \alpha_u^{n, i_0}) - L^{\alpha_u^{n, i_0}} \varphi_{i_0}(u, X_u^{n, i_0}) \right] \mathbb{1}_{A_n}(u) du \\
&+ \int_{(t_n, \theta_n \wedge t]} \sum_{k \geq 0} \left( (k-1) Y_u^{n, i_0} - \left( \sum_{\ell=0}^{k-1} \varphi_{i_0 \ell} - \varphi_{i_0} \right) (u, X_u^{n, i_0}) \right) Q^{i_0}(dudk) .
\end{aligned}$$

We then get

$$-\zeta \wedge \varepsilon \mathbb{1}_{\{\theta_n > t\} \cup \{\theta_n = \tau_n^r\}} \leq M_{t \wedge \theta_n}^{B, n} + M_{t \wedge \theta_n}^{Q, n}, \quad (4.43)$$

where

$$\begin{aligned}
M_s^{B, n} &= \beta_n - \zeta \wedge \varepsilon + \int_{t_n}^s b_u^n du + \int_{t_n}^s \psi_s^{n \top} dB_u^{i_0}, \\
b_s^n &= \left[ -\partial_t \varphi(s, X_s^{n, i_0}) + \lambda_Y(X_s^{n, i_0}, Y_s^{n, i_0}, \alpha_s^{n, i_0}) - L^{\alpha_s^{n, i_0}} \varphi(s, X_s^{n, i_0}) \right] \mathbb{1}_{A_n}(s) + \\
&+ \sum_{k \geq 0} \left( (k-1) Y_s^{n, i_0} - \left( \sum_{\ell=0}^{k-1} \varphi_{i_0 \ell} - \varphi_{i_0} \right) (s, X_s^{n, i_0}) \right) \gamma p_k, \\
M_s^{Q, n} &= \int_{(t_n, s]} \sum_{k \geq 0} \left( (k-1) Y_u^{n, i_0} - \left( \sum_{\ell=0}^{k-1} \varphi_{i_0 \ell} - \varphi_{i_0} \right) (u, X_u^{n, i_0}) \right) (Q^{i_0}(dudk) - \gamma p_k du) .
\end{aligned}$$

for  $s \in [t_n, T]$ . From to Step 1, the definition of  $\theta_n$ , the domination condition (4.34) and Assumption A1,  $M_{\cdot \wedge \theta_n}^{Q, n}$  is a pure jump martingale. Let  $L^n$  be the exponential local martingale defined by  $L_{t_n}^n = 1$  and

$$dL_s^n = -L_s^n b_s^n |\psi_s^n|^{-2} \psi_s^n \mathbb{1}_{A_n}(s)^\top dB_s^{i_0}$$

for  $s \in [t_n, T]$ .  $L^n$  is well defined by (4.41), Assumption A1 and the definition of the set of admissible controls  $\mathcal{A}$ . Moreover, From the definition of  $\theta_n$ ,  $L_{\cdot \wedge \theta_n}^n$  is a martingale. From Girsanov Theorem for jump diffusion processes (see e.g. Theorem 1.35 in [22]) and the definition of  $\theta_n$ , we get that  $L_{\cdot \wedge \theta_n}^n M_{\cdot \wedge \theta_n}^{B, n} + L_{\cdot \wedge \theta_n}^n M_{\cdot \wedge \theta_n}^{Q, n}$  is a martingale. It follows from (4.43) that

$$\begin{aligned}
-\zeta \wedge \varepsilon \mathbb{E}[\mathbb{1}_{\{\theta_n = \tau_n^r\}} L_{\theta_n}^n] &\leq \mathbb{E} \left[ L_{\theta_n}^n M_{\theta_n}^{B, n} + L_{\theta_n}^n M_{\theta_n}^{Q, n} \right] \\
&\leq L_{t_n}^n M_{t_n}^{B, n} + L_{t_n}^n M_{t_n}^{Q, n} = \beta_n - \zeta \wedge \varepsilon .
\end{aligned}$$

Since  $L_{\wedge \theta_n}^n$  is a martingale and  $\theta_n$  is a stopping time bounded by  $\varepsilon$ , we have  $\mathbb{E}[L_{\theta_n}^n] = L_{t_n}^n = 1$ . Therefore, the previous inequality becomes

$$\zeta \wedge \varepsilon \mathbb{E} [\mathbf{1}_{\{\theta_n < \tau_n^r\}} L_{\theta_n}^n] \leq \beta_n . \quad (4.44)$$

We next define the probability measure on  $\mathcal{F}_T$  by the Radon-Nikodym derivative

$$\left. \frac{d\mathbb{P}^n}{d\mathbb{P}} \right|_{\mathcal{F}_T} = L_{\theta_n}^n$$

and denote by  $\mathbb{E}^n$  the expectation under  $\mathbb{P}^n$ . Using Girsanov Theorem, we notice that  $\tau_n^r$  has the same law under  $\mathbb{P}$  and  $\mathbb{P}^n$ . In particular, we have

$$\mathbb{E} [\mathbf{1}_{\{\theta_n < \tau_n^r\}} L_{\theta_n}^n] \geq \mathbb{E} [\mathbf{1}_{\{\tau_n^r > \varepsilon\}} L_{\theta_n}^n] = \mathbb{E}^n [\mathbf{1}_{\{\tau_n^r > \varepsilon\}}] = \mathbb{E} [\mathbf{1}_{\{\tau_n^r > \varepsilon\}}] = \exp(-\varepsilon\gamma) .$$

Comparing with (4.44), we have

$$0 \leq \beta_n - \zeta \wedge \varepsilon \exp(-\varepsilon\gamma) ,$$

which contradicts (4.39) for  $n$  large enough.

### 4.3.2 Viscosity subsolution property

**Step 1.** Let  $\varphi \in C^0([0, T] \times \mathbb{R}^d)$ ,  $\varphi_i \in C^{1,2}([0, T] \times \mathbb{R}^d)$  for  $i \in \mathcal{I}$  and  $(t_0, x_0, i_0) \in [0, T] \times \mathbb{R}^d \times \mathcal{I}$  such that

$$\sup_i |\varphi_i| \leq \varphi$$

and

$$0 = (\bar{v}_{i_0}^* - \varphi_{i_0})(t_0, x_0) = \max_{(t, x, i) \in [0, T] \times \mathbb{R}^d \times \mathcal{I}} (\bar{v}_i^* - \varphi_i)(t, x) . \quad (4.45)$$

Without loss of generality we can assume that the maximum is strict in  $(t, x)$  once fixed  $i_0$ . We then argue by contradiction and assume that

$$4\eta = \min \left\{ (-\partial_t \varphi_{i_0} + F_* \varphi_{i_0})(t_0, x_0) ; \left( \varphi_{i_0} - \sup_{0 \leq k < \bar{K}} \varphi_{i_0 k} \right)(t_0, x_0) \right\} > 0 . \quad (4.46)$$

By (4.29), Assumption A3 and (4.46) we may find  $\varepsilon > 0$  such that

$$\rho(t, x, y) = -\partial_t \varphi_{i_0}(t, x) + \lambda_Y(x, y, \hat{a}(x, D\varphi_{i_0}(t, x))) - L^{\hat{a}(x, D\varphi_{i_0}(t, x))} \varphi_{i_0}(t, x) \geq \eta , \quad (4.47)$$

$$\left( \varphi_{i_0} - \sup_{0 \leq k < \bar{K}} \varphi_{i_0 k} \right)(t, x) \geq \eta \quad (4.48)$$

for all  $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$  such that  $(t, x) \in B_\varepsilon(t_0, x_0)$  and  $|y - \varphi_{i_0}(t, x)| \leq \varepsilon$ , where  $\hat{a}$  is a locally Lipschitz map satisfying

$$\hat{a}(x, D\varphi_{i_0}(t, x)) \in \mathcal{N}_0(x, D\varphi_{i_0}(t, x)) \text{ on } B_\varepsilon(t_0, x_0) . \quad (4.49)$$

Observe that, since  $(t_0, x_0)$  is a strict maximizer, we have

$$-\zeta = \max_{\partial_p B_\varepsilon(t_0, x_0)} (\bar{v}_{i_0}^* - \varphi_{i_0})(t, x) < 0 , \quad (4.50)$$

where  $\partial_p B_\varepsilon(t_0, x_0) = \{t_0 + \varepsilon\} \times \text{cl}(B_\varepsilon(t_0, x_0)) \cup [t_0, t_0 + \varepsilon) \times \partial B_\varepsilon(t_0, x_0)$  denotes the parabolic boundary of  $B_\varepsilon(t_0, x_0)$ .

**Step 2.** We now show that (4.47), (4.48), (4.49) and (4.50) lead to a contradiction to (3.25). Let  $(t_n, x_n)_{n \geq 1}$  be a sequence such that

$$(t_n, x_n) \rightarrow (t_0, x_0) \quad \text{and} \quad \bar{v}_{i_0}(t_n, x_n) \rightarrow \bar{v}_{i_0}^*(t_0, x_0) \quad \text{as} \quad n \rightarrow +\infty .$$

Set  $y_0 := \varphi_{i_0}(t_0, x_0)$ ,  $\hat{x}_0 := (x_0, y_0)$  and  $y_n := \bar{v}_{i_0}(t_n, x_n) - n^{-1}$ ,  $\hat{x}_n := (x_n, y_n)$  for  $n \geq 1$  and notice that

$$\beta_n := y_n - \varphi_{i_0}(t_n, x_n) \xrightarrow{n \rightarrow +\infty} 0 . \quad (4.51)$$

Define the following stopping times

$$\begin{aligned} \tau_n &:= \inf\{s \geq t_n : \exists i \in \mathcal{V}_s^n, (s, X_s^{n,i}) \notin B_\varepsilon(t_0, x_0)\} , \\ \tau_n^\varepsilon &:= \inf\{s \geq t_n : \exists i \in \mathcal{V}_s^n, |Y_s^{n,i} - \varphi_i(s, X_s^{n,i})| \geq \varepsilon\} , \\ \tau_n^r &:= \inf\{s \geq t_n : Q^{i_0}((t_n, s] \times \mathbb{N}) \geq 1\} , \\ \theta_n &:= \tau_n \wedge \tau_n^\varepsilon \wedge \tau_n^r . \end{aligned}$$

To alleviate the notations, we shall write

$$\begin{aligned} X^{n,i} &:= X^{t_n, \delta_{(i_0, x_n)}, \alpha^{n,i}} , \quad Y^{n,i} := Y^{t_n, \delta_{(i_0, x_n)}, y_n, \alpha^{n,i}} , \quad \hat{X}^{n,i} = (X^{n,i}, Y^{n,i}) \\ \hat{Z}^n &:= \hat{Z}^{t_n, \delta_{(i_0, \hat{x}_n)}, \hat{\alpha}^n} \quad \text{and} \quad \mathcal{V}^n = \mathcal{V}^{t_n, \delta_{(i_0, x_n)}, \hat{\alpha}^n} \end{aligned}$$

where  $\hat{\alpha}^n$  is the feedback control process given by  $\hat{\alpha}^{n,i} = \hat{\alpha}(X^{n,i}, D\varphi_{i_0}(\cdot, X^{n,i}))$  defined on  $[t_n, \theta_n)$  for  $n \geq 1$ . Since  $\hat{\alpha}$  is locally Lipschitz, this solution is well-defined. Since  $\bar{v}_i \leq \bar{v}_i^* \leq \varphi_i$ , we then deduce from (4.50) and the definition of  $\theta_n$  that on  $\{\theta_n < \tau_n^r\}$  we have

$$\begin{aligned} Y_{\theta_n}^{n,i_0} - \bar{v}_{i_0}(\theta_n, X_{\theta_n}^{n,i_0}) &\geq \mathbb{1}_{\{\theta_n = \tau_n^\varepsilon\}} \left( Y_{\theta_n}^{n,i_0} - \varphi_{i_0}(\theta_n, X_{\theta_n}^{n,i_0}) \right) \\ &\quad + \mathbb{1}_{\{\theta_n = \tau_n\}} \left( Y_{\theta_n}^{n,i_0} - \bar{v}_{i_0}^*(\theta_n, X_{\theta_n}^{n,i_0}) \right) \\ &= \varepsilon \mathbb{1}_{\{\theta_n = \tau_n^\varepsilon\}} + \mathbb{1}_{\{\theta_n = \tau_n < \tau_n^\varepsilon\}} \left( Y_{\theta_n}^{n,i_0} - \bar{v}_{i_0}^*(\theta_n, X_{\theta_n}^{n,i_0}) \right) \\ &\geq \varepsilon \mathbb{1}_{\{\theta_n = \tau_n^\varepsilon\}} + \mathbb{1}_{\{\theta_n = \tau_n < \tau_n^\varepsilon\}} \left( Y_{\theta_n}^{n,i_0} + \zeta - \varphi_{i_0}(\theta_n, X_{\theta_n}^{n,i_0}) \right) \\ &\geq \varepsilon \wedge \zeta + \mathbb{1}_{\{\theta_n = \tau_n < \tau_n^\varepsilon\}} \left( Y_{\theta_n}^{n,i_0} - \varphi_{i_0}(\theta_n, X_{\theta_n}^{n,i_0}) \right) . \end{aligned}$$

Secondly, on  $\{\theta_n = \tau_n^r\}$ , using the continuity of the trajectories of the particles  $Y_{\theta_n}^{i_0\ell} = Y_{\theta_n}^{i_0}$  and  $X_{\theta_n}^{i_0\ell} = X_{\theta_n}^{i_0}$  for all  $i_0\ell \in \mathcal{V}_{\theta_n}^n$ , we have

$$Y_{\theta_n}^{n,i_0\ell} - \varphi_{i_0\ell}(\theta_n, X_{\theta_n}^{n,i_0\ell}) = Y_{\tau_n^r}^{n,i_0} - \varphi_{i_0}(\tau_n^r, X_{\tau_n^r}^{n,i_0}) + \varphi_{i_0}(\tau_n^r, X_{\tau_n^r}^{n,i_0}) - \varphi_{i_0\ell}(\tau_n^r, X_{\tau_n^r}^{n,i_0}) ,$$

and from (4.48),

$$Y_{\theta_n}^{n,i_0\ell} - \varphi_{i_0\ell}(\theta_n, X_{\theta_n}^{n,i_0\ell}) \geq Y_{\theta_n}^{n,i_0} - \varphi_{i_0}(\theta_n, X_{\theta_n}^{n,i_0}) + \eta , \quad (4.52)$$

for all  $i_0\ell \in \mathcal{V}_{\theta_n}^n$ .

From (4.47) and (4.49), we get by Itô's formula

$$Y_{\theta_n}^{n,i} - \bar{v}_i(\theta_n, X_{\theta_n}^{n,i}) \geq \varepsilon \wedge \zeta \wedge \eta + \beta_n \quad \forall i \in \mathcal{V}_{\theta_n}^n .$$

Since  $y_n = \bar{v}_{i_0}(t_n, x_n) - n^{-1} < \bar{v}_{i_0}(t_n, x_n)$ , this is in contradiction with the dynamic programming principle (3.25) for sufficiently large  $n$  by (4.51).

## 4.4 Viscosity properties on $\{T\} \times \mathbb{R}^d \times \mathcal{I}$

### 4.4.1 Viscosity supersolution

Let  $(x_0, i_0) \in \mathbb{R}^d \times \mathcal{I}$  and  $\varphi_i \in C^2(\mathbb{R}^d)$  for  $i \in \mathcal{I}$  satisfying

$$0 = \bar{v}_{i_0,*}(T, x_0) - \varphi_{i_0}(x_0) = \min_{\mathcal{I} \times \mathbb{R}^d} (\bar{v}_{*,*}(T, \cdot) - \varphi.) .$$

Without loss of generalities we can take this minimum to be strict in  $x$  once fixed  $i_0$ .

**Step 1.** From the convention  $\sup \emptyset := -\infty$  and since  $\bar{v}$  is a viscosity supersolution for (4.28) on  $[0, T) \times \mathbb{R}^d \times \mathcal{I}$ , we have

$$\delta^* \bar{v}_{*,*} \geq 0 \text{ on } [0, T) \times \mathbb{R}^d \times \mathcal{I}$$

in the viscosity sense. From the upper-semicontinuity of  $\delta^*$ , we can then deduce by a standard argument (see e.g. proof of Lemma 5.2 in [25]) that  $\delta^* \varphi(x_0) \geq 0$ .

**Step 2.** We now prove

$$\varphi_{i_0}(x_0) - \sup_{0 \leq k < \bar{K}} \varphi_{i_0 k}(x_0) \geq 0 .$$

From the definition of  $\bar{v}_{*,*}$ , there exists a sequence  $(s_n, \xi_n)_{n \geq 1}$  converging to  $(T, x_0)$  such that  $s_n < T$  for  $n \geq 1$  and

$$\lim_{n \rightarrow \infty} \bar{v}_{i_0,*}(s_n, \xi_n) = \bar{v}_{i_0,*}(T, x_0) .$$

For  $n \geq 1$ , consider the auxiliary test function

$$\varphi_{n,i}(t, x) := \varphi_i(x) - \frac{1}{2}|x - x_0|^2 + \frac{T - t}{(T - s_n)^2} \quad (t, x, i) \in [0, T] \times \mathbb{R}^d \times \mathcal{I} .$$

Let  $B_1(x_0)$  be the unit open ball in  $\mathbb{R}^d$  centered at  $x_0$ . Choose  $(t_n, x_n) \in [s_n, T] \times \bar{B}_1(x_0)$ , which minimizes the difference  $\bar{v}_{i_0,*} - \varphi_{n,i_0}$  on  $[s_n, T] \times \bar{B}_1(x_0)$ .

We claim that, for  $n$  large enough  $t_n < T$ , and  $x_n$  converges to  $x_0$ . Indeed, for sufficiently large  $n$  we have

$$(\bar{v}_{i_0,*} - \varphi_{n,i_0})(s_n, x_0) \leq -\frac{1}{(T - s_n)} < 0 .$$

On the other hand, for any  $x \in \bar{B}_1(x_0)$

$$(\bar{v}_{i_0,*} - \varphi_{n,i_0})(T, x) = \bar{v}_{i_0,*}(T, x) - \varphi_{i_0}(x) + \frac{1}{2}|x - x_0|^2 \geq \bar{v}_{i_0,*}(T, x) - \varphi_{i_0}(x) \geq 0 .$$

Comparing the two inequalities leads us to conclude that  $t_n < T$  for large  $n$ . Let  $x^*$  be an adherence value of the sequence  $(x_n)_{n \geq 1}$ . Since  $t_n \geq s_n$  and  $(t_n, x_n)$  minimizes the difference  $(\bar{v}_{i_0,*} - \varphi_{n,i_0})$ , we have

$$\begin{aligned} & (\bar{v}_{i_0,*}(T, \cdot) - \varphi_{i_0})(x^*) - (\bar{v}_{i_0,*}(T, \cdot) - \varphi_{i_0})(x_0) \leq \\ & \liminf_{n \rightarrow \infty} (\bar{v}_{i_0,*} - \varphi_{n,i_0})(t_n, x_n) - (\bar{v}_{i_0,*} - \varphi_{n,i_0})(s_n, \xi_n) - \frac{1}{2}|x_n - x_0|^2 \leq \\ & \limsup_{n \rightarrow \infty} (\bar{v}_{i_0,*} - \varphi_{n,i_0})(t_n, x_n) - (\bar{v}_{i_0,*} - \varphi_{n,i_0})(s_n, \xi_n) - \frac{1}{2}|x_n - x_0|^2 \leq \\ & \qquad \qquad \qquad -\frac{1}{2}|x^* - x_0|^2 . \end{aligned}$$

Since  $x_0$  minimizes the difference  $\bar{v}_{i_0,*}(T, \cdot) - \varphi_{i_0}$  we have

$$0 \leq (\bar{v}_{i_0,*}(T, \cdot) - \varphi_{i_0})(x^*) - (\bar{v}_{i_0,*}(T, \cdot) - \varphi_{i_0})(x_0) \leq -\frac{1}{2}|x^* - x_0|^2.$$

Hence  $x^* = x_0$  and  $(x_n)_{n \geq 1}$  converges to  $x_0$ .

We now use the viscosity supersolution property of  $\bar{v}$  on  $[0, T] \times \mathbb{R}^d \times \mathcal{I}$  with the test function  $\tilde{\varphi}_n = \varphi_n + \bar{v}_{*,i_0}(t_n, x_n) - \varphi_{n,i_0}(t_n, x_n)$  and we have

$$\tilde{\varphi}_{n,i_0}(t_n, x_n) - \sup_{0 \leq k < \bar{K}} \tilde{\varphi}_{n,i_0k}(t_n, x_n) \geq 0 \quad (4.53)$$

for all  $n \geq 1$ . We clearly have

$$\varphi_{i_0}(x_n) - \sup_{0 \leq k < \bar{K}} \varphi_{i_0k}(x_n) = \tilde{\varphi}_{n,i_0}(t_n, x_n) - \sup_{0 \leq k < \bar{K}} \tilde{\varphi}_{n,i_0k}(t_n, x_n).$$

Since  $x_n$  converges to  $x_0$ , we get by sending  $n$  to infinity that  $\varphi_{i_0}(x_0) - \sup_{0 \leq k < \bar{K}} \varphi_{i_0k}(x_0) \geq 0$ .

**Step 3.** We now prove the last assertion. Assume that

$$F^* \varphi_{i_0}(x_0) < \infty \quad \text{and} \quad \varphi_{i_0}(x_0) = \bar{v}_{i_0,*}(T, x_0) < g_{i_0}$$

and let us work towards a contradiction. Since  $\bar{v}(T, \cdot) = g$  by the definition of the problem, there is a constant  $\eta > 0$  such that

$$\varphi_{i_0} - \bar{v}_{i_0}(T, \cdot) = \varphi_{i_0} - g_{i_0} \leq -\eta \quad \text{on } B_\varepsilon(x_0)$$

for some  $\varepsilon > 0$ . Since  $x_0$  is a strict minimizer, let  $\zeta$  be

$$2\zeta = \min_{x \in \partial B_\varepsilon(x_0)} \bar{v}_{i_0,*}(T, x) - \varphi_{i_0}(x) > 0.$$

It follows that there exists  $r > 0$  such that  $\bar{v}_{i_0}(t, x) - \varphi_{i_0}(x) \geq \zeta > 0$  for all  $(t, x) \in [T - r, T] \times \partial B_\varepsilon(x_0)$ . This holds, otherwise, for each  $r > 0$ , we could find  $(t_r, x_r) \in [T - r, T] \times \partial B_\varepsilon(x_0)$  such that  $\bar{v}_{i_0}(t_r, x_r) - \varphi_{i_0}(x_r) \leq \zeta$ . Sending  $r$  to 0, since  $\partial B_\varepsilon(x_0)$  is compact, up to a subsequence we would have  $\bar{v}_{i_0,*}(T, x^*) - \varphi_{i_0}(x^*) \leq \zeta$  for some  $x^* \in \partial B_\varepsilon(x_0)$ , in contradiction with the definition of  $\zeta$ .

Therefore, we have

$$\bar{v}_{i_0}(t, x) - \varphi_{i_0}(x) \geq \zeta \wedge \eta > 0 \quad \text{for } (t, x) \in ([T - r, T] \times \partial B_\varepsilon(x_0)) \cup (\{T\} \times B_\varepsilon(x_0)).$$

Since  $F^* \varphi_{i_0}(x_0) < \infty$ , up to smaller  $\varepsilon > 0$ , we have

$$\begin{aligned} \lambda_Y(x, y, a) - L^a \varphi_{i_0}(x) &\leq C \quad \text{for all } a \in \mathcal{N}_\varepsilon(x, D\varphi_{i_0}(x)) \\ \text{and } (x, y) \in \mathbb{R}^d \times \mathbb{R} \text{ such that } x \in B_\varepsilon(x_0) \text{ and } |y - \varphi_{i_0}(x)| &\leq \varepsilon. \end{aligned}$$

for some constant  $C > 0$ . Let  $\tilde{\varphi}_i(t, x) := \varphi_i(x) + 2C(t - T)$ . Then, for sufficiently small  $r > 0$ ,

$$\bar{v}_{i_0}(t, x) - \tilde{\varphi}_{i_0}(t, x) \geq \frac{1}{2}(\zeta \wedge \eta) > 0$$

for  $(t, x) \in ([T - r, T] \times \partial B_\varepsilon(x_0)) \cup (\{T\} \times B_\varepsilon(x_0))$ , and

$$-\partial_t \tilde{\varphi}_{i_0}(t, x) + \lambda_Y(x, y, a) - L^a \tilde{\varphi}_{i_0}(t, x) \leq -C$$

for all  $a \in \mathcal{N}_\varepsilon(x, D\tilde{\varphi}_{i_0}(t, x))$  and  $(x, y) \in \mathbb{R}^d \times \mathbb{R}$  such that  $x \in B_\varepsilon(x_0)$  and  $|y - \tilde{\varphi}_{i_0}(t, x)| \leq \varepsilon$ . By following the same arguments as in Step 3 of Section 4.3.1, the latter inequalities lead to a contradiction of (3.25).

#### 4.4.2 Viscosity subsolution

Let  $(x_0, i_0) \in \mathbb{R}^d \times \mathcal{I}$  and  $\varphi_i \in C^2(\mathbb{R}^d)$  for  $i \in \mathcal{I}$  satisfying

$$0 = \bar{v}_{i_0}^*(T, x_0) - \varphi_{i_0}(x_0) = \max_{\mathcal{I} \times \mathbb{R}^d} (\bar{v}^*(T, \cdot) - \varphi.) .$$

Without loss of generalities we can take this maximum to be strict in  $x$  once have fixed  $i_0$ . We argue by contradiction and assume  $\delta_* \varphi_{i_0}(x_0) > 0$  and

$$4\eta = \min \left\{ \varphi_{i_0}(x_0) - g_{i_0}(x_0) ; \left( \varphi_{i_0} - \sup_{0 \leq k < \bar{K}} \varphi_{i_0 k} \right) (x_0) \right\} > 0 . \quad (4.54)$$

**Step 1.** By (4.32) and Assumption A3, we can find  $r > 0$  and a locally Lipschitz map  $\hat{a}$  satisfying

$$\hat{a}(x, D\varphi_{i_0}(x)) \in \mathcal{N}_0(x, D\varphi_{i_0}(x)) \quad (4.55)$$

for all  $x \in B_r(x_0)$  Set  $\tilde{\varphi}_i(t, x) := \varphi_i(x) + \sqrt{T-t}$ . Since  $\partial \tilde{\varphi}_i(t, x) \rightarrow -\infty$  as  $t \rightarrow T$ , we deduce that, for  $r, \varepsilon > 0$  small enough,

$$\rho(t, x, y) = -\partial \tilde{\varphi}_{i_0}(t, x) + \lambda_Y(x, y, \hat{a}(x, D\tilde{\varphi}_{i_0}(t, x))) - L^{\hat{a}(x, D\tilde{\varphi}_{i_0}(t, x))} \tilde{\varphi}_{i_0}(t, x) \geq \eta , \quad (4.56)$$

for all  $(t, x, y) \in [T-r, T] \times \mathbb{R}^d \times \mathbb{R}$  such that  $x \in B_r(x_0)$  and  $|y - \tilde{\varphi}_{i_0}(t, x)| \leq \varepsilon$ . We can also notice that

$$\left( \tilde{\varphi}_{i_0} - \sup_{0 \leq k < \bar{K}} \tilde{\varphi}_{i_0 k} \right) (t, x_0) = \left( \varphi_{i_0} - \sup_{0 \leq k < \bar{K}} \varphi_{i_0 k} \right) (x_0) .$$

Therefore, we get from (4.54)

$$\left( \tilde{\varphi}_{i_0} - \sup_{0 \leq k < \bar{K}} \tilde{\varphi}_{i_0 k} \right) (t, x) \geq \eta , \quad \text{for all } (t, x) \in [T-r, T] \times B_r(x_0) . \quad (4.57)$$

for  $r > 0$  small enough.

Also observe that, since  $\bar{v}_{i_0}^* - \tilde{\varphi}_{i_0}$  is upper-semicontinuous and  $(\bar{v}_{i_0}^* - \tilde{\varphi}_{i_0})(T, x_0) = 0$ , we have

$$\bar{v}_{i_0}^*(t, x) \leq \tilde{\varphi}_{i_0}(t, x) + \varepsilon/2 \quad \text{for all } (t, x) \in [T-r, T] \times B_r(x_0) . \quad (4.58)$$

for  $r > 0$  small enough. Since  $\bar{v} \cdot (T, \cdot) = g \cdot$ , we have for  $r$  small enough

$$\tilde{\varphi}_{i_0} - \bar{v}_{i_0}(T, \cdot) = \tilde{\varphi}_{i_0} - g_{i_0} \geq \eta \quad \text{on } B_r(x_0) .$$

Since  $x_0$  is a strict maximizer for  $v_{i_0, *}(T, \cdot) - \varphi_{i_0}$ , we can define  $\zeta > 0$  such that

$$-2\zeta = \max_{x \in \partial B_r(x_0)} \bar{v}_{i_0}^*(T, x) - \varphi_{i_0}(x) < 0 .$$

for  $r > 0$  small enough. It follows that, for  $r > 0$  small enough,  $\bar{v}_{i_0}(t, x) - \tilde{\varphi}_{i_0}(x) \leq -\zeta < 0$  for all  $(t, x) \in [T-r, T] \times \partial B_r(x_0)$ . This means

$$\bar{v}_{i_0}(t, x) - \tilde{\varphi}_{i_0}(x) \leq -\zeta \wedge \eta \quad \text{for all } (t, x) \in ([T-r', T] \times \partial B_r(x_0)) \cup (\{T\} \times B_r(x_0)) . \quad (4.59)$$

By following the arguments in Step 2 of Section 4.3.2, we see that (4.55), (4.56), (4.57), (4.58), (4.59), lead to a contradiction of (3.25).



## 4.5 Uniqueness

We turn to the uniqueness of solution to the dynamic programming equation (4.28)-(4.33). To this end, we need to introduce additional assumptions. We first recall that the Hausdorff distance  $d_{\mathcal{H}}$  on closed subsets of  $A$  is defined by

$$d_{\mathcal{H}}(B, C) = \min \{r \geq 0 : B \subset C_r \text{ and } C \subset B_r\}$$

for  $B, C \subset A$  closed and nonempty, with

$$D_r = \{a \in A : \exists a' \in D, d_A(a, a') \leq r\} \quad (4.60)$$

for any  $D \subset A$  and any  $r \geq 0$ . We use the convention

$$d_{\mathcal{H}}(B, C) = +\infty$$

if  $B = \emptyset$  or  $C = \emptyset$ .

**Assumption A4.** (i) The functions  $\lambda$  and  $\sigma$  do not depend on the control, i.e.  $\lambda : \mathbb{R}^d \rightarrow \mathbb{R}^d$  and  $\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$ .

(ii) There exist two constants  $C > 0$  and  $\eta \in (0, 1]$  such that the function  $w$  appearing in Assumption A1(iii) satisfies  $w(x) \leq Cx^\eta$  for  $x \in \mathbb{R}_+$ .

(iii) There exists a constant  $C > 0$  such that

$$d_{\mathcal{H}}(\mathcal{N}_\varepsilon(x, p), \mathcal{N}_{\varepsilon'}(x', p')) \leq C(|p - p'| + \varepsilon + \varepsilon')(1 + |x|) + C|x - x'|$$

for all  $\varepsilon, \varepsilon' \geq 0, x, x', p, p' \in \mathbb{R}^d$ .

(iv)  $\mathcal{N}_0(0, 0) \neq \emptyset$ .

**Remark 4.3.** Since we use the convention (4.60), the combination of (iii) and (iv) implies that  $\mathcal{N}_\varepsilon(x, p) \neq \emptyset$  for any  $(\varepsilon, x, p) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^d$ .

In particular we always have that  $\delta\varphi \geq 0$  for any  $\varphi \in C^2(\mathbb{R}^d)$ . Therefore, the terminal viscosity supersolution solution property takes the following form

$$\min \left\{ \varphi_i(x) - g_i(x) ; \left( \varphi_i - \sup_{0 \leq k < K} \varphi_{ik} \right) (x) \right\} \geq 0 \quad (4.61)$$

for  $(x, i) \in \mathbb{R}^d \times \mathcal{I}$  and  $(\varphi_j)_{j \in \mathcal{I}}$  a test function according to Definition 4.2 (i).

**Lemma 4.2.** let  $u : [0, T] \times \mathbb{R}^d \times \mathcal{I}$  be a lower semi-continuous supersolution of (4.28)-(4.61). Define the function  $\Lambda : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$  by

$$\Lambda(t, x) = \theta e^{-\kappa t} (1 + |x|^{2\gamma+2}), \quad (t, x) \in [0, T] \times \mathbb{R}^d.$$

with  $\theta, \kappa, \gamma \in \mathbb{R}_+$ . Then, under Assumptions A1 and A4, for any  $\gamma \geq 0$  there exists  $\kappa_0 > 0$  such that for any  $\kappa \geq \kappa_0$  and  $\theta > 0$ , the function  $u + \Lambda$  is a supersolution to (4.28)-(4.33).

*Proof.* Let  $\varphi_j \in C^{1,2}([0, T] \times \mathbb{R}^d)$  for  $j \in \mathcal{I}$  be such that the function  $\varphi_j - (u + \Lambda)$  has a local maximum in  $(t, x, i)$  which is equal to 0 and  $\bar{\varphi} \in C^0([0, T] \times \mathbb{R}^d)$  such that  $\sup_{j \in \mathcal{I}} |\varphi_j| \leq \bar{\varphi}$ . Since  $u$  is a supersolution for (4.28), we have

$$\min \left\{ -\partial_t(\varphi_i - \Lambda)(t, x) + F^*(\varphi_i - \Lambda)(t, x) ; \left( (\varphi_i - \Lambda) - \sup_{0 \leq k < \bar{K}} (\varphi_{ik} - \Lambda) \right)(t, x) \right\} \geq 0.$$

We then have

$$\left( \varphi_i - \sup_{0 \leq k < \bar{K}} \varphi_{ik} \right)(t, x) = \left( (\varphi_i - \Lambda) - \sup_{0 \leq k < \bar{K}} (\varphi_{ik} - \Lambda) \right)(t, x) \geq 0. \quad (4.62)$$

We now prove that

$$-\partial_t \varphi_i(t, x) + F^* \varphi_i(t, x) \geq 0.$$

If  $F^* \varphi_i(t, x) = +\infty$ , then the inequality is obvious. Suppose that  $F^* \varphi_i(t, x) < +\infty$ . From Assumption A4, we get that  $F^*$  is locally bounded. Since  $u$  is a viscosity supersolution to (4.28), we have

$$-\partial_t(\varphi_i - \Lambda)(t, x) + F^*(\varphi_i - \Lambda)(t, x) \geq 0.$$

Using the definition of  $\Lambda$  and  $F$ , Assumption A4 and the continuity of the functions considered, we get

$$\begin{aligned} & -\partial_t \varphi_i(t, x) - \theta \kappa e^{-\kappa t} (1 + |x|^{2\gamma+2}) \\ & + \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x - x'| \leq \varepsilon \\ |(\varphi_i - \Lambda)(t, x) - y'| \leq \varepsilon \\ |D(\varphi_i - \Lambda)(t, x) - p| \leq \varepsilon}} \sup_{a \in \mathcal{N}_\varepsilon(x', p)} \{ \lambda_Y(x', y', a) \} \\ & - \lambda(x)^\top D \varphi_i(t, x) + \theta e^{-\kappa t} \lambda(x)^\top D |x|^{2\gamma+2} \\ & - \frac{1}{2} \text{Tr} \left( \sigma \sigma^\top(x) D^2 \varphi_i(t, x) \right) + \theta e^{-\kappa t} \frac{1}{2} \text{Tr} \left( \sigma \sigma^\top(x) D^2 |x|^{2\gamma+2} \right) \geq 0. \end{aligned} \quad (4.63)$$

We next define the function  $\Gamma_\varepsilon : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$  by  $\Gamma_\varepsilon(x', y', p) = \sup_{a \in \mathcal{N}_\varepsilon(x', p)} \{ \lambda_Y(x', y', a) \}$  for  $(x', y', p) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d$ . Then, we get from (4.63)

$$\begin{aligned} & -\partial_t \varphi_i(t, x) + \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x - x'| \leq \varepsilon \\ |(\varphi_i - \Lambda)(t, x) - y'| \leq \varepsilon \\ |D(\varphi_i - \Lambda)(t, x) - p| \leq \varepsilon}} \Gamma_\varepsilon(x', y', p) \\ & - \lambda(x)^\top D \varphi_i(t, x) - \frac{1}{2} \text{Tr} \left( \sigma \sigma^\top(x) D^2 \varphi_i(t, x) \right) \geq \\ & \theta \kappa e^{-\kappa t} (1 + |x|^{2\gamma+2}) + \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x - x'| \leq \varepsilon \\ |(\varphi_i - \Lambda)(t, x) - y'| \leq \varepsilon \\ |D(\varphi_i - \Lambda)(t, x) - p| \leq \varepsilon}} \Gamma_\varepsilon(x', y', p) \\ & - \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x - x'| \leq \varepsilon \\ |(\varphi_i - \Lambda)(t, x) - y'| \leq \varepsilon \\ |D(\varphi_i - \Lambda)(t, x) - p| \leq \varepsilon}} \Gamma_\varepsilon(x', y', p) \\ & - \theta e^{-\kappa t} \lambda(x)^\top D |x|^{2\gamma+2} - \theta e^{-\kappa t} \frac{1}{2} \text{Tr} \left( \sigma \sigma^\top(x) D^2 |x|^{2\gamma+2} \right) = \\ & \theta \kappa e^{-\kappa t} (1 + |x|^{2\gamma+2}) - \theta e^{-\kappa t} \lambda(x)^\top D |x|^{2\gamma+2} \\ & - \theta e^{-\kappa t} \frac{1}{2} \text{Tr} \left( \sigma \sigma^\top(x) D^2 |x|^{2\gamma+2} \right) + \Delta \Gamma^1(t, x) + \Delta \Gamma^2(t, x), \end{aligned} \quad (4.64)$$

where

$$\begin{aligned}\Delta\Gamma^1(t, x) &= \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x-x'| \leq \varepsilon \\ |\varphi_i(t, x) - y'| \leq \varepsilon \\ |D\varphi_i(t, x) - p| \leq \varepsilon}} \Gamma_\varepsilon(x', y', p) - \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x-x'| \leq \varepsilon \\ |\varphi_i(t, x) - y'| \leq \varepsilon \\ |D(\varphi_i - \Lambda)(t, x) - p| \leq \varepsilon}} \Gamma_\varepsilon(x', y', p), \\ \Delta\Gamma^2(t, x) &= \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x-x'| \leq \varepsilon \\ |\varphi_i(t, x) - y'| \leq \varepsilon \\ |D(\varphi_i - \Lambda)(t, x) - p| \leq \varepsilon}} \Gamma_\varepsilon(x', y', p) - \lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x-x'| \leq \varepsilon \\ |(\varphi_i - \Lambda)(t, x) - y'| \leq \varepsilon \\ |D(\varphi_i - \Lambda)(t, x) - p| \leq \varepsilon}} \Gamma_\varepsilon(x', y', p).\end{aligned}$$

From Assumptions A1 and A4, we get a constant  $C_1$  that does not depend on  $(t, x, i)$  such that

$$\begin{aligned}\Delta\Gamma^1(t, x) &\geq -\lim_{\varepsilon \rightarrow 0} \sup_{\substack{|x-x'| \leq \varepsilon \\ |\varphi_i(t, x) - y'| \leq \varepsilon \\ |D\varphi_i(t, x) - p| \leq \varepsilon}} \sup_{\substack{|x-\tilde{x}'| \leq \varepsilon \\ |\varphi_i(t, x) - \tilde{y}'| \leq \varepsilon \\ |D(\varphi_i - \Lambda)(t, x) - \tilde{p}| \leq \varepsilon}} \Gamma_\varepsilon(x', y', p) - \Gamma_\varepsilon(\tilde{x}', \tilde{y}', \tilde{p}) \\ &\geq -C_1 |D\Lambda(t, x)|^\eta (1 + |x|^\eta),\end{aligned}$$

Analogously, for the second term we get a constant  $C_2 > 0$  that does not depend on  $(t, x, i)$  such that

$$\Delta\Gamma^2(t, x) \geq -C_2 \Lambda(t, x).$$

Considering the right-hand side of (4.64) and taking into account the growth condition of the different terms, there exists a constant  $\kappa_0$ , which does not depend on  $\theta$ , such that if  $\kappa \geq \kappa_0$  this expression is non-negative. Henceforth, with (4.62), we obtain that  $u + \Lambda$  is a viscosity supersolution to (4.28).

We finally take  $(i, x) \in \mathcal{I} \times \mathbb{R}^d$  and functions  $\varphi_j \in C^2(\mathbb{R}^d)$  and  $\bar{\phi} \in C^0(\mathbb{R}^d)$  such that  $\sup_{i \in \mathcal{I}} |\varphi_i| \leq \bar{\phi}$  and

$$0 = u_{i_0, *}(T, x) + \Lambda(T, x) - \varphi_i(x) = \max_{\mathcal{I} \times \mathbb{R}^d} (u_{*, *}(T, \cdot) + \Lambda(T, \cdot) - \varphi).$$

Since  $u$  is a supersolution to (4.33), we have

$$\varphi_i(x) - \Lambda(T, x) \geq g_i(x),$$

since  $\Lambda \geq 0$ , we get  $\varphi_i(T, x) \geq g_i(x)$ . Combining it with (4.62), we obtain from Remark 4.3 that  $u + \Lambda$  is a viscosity supersolution to (4.33).  $\square$

We turn to the main result of this section which is a comparison theorem. We recall that the definition of  $|\cdot|$  on  $\mathcal{I}$  is given in Section 2.1.

**Theorem 4.5.** *Let  $\bar{w}$ . (resp.  $\bar{u}$ .) be a lsc (resp. usc) viscosity supersolution (resp. subsolution) to (4.28)-(4.61). Suppose that there exists  $\gamma > 0$  such that*

$$\sup_{(t, x, i) \in [0, T] \times \mathbb{R}^d \times \mathcal{I}} \frac{|\bar{w}_i(t, x)| + |\bar{u}_i(t, x)|}{1 + |x|^\gamma} < +\infty, \quad (4.65)$$

and

$$\sup_{(t, x) \in [0, T] \times \mathbb{R}^d} |\bar{w}_i(t, x)| + |\bar{u}_i(t, x)| \xrightarrow{|i| \rightarrow \infty} 0. \quad (4.66)$$

Then, under Assumption A1-A2-A4, we have  $\bar{u} \leq \bar{w}$ . on  $[0, T] \times \mathbb{R}^d \times \mathcal{I}$ .

*Proof.* We proceed in six steps.

**Step 1.** We define  $\Lambda_{\theta,\kappa}(t,x) = \theta e^{-\kappa t}(1 + |x|^{2\gamma+2})$  for  $(t,x) \in [0,T] \times \mathbb{R}^d$  with  $\theta, \kappa \in \mathbb{R}_+$ . From Lemma 4.2, there exist  $\kappa$  large enough such that for any  $\theta > 0$ ,  $\bar{w}_i + \Lambda_{\theta,\kappa}$  is also a supersolution for (4.28)-(4.33). Set  $\bar{w}_{i,\theta,\kappa}(t,x) = \bar{w}_i(t,x) + \Lambda_{\theta,\kappa}(t,x)$ ,  $(i,t,x) \in \mathcal{I} \times [0,T] \times \mathbb{R}^d$ .

For some  $\eta, \eta' > 0$  to be chosen below, let  $\beta_t = e^{(\eta+\eta')t}$  for  $t \in [0,T]$ . A straightforward derivation shows that  $\beta_t \bar{w}_{i,\theta,\kappa}$  (resp.  $\beta_t \bar{u}_i$ ) is a viscosity supersolution (resp. subsolution) to

$$\min \left\{ \eta w_i - \partial_t w_i + \tilde{F}(t,x,w_i,Dw_i) - \lambda^\top Dw_i - \frac{1}{2} \text{Tr} \left( \sigma \sigma^\top D^2 w_i \right) ; \right. \\ \left. w_i - \sup_{0 \leq k < \bar{K}} w_{ik} \right\} = 0 \text{ on } [0,T] \times \mathbb{R}^d, \quad (4.67)$$

$$\min \left\{ w_i - \tilde{g} ; \delta w_i ; w_i - \sup_{0 \leq k < \bar{K}} w_{ik} \right\} = 0 \text{ on } \{T\} \times \mathbb{R}^d. \quad (4.68)$$

where

$$\tilde{F}(t,x,y,p) = \sup_{a \in \tilde{\mathcal{N}}_0(t,x,p)} \tilde{\lambda}_Y(x,y,a) \quad , \quad \tilde{\mathcal{N}}_0(t,x,p) = \mathcal{N}_0(x,\beta_t^{-1}p) \quad , \\ \tilde{\lambda}_Y(t,x,y,a) = \beta_t \lambda_Y(x,\beta_t^{-1}y,a) + \eta' y \quad , \quad \tilde{g}_i(x) = \beta_T g_i(x) \quad ,$$

for all  $(t,x,i,y,p,a) \in [0,T] \times \mathbb{R}^d \times \mathcal{I} \times \mathbb{R} \times \mathbb{R}^d \times A$ . Since  $\lambda_Y$  is Lipschitz, we can choose  $\eta'$  large enough so that  $\tilde{\lambda}_Y$  and, consequently,  $\tilde{F}$  are nondecreasing in  $y$ .

Let  $\varepsilon > 0$ . From an analogous computation, using the monotonicity of  $\tilde{F}$ , we see that  $\beta_t \bar{w}_i + \varepsilon/2^{|\bar{i}|}$  is a viscosity supersolution to

$$\eta w_i - \partial_t w_i + \tilde{F}(t,x,w_i,Dw_i) - \lambda(x)^\top Dw_i - \frac{1}{2} \text{Tr} \left( \sigma \sigma(x)^\top D^2 w_i \right) \geq 0 \quad , \quad (4.69)$$

$$\min \{ w_i(T, \cdot) - \tilde{g} ; \delta w_i \} \geq 0 \quad , \quad (4.70)$$

$$w_i - \sup_{0 \leq k < \bar{K}} w_{ik} \geq \frac{\varepsilon}{2^{|\bar{i}|+1}} =: \Delta_i > 0 \quad . \quad (4.71)$$

**Step 2.** Set  $\tilde{u}_i = \beta_t \bar{u}_i$  and  $\tilde{w}_{i,\theta,\kappa,\varepsilon} = \beta_t \bar{w}_i + \beta_t \Lambda_{\theta,\kappa} + \varepsilon/2^{|\bar{i}|} = \beta_t \bar{w}_{i,\theta,\kappa} + \varepsilon/2^{|\bar{i}|}$ . To prove our result, it is enough to show that

$$\tilde{u}_i(t,x) \leq \tilde{w}_{i,\theta,\kappa,\varepsilon}(t,x)$$

for each  $(i,t,x) \in \mathcal{I} \times [0,T] \times \mathbb{R}^d$  and  $\theta, \varepsilon > 0$ . Then taking the limit as  $\theta \rightarrow 0$  and  $\varepsilon \rightarrow 0$ , we obtain the desired result. For simplicity, we write  $\tilde{w}_{i,\theta,\kappa,\varepsilon}$  for  $\tilde{w}_i$  in the sequel. We argue by contradiction and suppose that

$$\sup_{\mathcal{I} \times [0,T] \times \mathbb{R}^d} \tilde{u}_i - \tilde{w}_i > 0 \quad . \quad (4.72)$$

Due to the growth condition on  $\tilde{u}_i$  and  $\tilde{w}_i$ , there exist  $R > 0$  such that

$$\tilde{u}_i(t,x) - \tilde{w}_i(t,x) < 0 \quad (4.73)$$

for all  $(i,t,x) \in \mathcal{I} \times [0,T] \times \mathbb{R}^d$  such that  $|x| \geq R$ . Then from (4.66) and since  $u_i - \tilde{w}_i$  is upper semicontinuous, there exist  $(i_0, t_0, x_0) \in \mathcal{I} \times [0,T] \times \mathbb{R}^d$  such that

$$\sup_{(i,t,x) \in \mathcal{I} \times [0,T] \times \mathbb{R}^d} (\tilde{u}_i - \tilde{w}_i)(t,x) = (\tilde{u}_{i_0} - \tilde{w}_{i_0})(t_0, x_0) > 0 \quad . \quad (4.74)$$

**Step 3.** For  $n \geq 1$ , we define the function

$$\Theta_n(t, x, y, i) = \tilde{u}_i(t, x) - \tilde{w}_i(t, y) - \varphi_n(t, x, y, i)$$

with

$$\varphi_n(t, x, y, i) = n|x - y|^2 + |x - x_0|^4 + |t - t_0|^2 + \mathbf{1}_{i \neq i_0}.$$

for all  $(t, x, y, i) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{I}$ . By the growth assumption on  $\tilde{u}$  and  $\tilde{w}$  and (4.66), for all  $n$ , there exists  $(t_n, x_n, y_n, i_n) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{I}$  attaining the maximum of  $\Theta_n$  on  $[0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{I}$ . We have

$$\Theta_n(t_n, x_n, y_n, i_n) \geq \Theta_n(t_0, x_0, x_0, i_0) = (\tilde{u}_{i_0} - \tilde{w}_{i_0})(t_0, x_0).$$

By (4.73) and (4.66), up to a subsequence,  $(t_n, x_n, y_n, i_n)$  converge to  $(\hat{t}, \hat{x}, \hat{y}, \hat{i})$ . Sending  $n$  to infinity, this provides

$$\begin{aligned} \bar{\ell} := \limsup_{n \rightarrow \infty} \varphi_n(t_n, x_n, y_n, i_n) &\leq \limsup_{n \rightarrow \infty} [\tilde{u}_{i_n}(t_n, x_n) - \tilde{w}_{i_n}(t_n, y_n) - (\tilde{u}_{i_0} - \tilde{w}_{i_0})(t_0, x_0)] \\ &\leq \tilde{u}_{\hat{i}}(\hat{t}, \hat{x}) - \tilde{w}_{\hat{i}}(\hat{t}, \hat{y}) - (\tilde{u}_{i_0} - \tilde{w}_{i_0})(t_0, x_0). \end{aligned}$$

In particular,  $\bar{\ell} < +\infty$  and  $\hat{x} = \hat{y}$ . Using the definition of  $(t_0, x_0, i_0)$  as a maximizer of  $\tilde{u} - \tilde{w}$ , we see that:

$$0 \leq \bar{\ell} \leq (\tilde{u}_{\hat{i}} - \tilde{w}_{\hat{i}})(\hat{t}, \hat{x}) - (\tilde{u}_{i_0} - \tilde{w}_{i_0})(t_0, x_0) \leq 0,$$

which implies

$$(t_n, x_n, y_n, i_n) \rightarrow (t_0, x_0, x_0, i_0), \quad (4.75)$$

$$n|x_n - y_n|^2 \rightarrow 0, \quad (4.76)$$

$$\tilde{u}_{i_n}(t_n, x_n) - \tilde{w}_{i_n}(t_n, y_n) \rightarrow (\tilde{u}_{i_0} - \tilde{w}_{i_0})(t_0, x_0). \quad (4.77)$$

Being  $\mathcal{I}$  endowed with the discrete topology, we can assume  $i_n = i_0$  for all  $n \geq 1$ .

**Step 4.** We now show that for  $n$  large enough

$$\tilde{u}_{i_0}(t_n, x_n) - \sup_{0 \leq k < \bar{K}} \tilde{u}_{i_0 k}(t_n, x_n) > 0. \quad (4.78)$$

On the contrary, up to a subsequence, we would have for all  $n$ ,

$$\tilde{u}_{i_0}(t_n, x_n) - \sup_{0 \leq k < \bar{K}} \tilde{u}_{i_0 k}(t_n, x_n) \leq 0. \quad (4.79)$$

Moreover, by the viscosity supersolution property of  $\tilde{w}$  to (4.71), we have

$$\tilde{w}_{i_0}(t_n, y_n) - \sup_{0 \leq k < \bar{K}} \tilde{w}_{i_0 k}(t_n, y_n) \geq \Delta_{i_0} > 0.$$

We deduce from the two previous inequalities

$$\begin{aligned} \tilde{u}_{i_0}(t_n, x_n) - \sup_{0 \leq k < \bar{K}} \tilde{u}_{i_0 k}(t_n, x_n) &\leq \tilde{w}_{i_0}(t_n, y_n) - \sup_{0 \leq k < \bar{K}} \tilde{w}_{i_0 k}(t_n, y_n) - \Delta_{i_0} \\ \tilde{u}_{i_0}(t_n, x_n) - \tilde{w}_{i_0}(t_n, y_n) + \Delta_{i_0} &\leq \sup_{0 \leq k < \bar{K}} \tilde{u}_{i_0 k}(t_n, x_n) - \sup_{0 \leq k < \bar{K}} \tilde{w}_{i_0 k}(t_n, y_n) \\ &\leq \sup_{0 \leq k < \bar{K}} [\tilde{u}_{i_0 k}(t_n, x_n) - \tilde{w}_{i_0 k}(t_n, y_n)]. \end{aligned} \quad (4.80)$$

Since  $\Delta_{i_0} > 0$ , for all  $n$  there exists  $k_n$  such that

$$\sup_{0 \leq k < \bar{K}} [\tilde{u}_{i_0 k}(t_n, x_n) - \tilde{w}_{i_0 k}(t_n, y_n)] - \frac{\Delta_{i_0}}{2} \leq \tilde{u}_{i_0 k_n}(t_n, x_n) - \tilde{w}_{i_0 k_n}(t_n, y_n) .$$

From (4.66), up to a subsequence, we may assume that  $(k_n)$  converges to  $k_0$  in  $\mathbb{N}$ . Hence, by sending  $n$  to infinity into (4.80), it follows with (4.77) and the upper (resp. lower)-semicontinuity of  $\tilde{u}$  (resp.  $\tilde{w}$ ) that :

$$(\tilde{u}_{i_0} - \tilde{w}_{i_0})(t_0, x_0) + \frac{\Delta_{i_0}}{2} \leq (\tilde{u}_{i_0 k_0} - \tilde{w}_{i_0 k_0})(t_0, x_0) ,$$

which is a contradiction to (4.74).

**Step 5.** Let us check that, up to a subsequence,  $t_n < T$  for all  $n$ . On the contrary,  $t_n = t_0 = T$  for  $n$  large enough, and from (4.78), and the viscosity subsolution property of  $\tilde{u}$  to (4.68), we would get

$$\tilde{u}_{i_0}(T, x_n) \leq \tilde{g}_{i_0}(x_n) .$$

On the other hand, by the viscosity supersolution property of  $\tilde{w}$  to (4.68), we have  $\tilde{w}(T, y_n) \geq \tilde{g}_{i_0}(y_n)$ , and so

$$\tilde{u}_{i_0}(T, x_n) - \tilde{w}(T, y_n) \leq \tilde{g}_{i_0}(x_n) - \tilde{g}_{i_0}(y_n) .$$

By sending  $n$  to infinity, and from Assumption (A2) and (4.77), this would imply  $\tilde{u}_{i_0}(t_0, x_0) - \tilde{v}(t_0, x_0) \leq 0$ , a contradiction to (4.72).

**Step 6.** We may then apply Ishii's lemma (see Theorem 8.3 in [7]) to  $(t_n, x_n, y_n) \in [0, T) \times \mathbb{R}^d \times \mathbb{R}^d$  that attains the maximum of  $\Theta_n(\cdot, i_0)$  and we get  $(p_{\tilde{u}}^n, q_{\tilde{u}}^n, M_n) \in \bar{J}^{2,+} \tilde{u}_{i_0}(t_n, x_n)$  and  $(p_{\tilde{v}}^n, q_{\tilde{v}}^n, N_n) \in \bar{J}^{2,-} \tilde{v}_{i_0}(t_n, y_n)$  such that

$$\begin{aligned} p_{\tilde{u}}^n - p_{\tilde{v}}^n &= \partial_t \varphi_n(t_n, x_n, y_n, i_0) = 2(t_n - t_0) , \\ q_{\tilde{u}}^n &= D_x \varphi_n(t_n, x_n, y_n, i_0) = n(x_n - y_n) + 4(x_n - x_0)|x_n - x_0|^2 , \\ q_{\tilde{v}}^n &= -D_y \varphi_n(t_n, x_n, y_n, i_0) = n(x_n - y_n) , \end{aligned}$$

andx

$$\begin{pmatrix} M_n & 0 \\ 0 & -N_n \end{pmatrix} \leq A_n + \frac{1}{2n} A_n^2 , \quad (4.81)$$

where

$$A_n = D_{(x,y)}^2 \varphi_n(t_n, x_n, y_n, i_0) = n \begin{pmatrix} \mathbb{I}_d & -\mathbb{I}_d \\ -\mathbb{I}_d & \mathbb{I}_d \end{pmatrix} - \begin{pmatrix} 4|x_n - x_0|^2 \mathbb{I}_d + 8(x_n - x_0)(x_n - x_0)^\top & \mathbb{O}_d \\ \mathbb{O}_d & \mathbb{O}_d \end{pmatrix} ,$$

with  $\mathbb{I}_d$  and  $\mathbb{O}_d$  respectively the identity and the zero matrix of  $\mathbb{R}^{d \times d}$ . We can therefore bound the right-hand side of (4.81) by

$$A_n + \frac{1}{2n} A_n^2 \leq 3n \begin{pmatrix} \mathbb{I}_d & -\mathbb{I}_d \\ -\mathbb{I}_d & \mathbb{I}_d \end{pmatrix} + \begin{pmatrix} A_n' & \mathbb{O}_d \\ \mathbb{O}_d & \mathbb{O}_d \end{pmatrix} , \quad (4.82)$$

with  $A'_n$  such that  $\limsup_{n \rightarrow \infty} \frac{1}{|x_n - x_0|^2} |A'_n| < +\infty$ . From the viscosity supersolution property of  $\tilde{w}_{i_0}$  to (4.67), we have

$$\eta \tilde{w}_{i_0}(t_n, y_n) - p_{\tilde{w}}^n + \tilde{F}^*(t_n, y_n, \tilde{w}_{i_0}(t_n, y_n), q_{\tilde{w}}^n) - \lambda(y_n)^\top q_{\tilde{w}}^n - \frac{1}{2} \text{Tr}(\sigma \sigma^\top(y_n) N_n) \geq 0 .$$

On the other hand, from (4.78) and the viscosity subsolution property of  $\tilde{u}$  to (4.67), we have

$$\eta \tilde{u}_{i_0}(t_n, x_n) - p_{\tilde{u}}^n + \tilde{F}_*(t_n, x_n, \tilde{u}_{i_0}(t_n, x_n), q_{\tilde{u}}^n) - \lambda(x_n)^\top q_{\tilde{u}}^n - \frac{1}{2} \text{Tr}(\sigma \sigma^\top(x_n) M_n) \leq 0 .$$

By subtracting the two previous inequalities, we obtain

$$\begin{aligned} \eta(\tilde{u}_{i_0}(t_n, x_n) - \tilde{w}_{i_0}(t_n, y_n)) &\leq p_{\tilde{u}}^n - p_{\tilde{w}}^n + \tilde{F}^*(t_n, y_n, \tilde{w}_{i_0}(t_n, y_n), q_{\tilde{w}}^n) - \tilde{F}_*(t_n, x_n, \tilde{u}_{i_0}(t_n, x_n), q_{\tilde{u}}^n) + \\ &\quad + \lambda(x_n)^\top q_{\tilde{u}}^n - \lambda(y_n)^\top q_{\tilde{w}}^n + \frac{1}{2} \text{Tr}(\sigma \sigma^\top(x_n) M_n) - \frac{1}{2} \text{Tr}(\sigma \sigma^\top(y_n) N_n) \\ &= p_{\tilde{u}}^n - p_{\tilde{w}}^n + \Delta C_n^1 + \Delta C_n^2 + \Delta C_n^3 \end{aligned} \quad (4.83)$$

where

$$\begin{aligned} \Delta C_n^1 &= \tilde{F}^*(t_n, y_n, \tilde{w}_{i_0}(t_n, y_n), q_{\tilde{w}}^n) - \tilde{F}_*(t_n, x_n, \tilde{u}_{i_0}(t_n, x_n), q_{\tilde{u}}^n) , \\ \Delta C_n^2 &= \lambda(x_n)^\top q_{\tilde{u}}^n - \lambda(y_n)^\top q_{\tilde{w}}^n , \\ \Delta C_n^3 &= \frac{1}{2} \text{Tr}(\sigma \sigma^\top(x_n) M_n) - \frac{1}{2} \text{Tr}(\sigma \sigma^\top(y_n) N_n) . \end{aligned}$$

From (4.75)), we have  $p_{\tilde{u}}^n - p_{\tilde{w}}^n \rightarrow 0$  as  $n \rightarrow \infty$ . From the Lipschitz continuity of  $\lambda$  and (4.76), we have  $\Delta C^2 \rightarrow 0$  as  $n \rightarrow \infty$ . From (4.82), (4.75), (4.76), and the Lipschitz property of  $\sigma$ , we also have  $\Delta C^3 \rightarrow 0$  as  $n \rightarrow \infty$ . Following the same argument as in the proof of lemma 4.2, we get from Assumptions A1 and A4 and (4.75)  $\Delta C_n^1 \rightarrow 0$  as  $n \rightarrow \infty$ .

Therefore, by sending  $n \rightarrow \infty$  into (4.83), we conclude with (4.77) that  $\eta(\tilde{u}_{i_0}(t_0, x_0) - \tilde{w}_{i_0}(t_0, y_0)) \leq 0$ , a contradiction with (4.74).  $\square$

From Theorems 4.3, 4.4 and 4.5, we get the following characterisation of the function  $\bar{v}$ .

**Corollary 4.1.** *Suppose that  $\bar{v}$  satisfies*

$$\sup_{(t,x,i) \in [0,T] \times \mathbb{R}^d \times \mathcal{I}} \frac{|\bar{v}_i(t,x)| + |\bar{u}_i(t,x)|}{1 + |x|^\gamma} < +\infty , \quad (4.84)$$

for some  $\gamma > 0$  and

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\bar{v}_i(t,x)| + |\bar{u}_i(t,x)| \xrightarrow{|i| \rightarrow \infty} 0 . \quad (4.85)$$

Under Assumptions A1-A2-A3-A4,  $\bar{v}$  is the unique viscosity solution to (4.28)-(4.61) satisfying (4.84)-(4.85). Moreover,  $\bar{v}$  is continuous on  $[0, T] \times \mathbb{R}^d \times \mathcal{I}$ .

We recall that Section 2.3 provides an example of a value function satisfying conditions (4.28)-(4.61).

# A Appendix

## A.1 Set of atomic finite measures

**Proposition A.6.** For  $\ell \geq 1$ ,  $E_\ell$  is a closed subset of  $\mathcal{M}_F(\mathcal{I} \times \mathbb{R}^\ell)$  for the topology of the weak convergence of measures.

*Proof.* Let  $(\mu_n)_{n \in \mathbb{N}}$  be a sequence of  $E_\ell$  such that  $\mu_n = \sum_{i \in V_n} \delta_{(i, x_n^i)} \xrightarrow{w} \mu \in \mathcal{M}_F(\mathcal{I} \times \mathbb{R}^\ell)$ . We prove that  $\mu$  is an element of  $E_\ell$ , i.e. it can be written as  $\mu = \sum_{i \in V} \delta_{(i, x^i)}$  for some set  $V \subseteq \mathcal{I}$ ,  $|V| < \infty$  and some points  $(x_i)_{i \in V}$ .

Consider the continuous functions  $\mathbb{1}_{\{i\} \times \mathbb{R}^\ell}$  for  $i \in \mathcal{I}$ . We then have

$$\langle \mu_n, \mathbb{1}_{\{i\} \times \mathbb{R}^\ell} \rangle = \begin{cases} 1 & \text{if } i \in V_n \\ 0 & \text{if } i \notin V_n \end{cases}$$

For each  $i \in \mathcal{I}$ , we have that the sequence  $(\langle \mu_n, \mathbb{1}_{\{i\} \times \mathbb{R}^\ell} \rangle)_n$  is a convergent sequence in  $\{0, 1\}$ , which is in particular stationary. Let  $V$  be defined as follow:

$$V := \left\{ i \in \mathcal{I} : \langle \mu_n, \mathbb{1}_{\{i\} \times \mathbb{R}^\ell} \rangle \xrightarrow{n \rightarrow \infty} 1 \right\}.$$

Let  $i \in V$ . Since the functions previously described converge, they are constant from a certain rank and there exists  $n_i \in \mathbb{N}$  such that for  $n \geq n_i$  we have  $i \in V_n$ . For  $f \in C(\mathbb{R}^\ell)$  and consider the function  $\mathbb{1}_{\{i\}} \otimes f : \mathcal{I} \times \mathbb{R}^\ell \rightarrow \mathbb{R}$ . We have:

$$f(x_n^i) = \langle \mu_n, \mathbb{1}_{\{i\}} \otimes f \rangle \longrightarrow \langle \mu, \mathbb{1}_{\{i\}} \otimes f \rangle \in \mathbb{R}.$$

This means that for each  $i \in V$  and  $f \in C(\mathbb{R}^\ell)$  the sequence  $(f(x_n^i))_n$  converges, therefore  $(x_n^i)_n$  converges to a point  $x^i \in \mathbb{R}^\ell$ .

We then notice that any continuous and bounded function  $f$  on  $\mathcal{I} \times \mathbb{R}^\ell$  is of the form  $f = \sum_{i \in \mathcal{I}} f_i$  with  $f_i$  is continuous and bounded on  $\mathbb{R}^\ell$ . In particular, we get

$$\int_{\mathcal{I} \times \mathbb{R}^\ell} f d\mu_n = \sum_{i \in V} f_i(x_n^i)$$

for  $n$  large enough, and

$$\int_{\mathcal{I} \times \mathbb{R}^\ell} f d\mu_n \xrightarrow{n \rightarrow +\infty} \int_{\mathcal{I} \times \mathbb{R}^\ell} f d \left( \sum_{i \in V} \delta_{(i, x^i)} \right)$$

so we have  $\mu = \sum_{i \in V} \delta_{(i, x^i)}$ .

To finally prove that  $\mu \in E_\ell$  we need to show that there do not exist  $i, j \in V$  such that  $i \prec j$ . Fix  $i, j \in V$ . From the previous steps, there exists some  $n$  such that  $i, j \in V_n$ . Since  $\mu_n \in E_\ell$ , we get  $i \not\prec j$  and  $j \not\prec i$ . Therefore, we have  $\mu \in E_\ell$ .  $\square$

## A.2 Branching martingale controlled problem

We first set our controlled martingale problem. We define the set  $\tilde{E}_{m+1}$  as the set of finite measure  $\mu$  on  $\mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1}$  of the form

$$\mu = \sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \delta_{(i, n, b^i, q^i, n)}$$



with  $b^i \in \mathbb{R}^m$  and  $q^{i,n} \in \mathbb{R}$  for  $i \in \mathcal{I}$  and  $n \in \mathbb{N}$ . From the same argument as for  $E_\ell$ , we have that  $\tilde{E}_{m+1}$  is Polish.

We then set  $\mathbf{X} = \mathbb{D}([0, T], E_{d+1}) \times \mathbb{D}([0, T], \tilde{E}_{m+1})$  the space pairs of càdlàg functions from  $[0, T]$  to  $E_{d+1}$  and  $\tilde{E}_{m+1}$ . We denote by  $\mathbf{x}$  the canonical process and by  $\mathbb{G} = (\mathcal{G}_t)_{t \in [0, T]}$  the canonical filtration on  $\mathbf{X}$ .

For  $\bar{x} = (\sum_{i \in \mathcal{V}_s} \delta_{(i, \hat{x}_s^i)}, \sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \delta_{(i, n, b_s^i, q_s^{i,n})})_{s \in [0, T]} \in \mathbf{X}$ , we write

$${}^1\bar{x} = \left( \sum_{i \in \mathcal{V}_s} \delta_{(i, \hat{x}_s^i)} \right)_{s \in [0, T]} \quad \text{and} \quad {}^2\bar{x} = \left( \sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \delta_{(i, n, b_s^i, q_s^{i,n})} \right)_{s \in [0, T]} \quad (\text{A.86})$$

We also define  ${}^1\mathbf{x}$  and  ${}^2\mathbf{x}$  the first and second component of the canonical process.

Let  $C^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1})$  (resp.  $C^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$ ) be the set of functions  $f : [0, T] \times \mathcal{I} \times \mathbb{R}^{d+1} \rightarrow \mathbb{R}$  (resp.  $g : [0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ ) such that  $f_i(\cdot) \in C^{1,2}([0, T] \times \mathbb{R}^{d+1})$  (resp.  $g_{i,n}(\cdot) \in C^{1,2}([0, T] \times \mathbb{R}^{m+1})$ ) for all  $i \in \mathcal{I}$  (resp.  $(i, n) \in \mathcal{I} \times \mathbb{N}$ ) and  $C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1})$  (resp.  $C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$ ) the set of  $f \in C^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1})$  (resp.  $g \in C^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$ ) such that there exists a compact  $K \subset \mathbb{R}^{d+1}$  (resp.  $K \subset \mathbb{R}^{m+1}$ ) satisfying  $f_i(t, x) = 0$  (resp.  $g_{i,n}(t, x) = 0$ ) for  $(t, i, x) \in [0, T] \times \mathcal{I} \times \mathbb{R}^{d+1}$  such that  $x \notin K$  (resp.  $(t, i, n, x) \in [0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1}$  such that  $x \notin K$ ).

We define the operator  $\Delta$  by

$$\Delta g(s, \mu) = \sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \Delta_b g_{i,n}(b^i, q^{i,n})$$

for  $g \in C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$  and  $\mu \in \sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \delta_{(i, n, b^i, q^{i,n})} \in \tilde{E}_{m+1}$ , where  $\Delta_b$  stands for the Laplacian operator with respect to the third variable of the function  $g$ . For a given control  $\alpha \in \mathcal{A}$ , let  $\bar{L}^\alpha$  be the following second order local operator

$$\begin{aligned} \bar{L}^\alpha F_{f,g}(s, \bar{x}) &= \partial_1 F(f(s, {}^1\bar{x}_s), g(s, {}^2\bar{x}_s)) \sum_{i \in \mathcal{V}_s} (\partial_t + \hat{L}^{\alpha^i(s, {}^2\bar{x}_s)}) f_i(s, \hat{x}_s^i) \\ &+ \partial_1 F(f(s, {}^1\bar{x}_s), g(s, {}^2\bar{x}_s)) \left( \partial_t + \frac{1}{2} \Delta \right) g(s, {}^2\bar{x}_s) \\ &+ \frac{1}{2} \partial_{11} F(f(s, {}^1\bar{x}_s), g(s, {}^2\bar{x}_s)) \sum_{i \in \mathcal{V}_s} |\hat{\sigma}(\hat{x}_s^i, \alpha(s, {}^2\bar{x}))^\top D f_i(s, \hat{x}_s^i)|^2 \\ &+ \partial_{12} F(f(s, {}^1\bar{x}_s), g(s, {}^2\bar{x}_s)) \\ &\sum_{i \in \mathcal{V}_s, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \partial_b g_{i,n}(s, b_s^i, q_s^{i,n})^\top \hat{\sigma}(\hat{x}_s^i, \alpha(s, {}^2\bar{x}))^\top D f_i(s, \hat{x}_s^i) \\ &+ \frac{1}{2} \partial_{22} F(f(s, {}^1\bar{x}_s), g(s, {}^2\bar{x}_s)) \sum_{i \in \mathcal{I}, n \in \mathbb{N}} \left| \frac{1}{2^{2(|i|+n)}} \partial_b g_{i,n}(s, b_s^i, q_s^{i,n}) \right|^2 \\ &+ \gamma \sum_{i \in \mathcal{I}, n \in \mathbb{N}} p_n \left\{ F \left( f(s, {}^1\bar{x}_s) + \mathbf{1}_{\mathcal{V}_s}(i) \sum_{\ell=0}^{n-1} (f_{i\ell} - f_i)(s, \hat{x}_s^i), \right. \right. \\ &\quad \left. \left. g(s, {}^2\bar{x}_s) + \frac{1}{2^{2(|i|+n)}} (g(i, n, s, b_s^i, q_s^{i,n} + 1) - g(i, n, s, b_s^i, q_s^{i,n})) \right. \right. \\ &\quad \left. \left. - F(f(s, {}^1\bar{x}_s), g(s, {}^2\bar{x}_s)) \right\} \end{aligned}$$

for  $s \in [0, T]$ ,  $\bar{x} = \left( \sum_{i \in \mathcal{V}_s} \delta_{(i, \hat{x}_u^i)}, \sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \delta_{(i, n, b_u^i, q_u^{i, n})} \right)_{u \in [0, T]} \in \mathbf{X}$ ,  $f \in C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1})$ ,  $g \in C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$  and  $F \in C_c^2(\mathbb{R}^2)$  with  $F_{f,g} = F \circ (f, g)$ . We then define the process  $\bar{M}^{t, \alpha, F_f}$  by

$$\bar{M}_s^{t, \alpha, F_{f,g}} = F_f(s, \mathbf{x}_s) - \int_t^s \bar{L}^\alpha F_{f,g}(u, \mathbf{x}) du, \quad s \in [t, T].$$

**Definition A.3** (Martingale problem). *Consider the initial condition  $(t, \bar{x}) \in [0, T] \times \mathbf{X}$ , and a control  $\alpha \in \mathcal{A}$ . A probability measure  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}$  is a solution to the controlled martingale problem if the process  $\bar{M}^{t, \alpha, F_{f,g}}$  is a  $\mathbb{G}$ -martingale under  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}$  for any  $f \in C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1})$ ,  $g \in C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$ , and any  $F \in C_c^2(\mathbb{R}^2)$ ,  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}(\mathbf{1}_{\mathbf{x}_s} = \bar{x}_s \text{ for } s \in [0, t]) = 1$  and  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}({}^2\mathbf{G} \in {}^2G) = \mathbb{W}({}^2G)$  for any  ${}^2G \in {}^2\mathcal{G}_t$  where  $\mathbb{W}$  stands for the law of the process  $\xi$  and  ${}^2\mathbb{G} = ({}^2\mathcal{G}_t)_{t \in [0, T]}$  stands for the canonical filtration on  $\mathbb{D}([0, T], \tilde{E}_{m+1})$ .*

**Definition A.4** (Shifted martingale problem). *Consider the initial condition  $(t, \bar{x}) \in [0, T] \times \mathbf{X}$ , and a control  $\alpha \in \mathcal{A}$ . A probability measure  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}$  is a solution to the shifted controlled martingale problem if the process  $\bar{M}^{t, \alpha, F_{f,g}}$  is a  $\mathbb{G}$ -martingale under  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}$  for any  $f \in C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1})$ ,  $g \in C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{m+1})$ , and any  $F \in C_c^2(\mathbb{R}^2)$ , and  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}(\mathbf{x}_s = \bar{x}_s \text{ for } s \in [0, t]) = 1$ .*

We are now able to state the main result of this section. For that we need the following notations. We first extend the definition of the concatenation operator  $\oplus$  on  $\mathbb{D}([0, T], \tilde{E}_{m+1})$  as follows:

$$(y \oplus_t \tilde{y})_s = \sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \delta_{(i, n, (b^i \oplus_t \tilde{b}^i)_s, (q^{i, n} \oplus_t \tilde{q}^{i, n})_s)}$$

with

$$\begin{aligned} (b^i \oplus_t \tilde{b}^i)_s &= b_{s \wedge t}^i \mathbf{1}_{s < t} + (\tilde{b}_s^i - \tilde{b}_t^i + b_t^i) \mathbf{1}_{s \geq t} \\ (q^{i, n} \oplus_t \tilde{q}^{i, n})_s &= q_{s \wedge t}^{i, n} \mathbf{1}_{s < t} + (\tilde{q}_s^{i, n} - \tilde{q}_t^{i, n} + q_t^{i, n}) \mathbf{1}_{s \geq t} \end{aligned}$$

for  $s \in [0, T]$ ,  $y = (\sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \delta_{(i, n, b_u^i, q_u^{i, n})})_{u \in [0, T]}$  and  $\tilde{y} = (\sum_{i \in \mathcal{I}, n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \delta_{(i, n, \tilde{b}_u^i, \tilde{q}_u^{i, n})})_{u \in [0, T]}$ . In particular, we have

$$\xi(\omega \oplus_t \tilde{\omega}) = \xi(\omega) \oplus_t \xi(\tilde{\omega})$$

for  $\omega, \tilde{\omega} \in \Omega$  and  $t \in [0, T]$ . For  $\eta : [0, T] \times {}^2\mathbf{X} \rightarrow \mathbb{R}$  and  $\bar{x} \in \mathbf{X}$  we finally define the function  $\eta^{t, \bar{x}}$  by

$$\eta^{t, \bar{x}}(s, {}^2\bar{x}') = \eta(s, {}^2\bar{x} \oplus_t {}^2\bar{x}')$$

for  $\bar{x}' \in \mathbf{X}$  and  $s \in [0, T]$ .

**Theorem A.6.** *Suppose that Assumption A1 holds and that there exists a unique solution to the martingale problem and the shifted martingale problem for each initial condition and control. Let  $(t, \bar{x}, \alpha) \in [0, T] \times \mathbf{X} \times \mathcal{A}$  and  $\tau$  a  $\mathbb{G}$ -stopping time valued in  $[t, T]$ . Then, we have*

$$\bar{\mathbb{P}}_{\bar{x}'}^{t, \bar{x}, \alpha} = \mathbb{P}^{\tau(\bar{x}'), \bar{x}', \alpha^{\tau(\bar{x}'), \bar{x}'}} , \quad \bar{\mathbb{P}}^{t, \bar{x}, \alpha}(d\bar{x}') - \text{a.s.}$$

where  $(\bar{\mathbb{P}}_{\bar{x}'}^{t, \bar{x}, \alpha}, \bar{x}' \in \mathbf{X})$  is a regular conditional probability distribution of  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}$  given  $\mathcal{G}_\tau$ .

*Proof.* We first define the function  $\hat{\lambda}^\alpha$  and  $\hat{\sigma}^\alpha$  by

$$\hat{\lambda}^\alpha(s, \bar{x}) = \hat{\lambda}(1\bar{x}_s, \alpha(s, 2\bar{x})) \quad \text{and} \quad \hat{\sigma}^\alpha(s, \bar{x}) = \hat{\sigma}(1\bar{x}_s, \alpha(s, 2\bar{x})), \quad (s, \bar{x}) \in [0, T] \times \mathbf{X},$$

for  $\alpha \in \mathcal{A}$ . Since  $C_c^2(\mathbb{R}^2) \times C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1}) \times C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$  admits a dense countable subset, we can apply Theorem 6.1.3 of [26] to our framework and we get a negligible set  $N \in \mathcal{G}_\tau$  such that for any  $(F, f, g) \in C_c^2(\mathbb{R}^2) \times C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1}) \times C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$ , the process  $(\bar{M}_s^{t, \alpha, F, f, g})_{s \in [\tau(\bar{x}'), T]}$  is a  $\mathbb{G}$ -martingale under  $\bar{\mathbb{P}}_{\bar{x}'}^{t, 1, \bar{x}, \alpha}$  for any  $\bar{x}' \in \mathbf{X} \setminus N$ . We notice that

$$\bar{\mathbb{P}}_{\bar{x}'}^{t, 1, \bar{x}, \alpha} \left( \left\{ \bar{x}'' \in \mathbf{X} : \hat{\lambda}^{\alpha^{\tau(\bar{x}'), 2\bar{x}'}}(s, \bar{x}'') = \hat{\lambda}^\alpha(s, \bar{x}'') \text{ for all } s \in [\tau(\bar{x}'), T] \right\} \right) = 1$$

and

$$\bar{\mathbb{P}}_{\bar{x}'}^{t, 1, \bar{x}, \alpha} \left( \left\{ \bar{x}'' \in \mathbf{X} : \hat{\sigma}^{\alpha^{\tau(\bar{x}'), 2\bar{x}'}}(s, \bar{x}'') = \hat{\sigma}^\alpha(s, \bar{x}'') \text{ for all } s \in [\tau(\bar{x}'), T] \right\} \right) = 1$$

for any  $\bar{x}' \in \mathbf{X} \setminus N$ . Therefore, for any  $(F, f, g) \in C_c^2(\mathbb{R}^2) \times C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1}) \times C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$ , the process  $(\bar{M}_s^{\tau(\bar{x}'), \alpha^{\tau(\bar{x}'), 2\bar{x}'}, F, f, g})_{s \in [\tau(\bar{x}'), T]}$  is a  $\mathbb{G}$ -martingale under  $\bar{\mathbb{P}}_{\bar{x}'}^{t, 1, \bar{x}, \alpha}$  for any  $\bar{x}' \in \mathbf{X} \setminus N$ . By uniqueness to the shifted controlled martingale problem with initial condition  $(\tau(\bar{x}'), \bar{x}')$  and control  $\alpha^{\tau(\bar{x}'), 2\bar{x}'}$ , we get

$$\bar{\mathbb{P}}_{\bar{x}'}^{t, 1, \bar{x}, \alpha} = \mathbb{P}^{\tau(\bar{x}'), \bar{x}', \alpha^{\tau(\bar{x}'), 2\bar{x}'}}$$

for any  $\bar{x}' \in \mathbf{X} \setminus N$ . □

**Theorem A.7.** *Under Assumption A1, the martingale problem and the shifted martingale problem admit unique solutions for any initial condition  $(t, \bar{x}) \in [0, T] \times \mathbf{X}$  and any control  $\alpha \in \mathcal{A}$ .*

To prove Theorem A.7, we need to consider an extended process  $\mathfrak{r}$  defined by

$$\mathfrak{r}_s = (s, (\mathbf{x}_{u \wedge s})), \quad s \in [t, T]$$

The process  $\mathfrak{r}$  is valued in  $\mathfrak{X} = \mathbb{R} \times \mathbf{X}$  which is separable. We introduce the domain  $\mathcal{D}$  as the set of function  $h : \mathfrak{X} \rightarrow \mathbb{R}$  of the form

$$h(s, \bar{x}) = H \left( F_{f^1, g^1}^1(s, \bar{x}_{s \wedge t_1}), \dots, F_{f^p, g^p}^p(s, \bar{x}_{s \wedge t_p}) \right), \quad (s, \bar{x}) \in \mathfrak{X},$$

for some  $p \geq 1$ ,  $0 \leq t_1 < \dots < t_p \leq T$ ,  $H \in C^2(\mathbb{R}^p)$ ,  $F^1, \dots, F^p \in C_c^{1,2}(\mathbb{R}^2)$ ,  $f^1, \dots, f^p \in C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{R}^{d+1})$  and  $g^1, \dots, g^p \in C_c^{1,2}([0, T] \times \mathcal{I} \times \mathbb{N} \times \mathbb{R}^{m+1})$ . We then define on  $\mathcal{D}$  the operator  $\mathbf{L}^{t, \alpha}$  by

$$\begin{aligned} \mathbf{L}^{t, \alpha} h(s, \bar{x}) &= \mathfrak{L}^{t, \alpha}(s, \bar{x}) \cdot DH(h^1(s, \bar{x}_{s \wedge t_1}), \dots, h^p(s, \bar{x}_{s \wedge t_p})) \\ &+ \frac{1}{2} \sum_{i \in \mathcal{I}} \text{Tr} \left( \mathfrak{G}^\alpha (\mathfrak{G}^\alpha)^\top (s, \bar{x}) D^2 H(F_{f^1, g^1}^1(s, \bar{x}_{s \wedge t_1}), \dots, F_{f^p, g^p}^p(s, \bar{x}_{s \wedge t_p})) \right) \\ &+ \sum_{j=1}^p \mathbb{1}_{t_{j-1} < s \leq t_j} \sum_{i \in \mathcal{I}} \sum_{k \geq 0} \gamma p k \\ &\left( H \left( F_{f^1, g^1}^1(s, \bar{x}_{t_1}), \dots, F_{f^{j-1}, g^{j-1}}^{j-1}(s, \bar{x}_{t_{j-1}}), \mathfrak{G}_{i, k} F_{f^j, g^j}^j(s, \bar{x}_s), \dots, \mathfrak{G}_{i, k} F_{f^p, g^p}^p(s, \bar{x}_s) \right) \right. \\ &\quad \left. - H \left( F_{f^1, g^1}^1(s, \bar{x}_{s \wedge t_1}), \dots, F_{f^p, g^p}^p(s, \bar{x}_{s \wedge t_p}) \right) \right) \end{aligned}$$

with  $t_0 = 0$ , where

$$\mathfrak{L}^{t,\alpha}(s, \bar{x}) = \begin{pmatrix} \mathbf{1}_{s \leq t_1} \mathfrak{L}^{t,\alpha,1}(s, \bar{x}) \\ \vdots \\ \mathbf{1}_{s \leq t_1} \mathfrak{L}^{t,\alpha,p}(s, \bar{x}) \end{pmatrix}$$

with

$$\begin{aligned} \mathfrak{L}^{t,\alpha,q}(s, \bar{x}) &= \bar{L}^\alpha F_{f^q, g^q}^q(s, \bar{x}) \\ &\quad - \gamma \sum_{i \in \mathcal{I}, n \in \mathbb{N}} p_n \left\{ F^q \left( f^q(s, {}^1\bar{x}_s) + \mathbf{1}_{\mathcal{V}_s}(i) \sum_{\ell=0}^{n-1} (f_{i\ell}^q - f_i^q)(s, \hat{x}_s^i), \right. \right. \\ &\quad \left. \left. g^q(s, {}^2\bar{x}_s) + \frac{1}{2^{2(|i|+n)}} (g^q(i, n, s, b_s^i, q_s^{i,n} + 1) - g^q(i, n, s, b_s^i, q_s^{i,n})) \right) \right. \\ &\quad \left. - F^q(f^q(s, {}^1\bar{x}_s), g^q(s, {}^2\bar{x}_s)) \right\} \end{aligned}$$

and

$$\mathfrak{S}^{t,\alpha}(s, \bar{x}) = \begin{pmatrix} \mathfrak{S}^{t,\alpha,1}(s, \bar{x}) \\ \vdots \\ \mathfrak{S}^{t,\alpha,p}(s, \bar{x}) \end{pmatrix}$$

with

$$\begin{aligned} \mathfrak{S}^{t,\alpha,q}(s, \bar{x}) &= \sum_{i \in \mathcal{I}} \mathbf{1}_{s \leq t_q} \left( \partial_2 F^q(f^q(s, {}^1\bar{x}_s), g^q(s, {}^2\bar{x}_s)) \sum_{n \in \mathbb{N}} \frac{1}{2^{2(|i|+n)}} \partial_b g_{i,n}^q(s, b_s^i, q_s^{i,n}) \right. \\ &\quad \left. + \mathbf{1}_{i \in \mathcal{V}_s} \partial_1 F^q(f^q(s, {}^1\bar{x}_s), g^q(s, {}^2\bar{x}_s)) D f_i^q(s, \hat{x}_s^i)^\top \hat{\sigma}(\hat{x}_s^i, \alpha(s, {}^2\bar{x})) \right) \end{aligned}$$

for  $q = 1, \dots, p$ ,  $(s, \bar{x}) \in [0, T] \times \mathbf{X}$  and

$$\begin{aligned} \mathfrak{S}_{i,n} F_{f^j, g^j}^j(s, \bar{x}) &= F \left( f(s, {}^1\bar{x}_s) + \mathbf{1}_{\mathcal{V}_s}(i) \sum_{\ell=0}^{n-1} (f_{i\ell} - f_i)(s, \hat{x}_s^i), \right. \\ &\quad \left. g(s, {}^2\bar{x}_s) + \frac{1}{2^{2(|i|+n)}} (g(i, n, s, b_s^i, q_s^{i,n} + 1) - g(i, n, s, b_s^i, q_s^{i,n})) \right) \end{aligned}$$

for  $(s, \bar{x}) \in [0, T] \times \mathbf{X}$ ,  $i \in \mathcal{I}$  and  $n \geq 0$ . We then notice that for  $\bar{\mathbb{P}}^{t, {}^1\bar{x}, \alpha}$  (resp.  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}$ ) solution to the martingale problem (resp. shifted martingale problem) with initial condition  $(t, {}^1\bar{x})$  (resp.  $(t, \bar{x})$ ) and control  $\alpha$  the process

$$h(\mathbf{x}_s) - \int_t^s \mathbf{L}^{t,\alpha} h(\mathbf{x}_u) du, \quad t \leq u \leq T,$$

is a  $\mathbb{G}$ -martingale under  $\bar{\mathbb{P}}^{t, {}^1\bar{x}, \alpha}$  (resp.  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}$ )

**Lemma A.3.** *Let  $(t, \bar{x}, \alpha) \in [0, T] \times \mathbf{X} \times \mathcal{A}$  and  $\bar{\mathbb{P}}_1^{t, {}^1\bar{x}, \alpha}$  and  $\bar{\mathbb{P}}_2^{t, {}^1\bar{x}, \alpha}$  (resp.  $\bar{\mathbb{P}}_1^{t, \bar{x}, \alpha}$  and  $\bar{\mathbb{P}}_2^{t, \bar{x}, \alpha}$ ) two solutions to the martingale problem (resp. shifted martingale problem) with initial condition  $(t, {}^1\bar{x})$  (resp.  $(t, \bar{x})$ ) and control  $\alpha$ . Then,  $\bar{\mathbb{P}}_1^{t, {}^1\bar{x}, \alpha}$  and  $\bar{\mathbb{P}}_2^{t, {}^1\bar{x}, \alpha}$  have the same one dimensional marginals:*

$$\bar{\mathbb{P}}_1^{t, {}^1\bar{x}, \alpha}(\mathbf{x}_s \in B) = \bar{\mathbb{P}}_2^{t, {}^1\bar{x}, \alpha}(\mathbf{x}_s \in B) \tag{A.87}$$

$$\text{(resp. } \bar{\mathbb{P}}_1^{t, \bar{x}, \alpha}(\mathbf{x}_s \in B) = \bar{\mathbb{P}}_2^{t, \bar{x}, \alpha}(\mathbf{x}_s \in B) \tag{A.88}$$

for  $s \in [t, T]$  and  $B \in \mathcal{B}(\mathfrak{X})$ .

*Proof.* We first endow the measurable space  $(\mathbf{X} \times \mathbf{X}, \mathcal{G}_T \otimes \mathcal{G}_T)$  with the probability measure  $\bar{\mathbb{P}} = \bar{\mathbb{P}}_1^{t, \bar{x}, \alpha} \otimes \bar{\mathbb{P}}_2^{t, \bar{x}, \alpha}$  (resp.  $\bar{\mathbb{P}} = \bar{\mathbb{P}}_1^{t, \bar{x}, \alpha} \otimes \bar{\mathbb{P}}_2^{t, \bar{x}, \alpha}$ ). For  $h \in \mathcal{D}$ , we have

$$\mathbb{E}^{\bar{\mathbb{P}}} [h \otimes h(\mathbf{x}_s, \mathbf{x}_t)] = \mathbb{E}^{\bar{\mathbb{P}}} [h \otimes h(\mathbf{x}_t, \mathbf{x}_s)]$$

Indeed, the processes

$$h \otimes h(\mathbf{x}_s, \mathbf{x}_t) - \int_t^s \mathbf{L}^{t, \alpha} h(\mathbf{x}_u) h(\mathbf{x}_t) du, \quad t \leq s \leq T$$

and

$$h \otimes h(\mathbf{x}_t, \mathbf{x}_s) - \int_t^s h(\mathbf{x}_t) \mathbf{L}^{t, \alpha} h(\mathbf{x}_u) du, \quad t \leq s \leq T$$

are both martingales under  $\bar{\mathbb{P}}$ . Since all the considered functions are bounded, we can take the expectation and we get

$$\mathbb{E}^{\bar{\mathbb{P}}} [h \otimes h(\mathbf{x}_t, \mathbf{x}_s)] = \mathbb{E}^{\bar{\mathbb{P}}} [h \otimes h(\mathbf{x}_s, \mathbf{x}_t)]$$

and

$$\begin{aligned} \mathbb{E}^{\bar{\mathbb{P}}_1^{t, \bar{x}, \alpha}} [h(\mathbf{x}_s)] &= \mathbb{E}^{\bar{\mathbb{P}}_2^{t, \bar{x}, \alpha}} [h(\mathbf{x}_s)] \\ (\text{resp. } \mathbb{E}^{\bar{\mathbb{P}}_1^{t, \bar{x}, \alpha}} [h(\mathbf{x}_s)] &= \mathbb{E}^{\bar{\mathbb{P}}_2^{t, \bar{x}, \alpha}} [h(\mathbf{x}_s)] ) . \end{aligned}$$

Since any bounded  $\mathcal{B}(\mathcal{X})$ -measurable function can be approximated almost everywhere for  $\bar{\mathbb{P}}_1^{t, \bar{x}, \alpha}$  and  $\bar{\mathbb{P}}_2^{t, \bar{x}, \alpha}$  (resp.  $\bar{\mathbb{P}}_1^{t, \bar{x}, \alpha}$  and  $\bar{\mathbb{P}}_2^{t, \bar{x}, \alpha}$ ) by a sequence of  $\mathcal{D}$  we get (A.87) (resp. (A.88)).  $\square$

*Proof of Theorem A.7.* The proof is a direct consequence of Theorem 4.2 in [9] and Lemma A.3.  $\square$

### A.3 Proof of Theorem 2.1

We keep the notations of Section A.2. Fix  $(t, \hat{\mu}, \alpha) \in [0, T] \times E_{d+1} \times \mathcal{A}$ . From Proposition 2.1, the law  $\mathcal{L}^{\mathbb{P}}(\hat{Z}^{t, \hat{\mu}, \alpha}, \xi)$  under  $\mathbb{P}$  of  $(\hat{Z}^{t, \hat{\mu}, \alpha}, \xi)$  provides a solution to the controlled martingale problem with initial condition  $(t, \bar{x})$ , where  $\bar{x} \in \mathbf{X}$  such that  ${}^1\bar{x}_s = \hat{\mu}$  for  $s \in [0, T]$ , and control  $\alpha$  given by Definition A.3. Therefore, we get from Theorem A.7

$$\mathcal{L}^{\mathbb{P}}(\hat{Z}^{t, \hat{\mu}, \alpha}, \xi) = \bar{\mathbb{P}}^{t, \bar{x}, \alpha} \tag{A.89}$$

In the same way, for  $\beta \in \mathcal{A}$ ,  $\bar{\omega} \in \Omega$  such that  $\xi(\bar{\omega}) = {}^2\bar{x}$ , the law  $\mathcal{L}^{\mathbb{P}}(\hat{Z}^{t, \hat{\mu}, \beta}, \xi(\bar{\omega} \oplus_t \cdot))$  under  $\mathbb{P}$  of  $(\hat{Z}^{t, \hat{\mu}, \beta}, \xi(\bar{\omega} \oplus_t \cdot))$  is the unique solution to the shifted controlled martingale problem with initial condition  $(t, \bar{x})$  and control  $\beta$  given by Definition A.4. Therefore, we also get from Theorem A.7 that

$$\mathcal{L}^{\mathbb{P}}(\hat{Z}^{t, \hat{\mu}, \beta}, \xi(\bar{\omega} \oplus_t \cdot)) = \bar{\mathbb{P}}^{t, \bar{x}, \beta} . \tag{A.90}$$

Fix now an  $\mathcal{F}_\tau$ -measurable random variable  $Y$ .

From Doob's functional representation Theorem (see Lemma 1.13 in [16]) there exists a random time  $\tilde{\tau} : \mathbb{D}([0, T], \tilde{E}_{m+1}) \rightarrow \mathbb{R}_+$  that is a stopping time with respect to the filtration generated by the canonical process on  $\mathbb{D}([0, T], \tilde{E}_{m+1})$ , and a measurable function  $g_Y : \mathbb{D}([0, T], \tilde{E}_{m+1}) \rightarrow \mathbb{R}$  such that

$$\tau(\omega) = \tilde{\tau}(\xi(\omega)) \quad \text{and} \quad Y(\omega) = g_Y(\xi(\omega_{\tau(\omega)\wedge \cdot})) = g_Y(\xi(\omega)) .$$

We then define  $\bar{\tau} : \mathbf{X} \rightarrow \mathbb{R}_+$  by  $\bar{\tau} = \tilde{\tau} \circ {}^2\mathbf{x}$  where we recall that  ${}^1\mathbf{x}$  and  ${}^2\mathbf{x}$  are given by (A.86). We observe that  $\bar{\tau}$  is a  $\mathbb{G}$ -stopping time and  $g_Y \circ {}^2\mathbf{x}$  is  $\mathcal{G}_{\bar{\tau}}$ -measurable. We therefore have from (A.89)

$$\begin{aligned} \mathbb{E} \left[ f \left( \hat{Z}^{t, \hat{\mu}, \alpha} \right) Y \right] &= \mathbb{E}^{\bar{\mathbb{P}}^{t, 1\bar{x}, \alpha}} \left[ f \left( {}^1\mathbf{x} \right) g_Y \left( {}^2\mathbf{x} \right) \right] \\ &= \int_{\mathbf{X}} \mathbb{E}^{\bar{\mathbb{P}}_{\bar{x}'}^{t, 1\bar{x}, \alpha}} \left[ f \left( {}^1\mathbf{x} \right) \right] g_Y \left( {}^2\mathbf{x}(\bar{x}') \right) d\bar{\mathbb{P}}^{t, 1\bar{x}, \alpha}(\bar{x}') . \end{aligned}$$

where  $(\bar{\mathbb{P}}_{\bar{x}'}^{t, \bar{x}, \alpha}, \bar{x}' \in \mathbf{X})$  stands for a *regular conditional probability distribution* of  $\bar{\mathbb{P}}^{t, \bar{x}, \alpha}$  given  $\mathcal{G}_{\bar{\tau}}$ . Using Theorem A.6 we finally get

$$\begin{aligned} \mathbb{E} \left[ f \left( \hat{Z}^{t, \hat{\mu}, \alpha} \right) Y \right] &= \int_{\mathbf{X}} \mathbb{E}^{\bar{\mathbb{P}}^{\bar{\tau}(\bar{x}'), \bar{x}', \alpha^{\bar{\tau}(\bar{x}'), 2\bar{x}'}}} \left[ f \left( {}^1\mathbf{x} \right) \right] g_Y \left( {}^2\mathbf{x}(\bar{x}') \right) d\bar{\mathbb{P}}^{t, 1\bar{x}, \alpha}(\bar{x}') \\ &= \int_{\mathbf{X}} F \left( \bar{\tau}(\bar{x}'), \bar{x}', \alpha^{\bar{\tau}(\bar{x}'), 2\bar{x}'} \right) g_Y \left( {}^2\mathbf{x}(\bar{x}') \right) d\bar{\mathbb{P}}^{t, 1\bar{x}, \alpha}(\bar{x}') . \end{aligned}$$

Then, using (A.90), we get

$$\mathbb{E} \left[ f \left( \hat{Z}^{t, \hat{\mu}, \alpha} \right) Y \right] = \int_{\Omega} F \left( \tau(\omega), \hat{Z}_{\cdot \wedge \tau}^{t, \hat{\mu}, \alpha}(\omega), \alpha^{\tau(\omega), \omega} \right) Y(\omega) d\mathbb{P}(\omega) .$$

## References

- [1] D. P. Bertsekas and S. E. Shreve. *Stochastic optimal control: the discrete time case*. Optimization and Neural Computation Series. Athena Scientific, 2007.
- [2] B. Bouchard. Stochastic targets with mixed diffusion processes and viscosity solutions. *Stochastic Processes and their Applications*, 101(2):273–302, 2002.
- [3] B. Bouchard, B. Djehiche, and I. Kharroubi. Quenched mass transport of particles towards a target. *Journal of Optimization Theory and Applications*, 186 (2):365–374, 2020.
- [4] B. Bouchard, R. Elie, and C. Imbert. Optimal Control under Stochastic Target Constraints. *SIAM Journal on Control and Optimization*, 48(5):3501–3531, 2010.
- [5] J. Claisse. Optimal control of branching diffusion processes: A finite horizon problem. *The Annals of Applied Probability*, 28 (1):1–34, 2018.
- [6] J. Claisse, D. Talay, and X. Tan. A pseudo-markov property for controlled diffusion processes. *SIAM Journal on Control and Optimization*, 54 (2):1017–1029, 2016.
- [7] M. G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bulletin of the American Mathematical Society*, 1992.
- [8] D. A. Dawson. Measure-valued markov processes. In *Ecole d’été de probabilités de Saint-Flour XXI*, number 1541 in Lecture Notes in Math., pages 1–260. Springer Berlin, 1993.
- [9] S. N. Ethier and T. G. Kurtz. *Markov Processes, Characterization and Convergence*. John Wiley & Sons, 1986.
- [10] S. Fahad. Blockchain without Waste: Proof-of-Stake. *The Review of Financial Studies*, 34(3):1156–1190, 2021.

- [11] W. H. Fleming and H. M. Soner. *Controlled Markov Processes and Viscosity Solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer Berlin, second edition edition, 2006.
- [12] P. Henry-Labordère, X. Tan, and N. Touzi. A numerical algorithm for a class of bsdes via the branching process. *Stochastic Process. Appl.*, 124:1112–1140, 2014.
- [13] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes I. *I. J. Math. Kyoto Univ.*, 8:233–278, 1968.
- [14] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes II. *I. J. Math. Kyoto Univ.*, pages 365–410, 1968.
- [15] N. Ikeda, M. Nagasawa, and S. Watanabe. Branching Markov processes III. *I. J. Math. Kyoto Univ.*, 9:95–160, 1969.
- [16] O. Kallenberg. *Foundation of Modern Probability*. Probability and its Applications. Springer-Verlag New York, second edition edition, 2002.
- [17] O. Kallenberg. *Random Measures, Theory and Applications*, volume 77 of *Probability Theory and Stochastic Modelling*. Springer International Publishing Switzerland, 2017.
- [18] N. V. Krylov. *Controlled Diffusion Processes*. Number 14 in Application of Mathematics. Springer-Verlag, 1980.
- [19] N. V. Krylov. *Nonlinear Elliptic and Parabolic Equations of the Second Order*, volume 7 of *Mathematics and Its Applications (Soviet Series)*. Reidel Dordrecht, 1987.
- [20] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. , 2008.
- [21] M. Nisio. Stochastic control related to branching diffusion processes. *J. Math. Kyoto Univ.*, 25:549–575, 1985.
- [22] B. Oksendal and A. Sulem. *Applied Stochastic Control for Jump Diffusions, Second Edition*. Universitext. Springer-Verlag Berlin Heidelberg, 2007.
- [23] A. V. Skorohod. Branching diffusion processes. *Teor. Veroyatnost. i Primenen.*, 9(3):492–497, 1964.
- [24] H. Mete Soner and Nizar Touzi. Dynamic programming for stochastic target problems and geometric flows. *Journal of the European Mathematical Society*, 4:201–236, 09 2002.
- [25] H. Mete Soner and Nizar Touzi. Stochastic Target Problems, Dynamic Programming, and Viscosity Solutions . *SIAM Journal on Control and Optimization*, 41(2):404–424, 2002.
- [26] D. Stroock and S.R.S. Varadhan. *Multidimensional Diffusion Processes*. Reprint of the 1997 Edition, Classics in Mathematics. Springer, 1997.
- [27] S. Ustunel. Construction of branching diffusion processes and their optimal stochastic control. *Appl. Math. Optim.*, 7:11–33, 1981.