BSDE Representation and Discretization for Hamilton-Jacobi-Bellman PDE

Idris KHARROUBI
CEREMADE, CNRS, UMR 7534
Université Paris Dauphine
and CREST

March 31, 2014
1 Introduction 2

1 BSDE representation for HJB equations 8

2 BSDE with nonpositive jumps 9
  2.1 Formulation and assumptions 9
  2.2 Existence and approximation by penalization 11
  2.3 Dual representation 15

3 Nonlinear IPDE and Feynman-Kac formula 19
  3.1 The Markovian framework 19
  3.2 Viscosity property of the penalized BSDE 25
  3.3 The non dependence of the function \( v \) in the variable \( a \) 29
  3.4 Viscosity properties of the minimal solution to the constrained BSDE . . 34

II Discretization of fully nonlinear HJB equations via BSDEs with nonpositive jumps 39

4 Discretization of the nonpositive jump constraint 40
  4.1 Discretely jump-constrained BSDE 40
  4.2 Convergence of discretely jump-constrained BSDE 46
    4.2.1 Convergence result 46
    4.2.2 Rate of convergence 51

5 Approximation scheme for jump-constrained BSDE and stochastic control problem 59
  5.1 The forward regime switching process 59
  5.2 BSDE scheme 62
  5.3 Error estimate for the discrete time scheme 63
  5.4 Approximate optimal control 66
Chapter 1

Introduction

Let us consider the Hamilton-Jacobi-Bellman (HJB) equation in the form:

\[
\frac{\partial v}{\partial t} + \sup_{a \in A} \left[ b(x, a) \cdot D_x v + \frac{1}{2} \text{tr} (\sigma^T(x, a) D_x^2 v) + f(x, a) \right] = 0, \quad \text{on } [0, T) \times \mathbb{R}^d, \quad (1.0.1)
\]

\[
v(T, x) = g(x), \quad x \in \mathbb{R}^d,
\]

where \(A\) is a subset of \(\mathbb{R}^q\). It is well-known (see e.g. [38]) that such nonlinear PDE is the dynamic programming equation associated to the stochastic control problem with value function defined by:

\[
v(t, x) = \sup_{\alpha} \mathbb{E} \left[ \int_t^T f(X^t,x,\alpha_s, \alpha_s) ds + g(X^t_t, x, \alpha_T) \right], \quad (1.0.2)
\]

where \(X^{t,x,\alpha}\) is the solution to the controlled diffusion:

\[
dX_s^\alpha = b(X_s^\alpha, \alpha_s) ds + \sigma(X_s^\alpha, \alpha_s) dW_s,
\]

starting from \(x\) at \(t\), and given a predictable control process \(\alpha\) valued in \(A\).

Our main goal is to provide a numerical scheme for the approximation of the nonlinear HJB equation using Backward Stochastic Differential Equation (BSDEs). To this end one has first to built a probabilistic representation via BSDEs, namely the so-called nonlinear Feynman-Kac formula, which involves a simulatable forward process. One can then hope to use such representation for deriving a probabilistic numerical scheme for the solution to HJB equation, whence the stochastic control problem. Such issues have attracted a lot of interest and generated an important literature over the recent years. Actually, there is a crucial distinction between the case where the diffusion coefficient is controlled or not.

Consider first the case where \(\sigma(x)\) does not depend on \(a \in A\), and assume that \(\sigma\sigma^T(x)\) is of full rank. Denoting by \(\theta(x,a) = \sigma^T(x)(\sigma\sigma^T(x))^{-1}b(x,a)\) a solution to \(\sigma(x)\theta(x,a) = b(x,a)\), we notice that the HJB equation reduces into a semi-linear PDE:

\[
\frac{\partial v}{\partial t} + \frac{1}{2} \text{tr} (\sigma^T(x) D_x^2 v) + F(x, \sigma^T D_x v) = 0, \quad (1.0.3)
\]

where \(F(x, z) = \sup_{a \in A} [f(x,a) + \theta(x,a) \cdot z]\) is the \(\theta\)-Fenchel-Legendre transform of \(f\). In this case, we know from the seminal works by Pardoux and Peng [33, 34], that the (viscosity)
solution $v$ to the semi-linear PDE (1.0.3) is connected to the BSDE:

$$Y_t = g(X^0_T) + \int_t^T F(X^0_s, Z_s)ds - \int_t^T Z_s dW_s, \quad t \leq T,$$

through the relation $Y_t = v(t, X^0_t)$, with a forward diffusion process

$$dX^0_s = \sigma(X^0_s)dW_s.$$

This probabilistic representation leads to a probabilistic numerical scheme for the resolution of (1.0.3) by discretization and simulation of the BSDE (1.0.4), see [9]. Alternatively, when the function $F(x, z)$ is of polynomial type on $z$, the semi-linear PDE (1.0.3) can be numerically solved by a forward Monte-Carlo scheme relying on marked branching diffusion, as recently pointed out in [22]. Moreover, as showed in [15], the solution to the BSDE (1.0.4) admits a dual representation in terms of equivalent change of probability measures as:

$$Y_t = \text{ess sup}_\alpha E^P_\alpha \left[ \int_t^T f(X^0_s, \alpha_s)ds + g(X^0_T) \bigg| \mathcal{F}_t \right],$$

where for a control $\alpha$, $P^\alpha$ is the equivalent probability measure to $P$ under which

$$dX^0_s = b(X^0_s, \alpha_s)ds + \sigma(X^0_s)dW^\alpha_s,$$

with $W^\alpha$ a $P^\alpha$-Brownian motion by Girsanov’s theorem. In other words, the process $X^0$ has the same dynamics under $P^\alpha$ than the controlled process $X^\alpha$ under $P$, and the representation (1.0.5) can be viewed as a weak formulation (see [14]) of the stochastic control problem (1.0.2) in the case of uncontrolled diffusion coefficient.

The general case with controlled diffusion coefficient $\sigma(x, a)$ associated to fully nonlinear PDE is challenging and led to recent theoretical advances. Consider the motivating example from uncertain volatility model in finance formulated here in dimension 1 for simplicity of notations:

$$dX^\alpha_s = \alpha_s dW_s,$$

where the control process $\alpha$ is valued in $A = [\underline{a}, \bar{a}]$ with $0 \leq \underline{a} \leq \bar{a} < \infty$, and define the value function of the stochastic control problem:

$$v(t, x) := \sup_\alpha \mathbb{E}[g(X^{t,x,\alpha}_T)], \quad (t, x) \in [0, T] \times \mathbb{R}.$$

The associated HJB equation takes the form:

$$\frac{\partial v}{\partial t} + G(D_x^2 v) = 0, \quad (t, x) \in [0, T) \times \mathbb{R}, \quad v(T, x) = g(x), \quad x \in \mathbb{R},$$

where $G(M) = \frac{1}{2} \sup_{a \in A} [a^2 M] = \bar{a}^2 M^+ - \underline{a}^2 M^-$. The unique (viscosity) solution to (1.0.6) is represented in terms of the so-called $G$-Brownian motion $B$, and $G$-expectation $\mathbb{E}_G$, concepts introduced in [36]:

$$v(t, x) = \mathbb{E}_G [g(x + B_{T-t})].$$
Moreover, G-expectation is closely related to second order BSDE studied in [41], namely the process \( Y_t = v(t, B_t) \) satisfies a 2BSDE, which is formulated under a non-dominated family of singular probability measures given by the law of \( X^\alpha \) under \( \mathbb{P} \). This gives a nice theory and representation for nonlinear PDE, but it requires a non degeneracy assumption on the diffusion coefficient, and does not cover general HJB equation (i.e. control both on drift and diffusion arising for instance in portfolio optimization). On the other hand, it is not clear how to simulate G-Brownian motion.

We provide here an alternative BSDE representation including general HJB equation, formulated under a single probability measure (thus avoiding non-dominated measures), and under which the forward process can be simulated. The idea, used in [25] for quasi variational inequalities arising in impulse control problems, is the following. We introduce a Poisson random measure \( \mu(dt, da) \) on \( \mathbb{R}_+ \times A \) with finite intensity measure \( \lambda(da)dt \) associated to the marked point process \( (\tau_i, \zeta_i) \), independent of \( W \), and consider the pure jump process \( (I_t) \) equal to the mark \( \zeta_i \) valued in \( A \) between two jump times \( \tau_i \) and \( \tau_{i+1} \).

We next consider the forward regime switching diffusion process

\[
dX_s = b(X_s, I_s)ds + \sigma(X_s, I_s)dW_s,
\]

and observe that the (uncontrolled) pair process \( (X, I) \) is Markov. Let us then consider the BSDE with jumps w.r.t the Brownian-Poisson filtration \( \mathbb{F} = \mathbb{F}^{W, \mu} \):

\[
Y_t = g(X_T) + \int_t^T f(X_s, I_s)ds - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \tilde{\mu}(ds, da),
\]

(1.0.7)

where \( \tilde{\mu} \) is the compensated measure of \( \mu \). This linear BSDE is the Feynman-Kac formula for the linear integro-partial differential equation (IPDE):

\[
\frac{\partial v}{\partial t} + b(x, a).D_xv + \frac{1}{2} \text{tr}(\sigma \sigma^T(x, a) D_x^2 v)
\]

\[
+ \int_A (v(t, x, a') - v(t, x, a)) \lambda(da') + f(x, a) = 0, \quad (t, x, a) \in [0, T) \times \mathbb{R}^d \times A,
\]

(1.0.8)

\[
v(T, x, a) = g(x), \quad (x, a) \in \mathbb{R}^d \times A,
\]

(1.0.9)

through the relation: \( Y_t = v(t, X_t, I_t) \). Now, in order to pass from the above linear IPDE with the additional auxiliary variable \( a \in A \) to the nonlinear HJB PDE (4.2.18), we constrain the jump component to the BSDE (1.0.7) to be nonpositive, i.e.

\[
U_t(a) \leq 0, \quad \forall (t, a).
\]

(1.0.10)

Then, since \( U_t(a) \) represents the jump of \( Y_t = v(t, X_t, I_t) \) induced by a jump of the random measure \( \mu \), i.e of \( I \), and assuming that \( v \) is continuous, the constraint (1.0.10) means that \( U_t(a) = v(t, X_t, a) - v(t, X_t, I_{t-}) \leq 0 \) for all \( (t, a) \). This formally implies that \( v(t, x) \) should not depend on \( a \in A \). Once we get the non dependence of \( v \) in \( a \), the equation (1.0.8) becomes a PDE on \( [0, T) \times \mathbb{R}^d \) with a parameter \( a \in A \). By taking the supremum over \( a \in A \) in (1.0.8), we then obtain the nonlinear HJB equation (4.2.18).
Inspired by the above discussion, we now introduce the following general class of BSDE with nonpositive jumps, which is a non Markovian extension of (1.0.7)-(1.0.10):

\[ Y_t = \xi + \int_t^T F(s, Y_s, Z_s, U_s)ds + K_T - K_t \]  
\[ - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \tilde{\mu}(ds, da), \quad 0 \leq t \leq T, \text{a.s.} \]  

with

\[ U_t(a) \leq 0, \quad d\mathbb{P} \otimes dt \otimes \lambda(da) \text{ a.e. on } \Omega \times [0,T] \times A. \]  

The solution to this BSDE is a quadruple \((Y, Z, U, K)\) where, besides the usual component \((Y, Z, U)\), the fourth component \(K\) is a predictable nondecreasing process, which makes the \(A\)-constraint (1.0.12) feasible. We thus look at the minimal solution \((Y, Z, U, K)\) in the sense that for any other solution \((\bar{Y}, \bar{Z}, \bar{U}, \bar{K})\) to (1.0.11)-(1.0.12), we must have \(Y \leq \bar{Y}\).

Through a penalization method, we construct the unique minimal solution as the limit of a sequence \((Y^n, Z^n, U^n, K^n)\) of Lipschitz BSDEs with jumps. In a markovian framework the general constrained BSDE takes the following form:

\[ Y_t = g(X_T) + \int_t^T f(X_s, I_s, Y_s, Z_s)ds + K_T - K_t \]  
\[ - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \tilde{\mu}(ds, da), \quad 0 \leq t \leq T, \]  

with

\[ U_t(a) \leq 0, \quad d\mathbb{P} \otimes dt \otimes \lambda(da) \text{ a.e. on } \Omega \times [0,T] \times A. \]  

Its minimal solution is then proved to be the Feynman-Kac representation of the PDE

\[ \frac{\partial v}{\partial t} + \sup_{a \in A} \left[ b(\cdot, a)D_x v + \frac{1}{2} \text{tr} \left( \sigma \sigma^t (\cdot, a)D_x^2 v \right) + f(x, a, v, \sigma^t (x, a)D_x v) \right] = 0 \]

on \([0,T] \times \mathbb{R}^d\), through the relation: \(Y_t = v(t, X_t)\). This equation clearly extends HJB equation (4.2.18) by incorporating the terms \(v\) and \(D_x v\) in the function \(f\).

In the second part of the lectures, we use this representation to set an approximation of solutions to HJB equations. Namely, we provide and analyze a discrete-time approximation scheme for the minimal solution to (1.0.13)-(1.0.14), and thus an approximation scheme for the HJB equation. In the non-constrained jump case, approximations schemes for BSDE have been studied in the papers [21], [8], which extended works in [9], [44] for BSDEs in a Brownian framework. The issue is now to deal with the nonpositive jump constraint in (1.0.13)-(1.0.14), and we propose a discrete time approximation scheme of the form:

\[
\begin{align*}
\bar{Y}_T^\pi &= \bar{Y}_T^\pi = g(\bar{X}_T^\pi) \\
\bar{Z}_t^\pi &= \mathbb{E} \left[ \bar{Y}_{t+1}^\pi - \bar{Y}_t^\pi \left| F_t \right. \right] \\
\bar{Y}_t^\pi &= \mathbb{E} \left[ \bar{Y}_{t+1}^\pi \left| F_t \right. \right] + (t_{k+1} - t_k) f(\bar{X}_{t_k}^\pi, I_{t_k}, \bar{Y}_{t_k}^\pi, \bar{Z}_{t_k}^\pi) \\
\bar{Y}_t^\pi &= \text{ess sup}_{a \in A} \mathbb{E} \left[ \bar{Y}_{t_k}^\pi \left| F_{t_k}, I_{t_k} = a \right. \right], \quad k = 0, \ldots, n - 1,
\end{align*}
\]
where \( \pi = \{ t_0 = 0 < \ldots < t_k < \ldots < t_n = T \} \) is a partition of the time interval \([0, T]\), with modulus \(|\pi|\), and \( \bar{X}^\pi \) is the Euler scheme of \( X \) (notice that \( I \) is perfectly simulatable once we know how to simulate the distribution \( \lambda(da)/\int\lambda(da) \) of the jump marks). The interpretation of this scheme is the following. The first three lines in (1.0.15) correspond to the standard scheme \((\bar{Y}^\pi, \bar{Z}^\pi)\) for a discretization of a BSDE with jumps (see [8]), where we omit here the computation of the jump component. The last line in (1.0.15) for computing the approximation \( \bar{Y}^\pi \) of the minimal solution \( Y \) corresponds precisely to the minimality condition for the nonpositive jump constraint and should be understood as follows. By the Markov property of the forward process \((X, I)\), the solution \((Y, Z, U)\) to the BSDE with jumps (without constraint) is in the form \( Y_t = \vartheta(t, X_t, I_t) \) for some deterministic function \( \vartheta \). Assuming that \( \vartheta \) is a continuous function, the jump component of the BSDE, which is induced by a jump of the forward component \( I \), is equal to \( U_t(a) = \vartheta(t, X_t, a) - \vartheta(t, X_t, I_t-). \) Therefore, the nonpositive jump constraint means that: \( \vartheta(t, X_t, I_t-) \geq \text{ess sup} \vartheta(t, X_t, a) \).

The minimality condition is thus written as:

\[
Y_t = \nu(t, X_t) = \text{ess sup} \vartheta(t, X_t, a) = \text{ess sup} \mathbb{E}[Y_t|X_t, I_t = a],
\]

whose discrete time version is the last line in scheme (1.0.15). We mainly consider the case where \( f(x, a, y) \) does not depend on \( z \), and our aim is to analyze the discrete time approximation error on \( Y \), where we split the error between the positive and negative parts:

\[
\text{Err}^+ (Y) := \left( \max_{k \leq n-1} \mathbb{E}\left[ \left( Y_{t_k} - \bar{Y}^\pi_{t_k}\right)^2 \right] \right)^{\frac{1}{2}}, \quad \text{Err}^- (Y) := \left( \max_{k \leq n-1} \mathbb{E}\left[ \left( Y_{t_k} - \bar{Y}^\pi_{t_k}\right)^2 \right] \right)^{\frac{1}{2}}.
\]

We do not study directly the error on \( Z \), and instead focus on the approximation of an optimal control for the HJB equation, which is more relevant in practice. It appears that the maximization step in the scheme (1.0.15) provides a control in feedback form \( \{a(t_k, X^\pi_{t_k})\}, k \leq n - 1 \), which approximates the optimal control with an estimated error bound. The analysis of the error on \( Y \) proceeds as follows. We first introduce the solution \((\bar{Y}^\pi, \bar{Y}^\pi, \bar{Z}^\pi, \bar{U}^\pi)\) of a discretely jump-constrained BSDE. This corresponds formally to BSDEs for which the nonpositive jump constraint operates only a finite set of times, and should be viewed as the analog of discretely reflected BSDEs defined in [1] and [7] in the context of the approximation for reflected BSDEs. By combining BSDE methods and PDE approach with comparison principles, and further with the shaking coefficients method of Krylov [29] and Barles, Jacobsen [5], we prove the monotone convergence of this discretely jump-constrained BSDE towards the minimal solution to the BSDE with nonpositive jump constraint. We also obtained a convergence rate without any ellipticity condition on the diffusion coefficient \( \sigma \). We next focus on the approximation error between the discrete time scheme in (1.0.15) and the discretely jump-constrained BSDE. The standard argument for studying rate of convergence of such error consists in getting an estimate of the error at time \( t_k \): \( \mathbb{E}[|Y^\pi_{t_k} - \bar{Y}^\pi_{t_k}|^2] \) in function of the same estimate at time \( t_{k+1} \), and then conclude by induction together with classical estimates for the forward Euler scheme.

However, due to the supremum in the conditional expectation in the scheme (1.0.15) for passing from \( \bar{Y}^\pi \) to \( \bar{Y}^\pi \), such argument does not work anymore. Indeed, taking the supremum is a nonlinear operation, which violates the law of iterated conditional expectations.
Therefore, we cannot obtain directly the error at time $t_k$ as a function of that at time $t_{k+1}$. Instead, we consider the auxiliary error control at time $t_k$:

$$\mathcal{E}_k^\pi(Y) := \mathbb{E}\left[\text{ess sup}_{a \in A} \mathbb{E}_{t_1,a} \left[ \ldots \text{ess sup}_{a \in A} \mathbb{E}_{t_k,a} \left[ |Y_{t_k}^\pi - \bar{Y}_{t_k}^\pi|^2 \ldots \right] \right] \right],$$

where $\mathbb{E}_{t_k,a}[]$ denotes the conditional expectation $\mathbb{E}[, |F_{t_k}, I_{t_k} = a|$, and we are able to express $\mathcal{E}_k^\pi(Y)$ in function of $\mathcal{E}_{k+1}^\pi(Y)$. We define similarly an error control $\mathcal{E}_k^\pi(X)$ for the forward Euler scheme, and prove that it converges to zero with a rate $|\pi|$. Proceeding by induction, we then obtain a rate of convergence $|\pi|$ for $\mathcal{E}_k^\pi(Y)$, and consequently for $\mathbb{E}[|Y_{t_k}^\pi - \bar{Y}_{t_k}^\pi|^2]$. This leads finally to a rate $|\pi|^{\frac{3}{2}}$ for $\text{Err}_\pi(Y)$, $|\pi|^{\frac{10}{11}}$ for $\text{Err}_\pi^+(Y)$, and so $|\pi|^{\frac{1}{10}}$ for the global error $\text{Err}_\pi(Y) = \text{Err}_\pi^+(Y) + \text{Err}_\pi^-(Y)$. In fact, as noticed in Remark 5.3.4, we believe that one can obtain a better rate at least of the order $|\pi|^{\frac{1}{6}}$. Anyway, our result improves the convergence rate of the mixed Monte-Carlo finite difference scheme proposed in [17], where the authors obtained a rate $|\pi|^{\frac{1}{4}}$ on one side and $|\pi|^{\frac{1}{10}}$ on the other side under a nondegeneracy condition.

We conclude this introduction by pointing out that the above discrete time scheme is not yet directly implemented in practice, and requires the estimation and computation of the conditional expectations together with the supremum. Actually, simulation-regression methods on basis functions defined on $\mathbb{R}^d \times A$ appear to be very efficient, and provide approximate optimal controls in feedback forms via the maximization operation in the last step of the scheme (1.0.15). We refer to [24] for analysis and illustrations with several numerical tests arising in superreplication of options under uncertain volatility and correlation. Notice that since it relies on the simulation of the forward process $(X, I)$, our scheme does not suffer the curse of dimensionality encountered in finite difference scheme or controlled Markov chains methods (see [30], [6]), and takes advantage of the high-dimensional properties of Monte-Carlo methods.
Part I

BSDE representation for HJB equations
Chapter 2

BSDE with nonpositive jumps

We introduce in this chapter a new class of BSDEs driven by a Brownian motion and a Poisson, where the jump component of the solution is subject to a constraint.

2.1 Formulation and assumptions

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a complete probability space on which are defined a \(d\)-dimensional Brownian motion \(W = (W_t)_{t \geq 0}\), and an independent integer valued Poisson random measure \(\mu\) on \(\mathbb{R}_+ \times A\), where \(A\) is a Borelian subset of \(\mathbb{R}^q\), endowed with its Borel \(\sigma\)-field \(\mathcal{B}(A)\). We assume that the random measure \(\mu\) has the intensity measure \(\lambda(da)dt\) for some \(\sigma\)-finite measure \(\lambda\) on \((A, \mathcal{B}(A))\) satisfying

\[
\lambda(A) := \int_A \lambda(da) < \infty.
\]

We set \(\tilde{\mu}(dt, da) = \mu(dt, da) - \lambda(da)dt\), the compensated martingale measure associated to \(\mu\), and denote by \(\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}\) the completion of the natural filtration generated by \(W\) and \(\mu\).

We fix a finite time duration \(T < \infty\) and we denote by \(\mathcal{P}\) the \(\sigma\)-algebra of \(\mathbb{F}\)-predictable subsets of \(\Omega \times [0, T]\). Let us introduce some additional notations. We denote by

- \(S^2\) the set of real-valued càdlàg adapted processes \(Y = (Y_t)_{0 \leq t \leq T}\) such that \(\|Y\|_{S^2} := \left(\mathbb{E} \left[ \sup_{0 \leq t \leq T} |Y_t|^2 \right] \right)^{\frac{1}{2}} < \infty\).

- \(L^p(0, T), p \geq 1\), the set of real-valued adapted processes \((\phi_t)_{0 \leq t \leq T}\) such that \(\mathbb{E} \left[ \int_0^T |\phi_t|^p dt \right] < \infty\).

- \(L^p(W), p \geq 1\), the set of \(\mathbb{R}^d\)-valued \(\mathcal{P}\)-measurable processes \(Z = (Z_t)_{0 \leq t \leq T}\) such that \(\|Z\|_{L^p(W)} := \left(\mathbb{E} \left[ \int_0^T |Z_t|^p dt \right] \right)^{\frac{1}{p}} < \infty\).

- \(L^p(\tilde{\mu}), p \geq 1\), the set of \(\mathcal{P} \otimes \mathcal{B}(A)\)-measurable maps \(U : \Omega \times [0, T] \times E \to \mathbb{R}\) such that \(\|U\|_{L^p(\tilde{\mu})} := \left(\mathbb{E} \left[ \int_0^T \int_A |U_t(a)|^2 \lambda(da) dt \right] \right)^{\frac{1}{p}} < \infty\).
• $L^2(\lambda)$ is the set of $\mathcal{B}(A)$-measurable maps $u : E \to \mathbb{R}$ such that $|u|_{L^2(\lambda)} := \left( \int_A |u(a)|^2 \lambda(da) \right)^{\frac{1}{2}} < \infty$.

• $K^2$ the closed subset of $S^2$ consisting of nondecreasing processes $K = (K_t)_{0 \leq t \leq T}$ with $K_0 = 0$.

We are then given two objects:

1. A terminal condition $\xi$, which is an $\mathcal{F}_T$-measurable random variable.

2. A generator function $F : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \times L^2(\lambda) \to \mathbb{R}$, which is a $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(L^2(\lambda))$-measurable map.

We shall impose the following assumption on these objects:

\textbf{(H0)}

(i) The random variable $\xi$ and the generator function $F$ satisfy the square integrability condition:

$$\mathbb{E}[|\xi|^2] + \mathbb{E}\left[ \int_0^T |F(t, 0, 0, 0)|^2 dt \right] < \infty.$$

(ii) The generator function $F$ satisfies the uniform Lipschitz condition: there exists a constant $C_F$ such that

$$|F(t, y, z, u) - F(t, y', z', u')| \leq C_F(|y - y'| + |z - z'| + |u - u'|_{L^2(\lambda)}),$$

for all $t \in [0, T]$, $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^d$ and $u, u' \in L^2(\lambda)$.

(iii) The generator function $F$ satisfies the monotonicity condition:

$$F(t, y, z, u) - F(t, y, z, u') \leq \int_A \gamma(t, e, y, z, u, u')(u(a) - u'(a))\lambda(da),$$

for all $t \in [0, T]$, $z \in \mathbb{R}^d$, $y \in \mathbb{R}$ and $u, u' \in L^2(\lambda)$, where $\gamma : [0, T] \times \Omega \times E \times \mathbb{R} \times \mathbb{R}^d \times L^2(\lambda) \times L^2(\lambda) \to \mathbb{R}$ is a $\mathcal{P} \otimes \mathcal{B}(A) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(L^2(\lambda)) \otimes \mathcal{B}(L^2(\lambda))$-measurable map satisfying: $C_1 \leq \gamma(t, a, y, z, u, u') \leq C_2$, for all $a \in A$, with two constants $-1 < C_1 \leq 0 \leq C_2$.

Let us now introduce our class of Backward Stochastic Differential Equations (BSDE) with partially nonpositive jumps written in the form:

$$Y_t = \xi + \int_t^T F(s, Y_s, Z_s, U_s)ds + K_T - K_t - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a)\tilde{\mu}(ds, da), \quad 0 \leq t \leq T, \text{ a.s.}$$

with

$$U_t(a) \leq 0, \quad d\mathbb{P} \otimes dt \otimes \lambda(da) \text{ a.e. on } \Omega \times [0, T] \times E.$$
Definition 2.1.1 A minimal solution to the BSDE with terminal data/generator \((\xi, F)\) and \(A\)-nonpositive jumps is a quadruple of processes \((Y, Z, U, K)\) \(\in S^2 \times L^2(W) \times L^2(\tilde{\mu}) \times K^2\) satisfying (2.1.1)-(2.1.2) such that for any other quadruple \((\bar{Y}, \bar{Z}, \bar{U}, \bar{K})\) \(\in S^2 \times L^2(W) \times L^2(\tilde{\mu}) \times K^2\) satisfying (2.1.1)-(2.1.2), we have

\[Y_t \leq \bar{Y}_t, \quad 0 \leq t \leq T, \text{ a.s.}\]

Remark 2.1.1 Notice that when it exists, there is a unique minimal solution. Indeed, by definition, we clearly have uniqueness of the component \(Y\). The uniqueness of \(Z\) follows by identifying the Brownian parts and the finite variation parts, and then the uniqueness of \((U, K)\) is obtained by identifying the predictable parts and by recalling that the jumps of \(\mu\) are inaccessible. By misuse of language, we say sometimes that \(Y\) (instead of the quadruple \((Y, Z, U, K)\)) is the minimal solution to (2.1.1)-(2.1.2).

In order to ensure that the problem of getting a minimal solution is well-posed, we shall need to assume:

\((H1)\) There exists a quadruple \((\bar{Y}, \bar{Z}, \bar{K}, \bar{U})\) \(\in S^2 \times L^2(W) \times L^2(\tilde{\mu}) \times K^2\) satisfying (2.1.1)-(2.1.2).

We shall see later in Lemma 3.1.6 how such condition is satisfied in a Markovian framework.

2.2 Existence and approximation by penalization

In this section, we prove the existence of a minimal solution to (2.1.1)-(2.1.2), based on approximation via penalization. For each \(n \in \mathbb{N}\), we introduce the penalized BSDE with jumps

\[Y^n_t = \xi + \int_t^T F(s, Y^n_s, Z^n_s, U^n_s)ds + K^n_T - K^n_t \]

\[- \int_t^T Z^n_s dW_s - \int_t^T \int_A U^n_s(a)\tilde{\mu}(ds, da), \quad 0 \leq t \leq T,\]

where \(K^n\) is the nondecreasing process in \(K^2\) defined by

\[K^n_t = n \int_0^t \int_A [U^n_s(a)]^+ \lambda(da) ds, \quad 0 \leq t \leq T.\]

Here \([u]^+ = \max(u, 0)\) denotes the positive part of \(u\). Notice that this penalized BSDE can be rewritten as

\[Y^n_t = \xi + \int_t^T F_n(s, Y^n_s, Z^n_s, U^n_s)ds - \int_t^T Z^n_s dW_s - \int_t^T \int_A U^n_s(a)\tilde{\mu}(ds, da), \quad 0 \leq t \leq T,\]

where the generator \(F_n\) is defined by

\[F_n(t, y, z, u) = F(t, y, z, u) + n \int_A [u(a)]^+ \lambda(da),\]
for all \((t, y, z, u) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d \times L^2(\lambda)\). Under \((H0)\)(ii)-(iii) and since \(\lambda(E) < \infty\), we see that \(F_n\) is Lipschitz continuous w.r.t. \((y, z, u)\) for all \(n \in \mathbb{N}\). Therefore, we obtain from Lemma 2.4 in [40], that under \((H0)\), BSDE (2.2.3) admits a unique solution \((Y^n, Z^n, U^n) \in S^2 \times L^2(W) \times L^2(\tilde{\mu})\) for any \(n \in \mathbb{N}\).

**Lemma 2.2.1** Let Assumption \((H0)\) holds. The sequence \((Y^n)_n\) is nondecreasing, i.e. \(Y^n_t \leq Y^{n+1}_t\) for all \(t \in [0, T]\) and all \(n \in \mathbb{N}\).

**Proof.** Fix \(n \in \mathbb{N}\), and observe that

\[
F_n(t, e, y, z, u) \leq F_{n+1}(t, e, y, z, u),
\]

for all \((t, e, y, z, u) \in [0, T] \times E \times \mathbb{R} \times \mathbb{R}^d \times L^2(\lambda)\). Under Assumption \((H0)\), we can apply the comparison Theorem 2.5 in [40], which shows that \(Y^n_t \leq Y^{n+1}_t\), \(0 \leq t \leq T\), a.s. \(\square\)

The next result shows that the sequence \((Y^n)_n\) is upper-bounded by any solution to the constrained BSDE.

**Lemma 2.2.2** Let Assumption \((H0)\) holds. For any quadruple \((\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K}) \in S^2 \times L^2(W) \times L^2(\tilde{\mu}) \times K^2\) satisfying (2.1.1)-(2.1.2), we have

\[
Y^n_t \leq \tilde{Y}_t, \quad 0 \leq t \leq T, \quad n \in \mathbb{N}.
\]

**Proof.** Fix \(n \in \mathbb{N}\), and consider a quadruple \((\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K}) \in S^2 \times L^2(W) \times L^2(\tilde{\mu}) \times K^2\) solution to (2.1.1)-(2.1.2). Then, \(\tilde{U}\) clearly satisfies \(\int_0^T \int_A [\tilde{U}_s(a)]^+ \lambda(da)ds = 0\) for all \(t \in [0, T]\), and so \((\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K})\) is a supersolution to the penalized BSDE (2.2.3), i.e:

\[
\tilde{Y}_t = \xi + \int_t^T F_n(s, \tilde{Y}_s, \tilde{Z}_s, \tilde{U}_s)ds + \tilde{K}_T - \tilde{K}_t
\]

\[
- \int_t^T \tilde{Z}_sdW_s - \int_t^T \int_A \tilde{U}_s(a)\tilde{\mu}(ds, da), \quad 0 \leq t \leq T.
\]

By a slight adaptation of the comparison Theorem 2.5 in [40] under \((H0)\), we obtain the required inequality: \(Y^n_t \leq \tilde{Y}_t\), \(0 \leq t \leq T\). \(\square\)

We now establish a priori uniform estimates on the sequence \((Y^n, Z^n, U^n, K^n)_n\).

**Lemma 2.2.3** Under \((H0)\) and \((H1)\), there exists some constant \(C\) depending only on \(T\) and the monotonicity condition of \(F\) in \((H0)\)(iii) such that

\[
\|Y^n\|_{S^2}^2 + \|Z^n\|_{L^2(W)}^2 + \|U^n\|_{L^2(\tilde{\mu})}^2 + \|K^n\|_{S^2}^2 \\
\leq C \left( \mathbb{E}[|\xi|^2] + \mathbb{E}[\int_0^T |F(t, 0, 0, 0)|^2 dt] + \mathbb{E}[\sup_{0 \leq t \leq T} |\tilde{Y}_t|^2] \right), \quad \forall n \in \mathbb{N}.
\]

**Proof.** In what follows we shall denote by \(C > 0\) a generic positive constant depending only on \(T\), and the linear growth condition of \(F\) in \((H0)\)(ii), which may vary from line to
line. By applying Itô’s formula to $|Y^n_t|^2$, and observing that $K^n$ is continuous and $\Delta Y^n_t = \int_A U^n_t(a) \mu(\{t\}, da)$, we have

$$E|\xi|^2 = E|Y^n_t|^2 - 2E\int_t^T Y^n_s F(s, Y^n_s, Z^n_s, U^n_s)ds - 2E\int_t^T Y^n_s dK^n_s + E\int_t^T |Z^n_s|^2 ds$$

$$+ E\int_t^T \int_A \{ |Y^n_s + U^n_s(a)|^2 - |Y^n_s|^2 - 2Y^n_s U^n_s(a) \} \mu(da, ds)$$

$$= E|Y^n_t|^2 + E\int_t^T |Z^n_s|^2 ds + E\int_t^T \int_A |U^n_s(a)|^2 \lambda(da) ds$$

$$- 2E\int_t^T Y^n_s F(s, Y^n_s, Z^n_s, U^n_s)ds - 2E\int_t^T Y^n_s dK^n_s, \quad 0 \leq t \leq T.$$  

From (H0)(iii), the inequality $Y^n_t \leq \tilde{Y}_t$ by Lemma 2.2.2 under (H1), and the inequality $2ab \leq \frac{a^2}{T} + \frac{b^2}{T}$ for any constant $a > 0$, we have:

$$E|Y^n_t|^2 + E\int_t^T |Z^n_s|^2 ds + E\int_t^T \int_A |U^n_s(a)|^2 \lambda(da) ds$$

$$\leq E|\xi|^2 + CE\int_t^T |Y^n_s| \left( |F(s, 0, 0, 0)| + |Y^n_s| + |Z^n_s| + |U^n_s|_{L^2(\lambda)} \right) ds$$

$$+ \frac{1}{\alpha} E \left[ \sup_{s \in [0,T]} |Y^n_s|^2 \right] + aE|K^n_T - K^n_t|^2.$$  

Using again the inequality $ab \leq \frac{a^2}{T} + \frac{b^2}{T}$, and (H0)(i), we get

$$E|Y^n_t|^2 + \frac{1}{2} E\int_t^T |Z^n_s|^2 ds + \frac{1}{2} E\int_t^T \int_A |U^n_s(a)|^2 \lambda(da) ds$$

$$\leq C E\int_t^T |Y^n_s|^2 ds + E|\xi|^2 + \frac{1}{2} E\int_0^T |F(s, 0, 0, 0)|^2 ds + \frac{1}{\alpha} E \left[ \sup_{t \in [0,T]} |Y^n_t|^2 \right] + aE|K^n_T - K^n_t|^2.$$  

Now, from the relation (2.2.3), we have:

$$K^n_T - K^n_t = Y^n_t - \xi - \int_t^T F(s, Y^n_s, Z^n_s, U^n_s)ds$$

$$+ \int_t^T Z^n_s dW_s + \int_t^T \int_A U^n_s(a) \tilde{\mu}(ds, da).$$

Thus, there exists some positive constant $C_1$ depending only on the linear growth condition of $F$ in (H0)(ii) such that

$$E|K^n_T - K^n_t|^2 \leq C_1 \left( E|\xi|^2 + E\int_0^T |F(s, 0, 0, 0)|^2 ds + E|Y^n_t|^2 \right.$$

$$+ E\int_t^T (|Y^n_s|^2 + |Z^n_s|^2 + |U^n_s|_{L^2(\lambda)}^2) ds, \quad 0 \leq t \leq T.$$  

Hence, by choosing $\alpha > 0$ s.t. $C_1 \alpha \leq \frac{1}{4}$, and plugging into (2.2.6), we get

$$\frac{3}{4} E|Y^n_t|^2 + \frac{1}{4} E\int_t^T |Z^n_s|^2 ds + \frac{1}{4} E\int_t^T \int_A |U^n_s(a)|^2 \lambda(da) ds$$

$$\leq C E\int_t^T |Y^n_s|^2 ds + \frac{5}{4} E|\xi|^2 + \frac{1}{4} E\int_0^T |F(s, 0, 0, 0)|^2 ds + \frac{1}{\alpha} E \left[ \sup_{s \in [0,T]} |\tilde{Y}_s|^2 \right], \quad 0 \leq t \leq T.$$
Thus application of Gronwall’s lemma to \( t \mapsto \mathbb{E}|Y^n_t|^2 \) yields:

\[
\sup_{0 \leq t \leq T} \mathbb{E}|Y^n_t|^2 + \mathbb{E} \int_0^T |Z^n_t|^2 dt + \mathbb{E} \int_0^T \int_A |U^n_t(a)|^2 \lambda(da) dt \\
\leq C \left( \mathbb{E}|\xi|^2 + \mathbb{E} \int_0^T |F(t,0,0,0)|^2 dt + \mathbb{E} \left[ \sup_{t \in [0,T]} |Y_t|^2 \right] \right), \tag{2.2.5}
\]

which gives the required uniform estimates \( (2.2.5) \) for \((Z^n, U^n)_n\) and also \((K^n)_n\) by \( (2.2.7) \).

Finally, by writing from \( (2.2.3) \) that

\[
\sup_{0 \leq t \leq T} |Y^n_t| \leq |\xi| + \int_0^T |F(t, Y^n_t, Z^n_t, U^n_t)| dt + K^n_T \\
+ \sup_{0 \leq t \leq T} \left| \int_0^t Z^n_s dW_s \right| + \sup_{0 \leq t \leq T} \left| \int_0^t \int_A U^n_s(a) \tilde{\mu}(ds, da) \right|
\]

we obtain the required uniform estimate \( (2.2.5) \) for \((Y^n)_n\) by Burkholder-Davis-Gundy inequality, linear growth condition in \((H0)\)\( (ii) \), and the uniform estimates for \((Z^n, U^n, K^n)_n\).

We can now state the main result of this paragraph.

**Theorem 2.2.1** Under \((H0)\) and \((H1)\), there exists a unique minimal solution \((Y, Z, U, K) \in S^2 \times L^2(W) \times L^2(\hat{\mu}) \times K^2\) with \(K\) predictable, to \( (2.1.1) \)\(-\)(2.1.2). \(Y\) is the increasing limit of \((Y^n)_n\) and also in \(L^2(0, T)\), \(K_t\) is the weak limit of \((K^n)_n\) in \(L^2(\Omega, F_t, \mathbb{P})\) for all \(t \in [0, T]\), and for any \(p \in [1, 2)\),

\[
\|Z^n - Z\|_{L^p(W)} + \|U^n - U\|_{L^p(\hat{\mu})} \rightarrow 0,
\]

as \(n\) goes to infinity.

**Proof.** By the Lemmata \(2.2.1\) and \(2.2.2\), \((Y^n)_n\) converges increasingly to some adapted process \(Y\), satisfying: \(\|Y\|_{s^2} \leq \infty\) by the uniform estimate for \((Y^n)_n\) in Lemma \(2.2.3\) and Fatou’s lemma. Moreover by dominated convergence theorem, the convergence of \((Y^n)_n\) to \(Y\) also holds in \(L^2(0, T)\). Next, by the uniform estimates for \((Z^n, U^n, K^n)_n\) in Lemma \(2.2.3\), we can apply the monotonic convergence Theorem 3.1 in \(10\), which extends to the jump case the monotonic convergence theorem of Peng \(35\) for BSDE. This provides the existence of \((Z, U) \in L^2(W) \times L^2(\hat{\mu})\), and \(K\) predictable, nondecreasing with \(\mathbb{E}[K^2_T] < \infty\), such that the sequence \((Z^n, U^n, K^n)_n\) converges in the sense of Theorem \(2.2.1\) to \((Z, U, K)\) satisfying:

\[
Y_t = \xi + \int_t^T F(s, Y_s, Z_s, U_s) ds + K_T - K_t \\
- \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \tilde{\mu}(ds, da), \quad 0 \leq t \leq T.
\]

Thus, the process \(Y\) is the difference of a càdlàg process and the nondecreasing process \(K\), and by Lemma 2.2 in \(35\), this implies that \(Y\) and \(K\) are also càdlàg, hence respectively
in $S^2$ and $K^2$. Moreover, from the strong convergence in $L^1(\tilde{\mu})$ of $(U^n)_n$ to $U$ and since $\lambda(E) < \infty$, we have

$$
\mathbb{E} \int_0^T \int_A [U^n_s(a)]^+ \lambda(da) ds \longrightarrow \mathbb{E} \int_0^T \int_A [U_s(a)]^+ \lambda(da) ds,
$$

as $n$ goes to infinity. Since $K^n_T = n \int_0^T \int_A [U^n_s(a)]^+ \lambda(da) ds$ is bounded in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, this implies

$$
\mathbb{E} \int_0^T \int_A [U_s(a)]^+ \lambda(da) ds = 0,
$$

which means that the $A$-nonpositive constraint (2.1.2) is satisfied. Hence, $(Y, Z, K, U)$ is a solution to the constrained BSDE (2.1.1)-(2.1.2), and by Lemma 2.2.2, $Y = \lim Y^n$ is the minimal solution. Finally, the uniqueness of the solution $(Y, Z, U, K)$ is given by Remark 2.1.1.

### 2.3 Dual representation

In this section, we consider the case where the generator function $F(t, \omega)$ does not depend on $y, z, u$. Our main goal is to provide a dual representation of the minimal solution to the BSDE with nonpositive jumps in terms of a family of equivalent probability measures.

Let $\mathcal{V}$ be the set of essentially bounded $\mathbb{P} \otimes \mathcal{B}(A)$-measurable processes valued in $(0, \infty)$, and consider for any $\nu \in \mathcal{V}$, the Doléans-Dade exponential local martingale

$$
L_t^\nu := \mathcal{E} \left( \int_0^t \int_A (\nu_s(a) - 1) \tilde{\mu}(ds, da) \right)_t
$$

$$
= \exp \left( \int_0^t \int_A \ln \nu_s(a) \mu(ds, da) - \int_0^t \int_A (\nu_s(a) - 1) \lambda(da) ds \right), \quad 0 \leq t \leq T(2.3.9)
$$

When $L^\nu$ is a true martingale, i.e. $\mathbb{E}[L_T^\nu] = 1$, it defines a probability measure $\mathbb{P}^\nu$ equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{F}_T)$ with Radon-Nikodym density:

$$
\frac{d\mathbb{P}^\nu}{d\mathbb{P}}|_{\mathcal{F}_t} = L_t^\nu, \quad 0 \leq t \leq T, \quad (2.3.10)
$$

and we denote by $\mathbb{E}^\nu$ the expectation operator under $\mathbb{P}^\nu$. Notice that $W$ remains a Brownian motion under $\mathbb{P}^\nu$, and the effect of the probability measure $\mathbb{P}^\nu$, by Girsanov’s Theorem, is to change the compensator $\lambda(da) dt$ of $\mu$ under $\mathbb{P}$ to $\nu_t(a) \lambda(da) dt$ under $\mathbb{P}^\nu$. We denote by $\tilde{\mu}^\nu(dt, da) = \mu(dt, da) - \nu_t(a) \lambda(da) dt$ the compensated martingale measure of $\mu$ under $\mathbb{P}^\nu$.

We then introduce the subset $\mathcal{V}^n$ of $\mathcal{V}$ as the elements $\nu \in \mathcal{V}$ essentially bounded by $n + 1$, for $n \in \mathbb{N}$.

**Lemma 2.3.4** For any $\nu \in \mathcal{V}$, $L^\nu$ is a uniformly integrable martingale, and $L_T^\nu$ is square integrable.
\textbf{Proof.} Several sufficient criteria for $L^{\nu}$ to be a uniformly integrable martingale are known. We refer for example to the recent paper [39], which shows that if

$$S_T^{\nu} := \exp \left( \int_0^T \int_A |\nu_t(a) - 1|^2 \lambda(da)dt \right)$$

is integrable, then $L^{\nu}$ is uniformly integrable. By definition of $\mathcal{V}$, we see that for $\nu \in \mathcal{V}$, $S_T^{\nu}$ is essentially bounded since $\nu$ is essentially bounded and $\lambda(E) < \infty$. Moreover, from the explicit form (2.3.9) of $L^{\nu}$, we have $|L_T^{\nu}|^2 = L_T^2 S_T^{\nu}$, and so $E|L_T^{\nu}|^2 \leq \|S_T^{\nu}\|_{\infty}$. \hfill \Box

We can then associate to each $\nu \in \mathcal{V}$ the probability measure $\mathbb{P}^{\nu}$ through (2.3.10). We first provide a dual representation of the penalized BSDEs in terms of such $\mathbb{P}^{\nu}$-martingales. To this end, we need the following Lemma.

\textbf{Lemma 2.3.5} Let $\phi \in L^2(\mathcal{W})$ and $\psi \in L^2(\tilde{\mu})$. Then for every $\nu \in \mathcal{V}$, the processes $\int_0^t \phi dW_t$ and $\int_0^t \int_A \psi_t(a) \tilde{\mu}^{\nu}(dt, da)$ are $\mathbb{P}^{\nu}$-martingales.

\textbf{Proof.} Fix $\phi \in L^2(\mathcal{W})$ and $\nu \in \mathcal{V}$ and denote by $M^{\phi}$ the process $\int_0^t \phi dW_t$. Since $W$ remains a $\mathbb{P}^{\nu}$-Brownian motion, we know that $M^{\phi}$ is a $\mathbb{P}^{\nu}$-local martingale. From Burkholder-Davis-Gundy and Cauchy Schwarz inequalities, we have

$$E^\nu \left[ \sup_{t \in [0,T]} |M_t^{\phi}| \right] \leq C E^\nu \left[ \sqrt{\langle M^\phi \rangle_T} \right] = C E^\nu \left[ L_T^{\nu} \sqrt{\int_0^T |\phi_t|^2 dt} \right] \leq C \sqrt{E^\nu \left[ |L_T^{\nu}|^2 \right]} \sqrt{E^\nu \left[ \int_0^T |\phi_t|^2 dt \right]} < \infty,$$

since $L_T^{\nu}$ is square integrable by Lemma 2.3.4 and $\phi \in L^2(\mathcal{W})$. This implies that $M^{\phi}$ is a $\mathbb{P}^{\nu}$-uniformly integrable, and hence a true $\mathbb{P}^{\nu}$-martingale. The proof for $\int_0^t \int_A \psi_t(a) \tilde{\mu}^{\nu}(dt, da)$ follows exactly the same lines and is therefore omitted. \hfill \Box

\textbf{Proposition 2.3.1} For all $n \in \mathbb{N}$, the solution to the penalized BSDE (2.2.3) is explicitly represented as

$$Y^n_t = \text{ess sup}_{\nu \in \mathcal{V}^n} E^\nu \left[ \xi + \int_t^T F(s)ds \bigg| \mathcal{F}_t \right], \quad 0 \leq t \leq T. \quad (2.3.11)$$

\textbf{Proof.} Fix $n \in \mathbb{N}$. For any $\nu \in \mathcal{V}^n$, and by introducing the compensated martingale measure $\tilde{\mu}^{\nu}(dt, da) = \tilde{\mu}(dt, da) - (\nu_t(a) - 1)\lambda(da)dt$ under $\mathbb{P}^{\nu}$, we see that the solution $(Y^n, \mathcal{Z}^n, U^n)$ to the BSDE (2.2.3) satisfies:

$$Y^n_t = \xi + \int_t^T \left[ F(s) + \int_A \left( n[U^n_s(a)]^+ - (\nu_s(a) - 1)U^n_s(a) \right) \lambda(da) \right] ds - \int_t^T \mathcal{Z}^n_s dW_s - \int_t^T \int_A U^n_s(a) \tilde{\mu}^{\nu}(ds, da). \quad (2.3.12)$$

By taking expectation in (2.3.12) under $\mathbb{P}^{\nu}$ ($\sim \mathbb{P}$), we then get from Lemma 2.3.5

$$Y^n_t = E^\nu \left[ \xi + \int_t^T \left( F(s) + \int_A \left( n[U^n_s(a)]^+ - (\nu_s(a) - 1)U^n_s(a) \right) \lambda(da) \right] ds \bigg| \mathcal{F}_t \right]. \quad (2.3.13)$$
Now, observe that for any $\nu \in V^n$, hence valued in $[1, n + 1]$, we have
\[ n[U^n_t(a)]^+ - (\nu_t(a) - 1)U^n_t(a) \geq 0, \quad d\mathbb{P} \otimes dt \otimes \lambda(da) \text{ a.e.} \]
which yields by (2.3.13):
\[ Y^n_t = \text{ess sup}_{\nu \in V^n} \mathbb{E}^\nu \left[ \xi + \int_t^T F(s)ds \bigg| \mathcal{F}_t \right]. \quad (2.3.14) \]
On the other hand, let us consider the process $\nu^* \in V^n$ defined by
\[ \nu^*_t(a) = \mathbbm{1}_{U_t(a) \leq 0} + (n + 1)\mathbbm{1}_{U_t(a) > 0}, \quad 0 \leq t \leq T, e \in E. \]
By construction, we clearly have
\[ n[U^n_t(a)]^+ - (\nu^*_t(a) - 1)U^n_t(a) = 0, \quad \text{for all} \ 0 \leq t \leq T, e \in E, \]
and thus for this choice of $\nu = \nu^*$ in (2.3.13):
\[ Y^n_t = \mathbb{E}^{\nu^*} \left[ \xi + \int_t^T F(s)ds \bigg| \mathcal{F}_t \right]. \]
Together with (2.3.14), this proves the required representation of $Y^n$.

**Remark 2.3.2** Arguments in the proof of Proposition 2.3.1 shows that the relation (2.3.11) holds for general generator function $F$ depending on $(y, z, u)$, i.e.
\[ Y^n_t = \text{ess sup}_{\nu \in V^n} \mathbb{E}^\nu \left[ \xi + \int_t^T F(s,Y^n_s,Z^n_s,U^n_s)ds \bigg| \mathcal{F}_t \right], \]
which is in this case an implicit relation for $Y^n$. Moreover, the essential supremum in this dual representation is attained for some $\nu^*$, which takes extreme values 1 or $n + 1$ depending on the sign of $U^n$, i.e. of bang-bang form.

Let us then focus on the limiting behavior of the above dual representation for $Y^n$ when $n$ goes to infinity.

**Theorem 2.3.2** Under $(H1)$, the minimal solution to (2.1.1)-(2.1.2) is explicitly represented as
\[ Y_t = \text{ess sup}_{\nu \in V} \mathbb{E}^\nu \left[ \xi + \int_t^T F(s)ds \bigg| \mathcal{F}_t \right], \quad 0 \leq t \leq T. \quad (2.3.15) \]

**Proof.** Let $(Y, Z, U, K) \in \mathcal{V}$ be the minimal solution to (2.1.1)-(2.1.2). Let us denote by $\hat{Y}$ the process defined in the r.h.s of (2.3.15). Since $V^n \subset V$, it is clear from the representation (2.3.11) that $Y^n_t \leq \hat{Y}_t$, for all $n$. Recalling from Theorem 2.2.1 that $Y$ is the pointwise limit of $Y^n$, we deduce that $Y_t = \lim_{n \to \infty} Y^n_t \leq \hat{Y}_t, \ 0 \leq t \leq T$. 

Conversely, for any $\nu \in \mathcal{V}$, let us consider the compensated martingale measure $\tilde{\mu}^\nu(dt, da) = \tilde{\mu}(dt, da) - (\nu_t(a) - 1)\lambda(da)dt$ under $\mathbb{P}^\nu$, and observe that $(Y, Z, U, K)$ satisfies:

$$Y_t = \xi + \int_t^T \left[ F(s) - \int_A (\nu_s(a) - 1)U_s(a)\lambda(da) \right] ds + K_T - K_t \quad (2.3.16)$$

$$- \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a)\tilde{\mu}^\nu(ds, da).$$

Thus, by taking expectation in (2.3.16) under $\mathbb{P}^\nu$ from Lemma 2.3.5, and recalling that $K$ is nondecreasing, we have:

$$Y_t \geq \mathbb{E}^\nu \left[ \xi + \int_t^T \left( F(s) - \int_A (\nu_s(a) - 1)U_s(a)\lambda(da) \right) ds \bigg| \mathcal{F}_t \right]$$

$$\geq \mathbb{E}^\nu \left[ \xi + \int_t^T F(s) ds \bigg| \mathcal{F}_t \right],$$

since $\nu$ is valued in $[1, \infty)$, and $U$ satisfies the nonpositive constraint (2.1.2). Since $\nu$ is arbitrary in $\mathcal{V}$, this proves the inequality $Y_t \geq \tilde{Y}_t$, and finally the required relation $Y = \tilde{Y}$.

$\square$
Chapter 3

Nonlinear IPDE and Feynman-Kac formula

In this chapter, we shall show how minimal solutions to our BSDE class with partially nonpositive jumps provides actually a new probabilistic representation (or Feynman-Kac formula) to fully nonlinear integro-partial differential equation (IPDE) of Hamilton-Jacobi-Bellman (HJB) type, when dealing with a suitable Markovian framework.

3.1 The Markovian framework

We first assume that

\((HA)\) \( A\) is compact, its interior \( \bar{A}\) is connex, and \( A = \text{Adh}(\bar{A})\), the closure of its interior.

\[
\int_{\bar{A}} \lambda(da) < \infty.
\]

We also assume that

\((H\lambda)\)

(i) The measure \( \lambda \) supports the whole set \( \bar{A} \): for any \( a \in \bar{A} \) and any open neighborhood \( O \) of \( a \) in \( \mathbb{R}^q \) we have \( \lambda(O \cap \bar{A}) > 0 \).

(ii) The boundary of \( A \): \( \partial A = A \setminus \bar{A} \), is negligible w.r.t. \( \lambda \), i.e. \( \lambda(\partial A) = 0 \).

Given some measurable functions \( b : \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^d \), \( \sigma : \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^{d \times d} \) and \( \beta : \mathbb{R}^d \times \mathbb{R}^q \times L \to \mathbb{R}^d \), we introduce the forward Markov regime-switching jump-diffusion process \((X, I)\) governed by:

\[
\begin{align*}
    dX_s &= b(X_s, I_s)ds + \sigma(X_s, I_s)dW_s, \\
    dI_s &= \int_{\bar{A}} (a - I_s^-)\mu(ds, da).
\end{align*}
\]
In other words, $I$ is the pure jump process valued in $A$ associated to the Poisson random measure $\pi$, which changes the coefficients of jump–diffusion process $X$. We make the usual assumptions on the forward jump–diffusion coefficients:

(HFC) There exists a constant $C$ such that

$$|b(x, a) - b(x', a')| + |\sigma(x, a) - \sigma(x', a')| \leq C(|x - x'| + |a - a'|),$$

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

Remark 3.1.3 We do not make any ellipticity assumption on $\sigma$. In particular, some lines and columns of $\sigma$ may be equal to zero, and so there is no loss of generality by considering that the dimension of $X$ and $W$ are equal. We can also make the coefficients $b, \sigma$ and $\beta$ depend on time with the following standard procedure: we introduce the time variable as a state component $\Theta_t = t$, and consider the forward Markov system:

$$\begin{align*}
    dX_s &= b(X_s, \Theta_s, I_s)ds + \sigma(X_s, \Theta_s, I_s)dW_s, \\
    d\Theta_s &= ds \\
    dI_s &= \int_A (a - I_s^-) \pi(ds, da).
\end{align*}$$

which is of the form given above, but with an enlarged state $(X, \Theta, I)$ (with degenerate noise), and with the resulting assumptions on $b(x, \theta, a)$ and $\sigma(x, \theta, a)$.

Under these conditions, existence and uniqueness of a solution $(X^{t,x,a}_s, I^{t,a}_s)_{t \leq s \leq T}$ to (3.1.1)-(3.1.2) starting from $(x, a) \in \mathbb{R}^d \times \mathbb{R}^q$ at time $s = t \in [0, T]$, is well-known, and we have the standard estimate: for all $p \geq 2$, there exists some positive constant $C_p$ s.t.

$$\mathbb{E}\left[ \sup_{t \leq s \leq T} |X^{t,x,a}_s|^p + |I^{t,a}_s|^p \right] \leq C_p(1 + |x|^p + |a|^p), \quad (3.1.3)$$

for all $(t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q$.

In this Markovian framework, the terminal data and generator of our class of BSDE are given by two continuous functions $g: \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}$ and $f: \mathbb{R}^d \times \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$. We make the following assumptions on the BSDE coefficients:

(HBC) There exists some constant $C$ s.t.

$$|f(x, a, y, z) - f(x', a', y', z')| \leq C(|x - x'| + |a - a'| + |y - y'| + |z - z'|),$$

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

In this framework, the BSDE problem (2.1.1)-(2.1.2) takes the following form: find the minimal solution $(Y, Z, U, K) \in S^2 \times L^2(W) \times L^2(\mu) \times K^2$ to

$$\begin{align*}
    Y_t &= g(X_T, I_T) + \int_t^T f(X_s, I_s, Y_s, Z_s)ds + K_T - K_t \\
    &\quad - \int_t^T Z_s dW_s - \int_t^T \int_L U_s(a)\tilde{\mu}(ds, da), \quad (3.1.4)
\end{align*}$$
with
\[ U_t(a) \leq 0, \quad d\mathbb{P} \otimes dt \otimes \lambda(da) \text{ a.e.} \quad (3.1.5) \]

The main goal of this chapter is to relate the BSDE \((3.1.4)\) with nonpositive jumps \((3.1.5)\) to the following nonlinear IPDE of HJB type:
\[
- \frac{\partial w}{\partial t} - \sup_{a \in A} \left[ \mathcal{L}^a w + f(\cdot, a, w, \sigma^a(\cdot, a) D_x w) \right] = 0, \quad \text{on } [0, T) \times \mathbb{R}^d, \quad (3.1.6)
\]
\[
w(T, x) = \sup_{a \in A} g(x, a), \quad x \in \mathbb{R}^d, \quad (3.1.7)
\]
where
\[
\mathcal{L}^a w(t, x) = b(x, a).D_x w(t, x) + \frac{1}{2} \text{tr}(\sigma(\sigma^a(x, a) D^2_x w(t, x))
\]
for \((t, x, a) \in [0, T) \times \mathbb{R}^d \times \mathbb{R}^q\).

Notice that under (HBC) and \((3.1.3)\) (which follows from (HFC)), the generator \(F(t, \omega, y, z, u) = f(X_t(\omega), I_t(\omega), y, z)\) and the terminal condition \(\xi = g(X_T, I_T)\) satisfy clearly Assumption (H0). Let us now show that Assumption (H1) is satisfied. More precisely, we have the following result.

Lemma 3.1.6 Let Assumptions (HFC), (HBC1) hold. Then, for any initial condition \((t, x, a) \in [0, T) \times \mathbb{R}^d \times \mathbb{R}^q\), there exists a solution \(\{\bar{Y}^s_{t,x,a}, \bar{Z}^s_{t,x,a}, \bar{U}^s_{t,x,a}, \bar{K}^s_{t,x,a}, t \leq s \leq T\}\) to the BSDE \((3.1.4) - (3.1.5)\) when \((X, I) = \{(X^s_{t,x,a}, I^s_{t,x,a}), t \leq s \leq T\}\), with \(\bar{Y}^s_{t,x,a} = \bar{v}(s, X^s_{t,x,a})\) for some deterministic function \(\bar{v}\) on \([0, T] \times \mathbb{R}^d\) satisfying a polynomial growth condition:
\[
\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \frac{\bar{v}(t,x)}{1 + |x|^2} < \infty. \quad (3.1.8)
\]

Proof. Under (HBC1) and since \(A\) is compact, we observe that
\[
C_{f,g} := \sup_{x \in \mathbb{R}^d, a \in A} \frac{|g(x, a)| + |f(x, a, y, z)|}{1 + |x| + |y| + |z|} < \infty. \quad (3.1.9)
\]
Let us then consider the smooth function \(\bar{v}(t, x) = \bar{C}e^{\rho(T-t)}(1 + |x|^2)\) for some positive constants \(\bar{C}\) and \(\rho\) to be determined later. We claim that for \(\bar{C}\) and \(\rho\) large enough, the function \(\bar{v}\) is a classical supersolution to \((3.1.6) - (3.1.7)\). Indeed, observe first that from the growth condition on \(g\) in \((3.1.9)\), there exists \(\bar{C} > 0\) s.t. \(\hat{g}(x) := \sup_{a \in A} g(x, a) \leq \bar{C}(1 + |x|^2)\) for all \(x \in \mathbb{R}^d\). For such \(\bar{C}\), we then have: \(\bar{v}(T, \cdot) \geq \hat{g}\). On the other hand, we see after straightforward calculation that there exists a positive constant \(C\) depending only on \(\bar{C}, C_{f,g}\), and the linear growth condition in \(x\) on \(b\) and \(\sigma\) by (HFC) (recall that \(A\) is compact), such that
\[
- \frac{\partial \bar{v}}{\partial t} - \sup_{a \in A} \left[ \mathcal{L}^a \bar{v} + f(\cdot, a, \bar{v}, \sigma^a(\cdot, a) D_x \bar{v}) \right] \geq (\rho - C)\bar{v} \geq 0,
\]
by choosing $\rho \geq C$. Let us now define the quadruple $(\bar{Y}, \bar{Z}, \bar{U}, \bar{K})$ by:

$$
\begin{align*}
\bar{Y}_t &= \bar{v}(t, X_t) \text{ for } t < T, \quad \bar{Y}_T = g(X_T, I_T), \\
\bar{Z}_t &= \sigma'(X_{t^-}, I_{t^-})D_x \bar{v}(t, X_{t^-}), \quad t \leq T, \\
\bar{U}_t &= 0, \quad t \leq T \\
\bar{K}_t &= \int_0^t \left[ -\frac{\partial \bar{v}}{\partial t}(s, X_s) - \mathcal{L}^s \bar{v}(s, X_s) - f(X_s, I_s, \bar{Z}_s) \right] ds, \quad t < T \\
\bar{K}_T &= \bar{K}_{T^-} + \bar{v}(T, X_T) - g(X_T, I_T).
\end{align*}
$$

From the supersolution property of $\bar{v}$ to (3.1.6)-(3.1.7), the process $\bar{K}$ is nondecreasing. Moreover, from the polynomial growth condition on $\bar{v}$, linear growth condition on $b$, $\sigma$, growth condition (3.1.9) on $f$, $g$ and the estimate (3.1.3), we see that $(\bar{Y}, \bar{Z}, \bar{U}, \bar{K})$ lies in $S^2 \times L^2(W) \times L^2([\bar{\mu}]) \times K^2$. Finally, by applying Itô’s formula to $\bar{v}(t, X_t)$, we conclude that $(\bar{Y}, \bar{Z}, \bar{U}, \bar{K})$ is a solution to (3.1.4), and the constraint (3.1.5) is trivially satisfied.

Under (HFC) and (HBC), we then get from Theorem 2.2.1 the existence of a unique minimal solution $\{Y_{s,t}^{l,x,a}, Z_{s,t}^{l,x,a}, U_{s,t}^{l,x,a}, K_{s,t}^{l,x,a}\}, t \leq s \leq T$ to (3.1.4)-(3.1.5) when $(X, I) = \{(X_{s,t}^{l,x,a}, I_{s,t}^{l,x,a})\}, t \leq s \leq T$. Moreover, as we shall see in the next paragraph, this minimal solution is written in this Markovian context as: $Y_{s,t}^{l,x,a} = v(s, X_{s,t}^{l,x,a}, I_{s,t}^{l,x,a})$ where $v$ is the deterministic function defined on $[0, T] \times \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}$ by:

$$
v(t, x, a) := Y_{t}^{l,x,a}, \quad (t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q. \quad (3.1.10)
$$

We aim at proving that the function $v$ defined by (3.1.10) does not depend actually on its argument $a$, and is a solution in a sense to be precised to the parabolic IPDE (3.1.6)-(3.1.7). Notice that we do not have a priori any smoothness or even continuity properties on $v$.

To this end, we first recall the definition of (discontinuous) viscosity solutions to (3.1.6)-(3.1.7). For a locally bounded function $w$ on $[0, T] \times \mathbb{R}^d$, we define its lower semicontinuous (lsc for short) envelope $w_*$ and upper semicontinuous (usc for short) envelope $w^*$ by:

$$
w_*(t, x) = \liminf_{(t', x') \to (t, x)} w(t', x') \quad \text{and} \quad w^*(t, x) = \limsup_{(t', x') \to (t, x)} w(t', x'),
$$

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

**Definition 3.1.2 (Viscosity solutions to (3.1.6)-(3.1.7))**

(i) A function $w$, lsc (resp. usc) on $[0, T] \times \mathbb{R}^d$, is called a viscosity supersolution (resp. subsolution) to (3.1.6)-(3.1.7) if

$$
\begin{align*}
\text{for any } x \in \mathbb{R}^d, \text{ and} \\
\left( -\frac{\partial \varphi}{\partial t} + \sup_{a \in A} \left[ \mathcal{L}^a \varphi + f(\cdot, a, w, \sigma(\cdot, a) D_x \varphi) \right] \right)(t, x) &\geq (\text{resp.} \leq) 0,
\end{align*}
$$

where $\mathcal{L}^a \varphi$ is the generator of the process $\bar{X}^a$.

(ii) In the case $w$ has no growth condition on $b$, $\sigma$, the notion of viscosity solution is understood in the weak sense (see [5]).
for any \((t, x) \in [0, T) \times \mathbb{R}^d\) and any \(\varphi \in C^{1,2}([0, T] \times \mathbb{R}^d)\) such that
\[
(w - \varphi)(t, x) = \min_{[0, T] \times \mathbb{R}^d} (w - \varphi) \quad \text{(resp.} \max_{[0, T] \times \mathbb{R}^d} (w - \varphi))
\].

(ii) A locally bounded function \(w\) on \([0, T) \times \mathbb{R}^d\) is called a viscosity solution to (3.1.6)-(3.1.7) if \(w_*\) is a viscosity supersolution and \(w^*\) is a viscosity subsolution to (3.1.6)-(3.1.7).

We can now state the main result of this chapter.

**Theorem 3.1.3** Assume that conditions (HA), (Ha), (HFC) and (HBC) hold. The function \(v\) in (3.1.10) does not depend on the variable \(a\) on \([0, T) \times \mathbb{R} \times \hat{A}\) i.e.
\[
v(t, x, a) = v(t, x, a'), \quad \forall \ a, a' \in \hat{A},
\]
for all \((t, x) \in [0, T) \times \mathbb{R}^d\). Let us then define by misuse of notation the function \(v\) on \([0, T) \times \mathbb{R}^d\) by:
\[
v(t, x) = v(t, x, a), \quad (t, x) \in [0, T) \times \mathbb{R}^d,
\]
(3.1.11)
for any \(a \in \hat{A}\). Then \(v\) is a viscosity solution to (3.1.6)-(3.1.7).

**Remark 3.1.4** 1. Once we have a uniqueness result for the fully nonlinear IPDE (3.1.6)-(3.1.7), Theorem 3.1.3 provides a Feynman-Kac representation of this unique solution by means of the minimal solution to the BSDE (3.1.4)-(3.1.5). This suggests consequently an original probabilistic numerical approximation of the nonlinear IPDE (3.1.6)-(3.1.7) by discretization and simulation of the minimal solution to the BSDE (3.1.4)-(3.1.5). This issue, especially the treatment of the nonpositive jump constraint, has been recently investigated in [23] and [24], where the authors analyze the convergence rate of the approximation scheme, and illustrate their results with several numerical tests arising for instance in the super-replication of options in uncertain volatilities and correlations models. We mention here that a nice feature of our scheme is the fact that the forward process \((X, I)\) can be easily simulated: indeed, notice that the jump times of \(I\) follow a Poisson distribution of parameter \(\lambda := \int_A \lambda(da)\), and so the pure jump process \(I\) is perfectly simulatable once we know how to simulate the distribution \(\lambda(da)/\bar{\lambda}\) of the jump marks. Then, we can use a standard Euler scheme for simulating the component \(X\). Our scheme does not suffer the curse of dimensionality encountered in finite difference methods or controlled Markov chains, and takes advantage of the high dimensional properties of Monte-Carlo methods.
2. We do not address here comparison principles (and so uniqueness results) for the general parabolic nonlinear IPDE (3.1.6)-(3.1.7). In the case where the generator function \(f(x, a)\) does not depend on \((y, z, u)\) (see Remark 3.1.5 below), comparison principle is proved in [37], and the result can be extended by same arguments when \(f(x, a, y, z)\) depends also on \(y, z\) under the Lipschitz condition of (HBC).

**Remark 3.1.5** Stochastic control problem
1. Let us now consider the particular and important case where the generator \( f(x, a) \) does not depend on \( (y, z) \). We then observe that the nonlinear IPDE (3.1.6) is the Hamilton-Jacobi-Bellman (HJB) equation associated to the following stochastic control problem: let us introduce the controlled jump-diffusion process:

\[
dX_s^\alpha = b(X_s^\alpha, \alpha_s)ds + \sigma(X_s^\alpha, \alpha_s)dW_s,
\]

where \( W \) is a Brownian motion independent of a random measure \( \vartheta \) on a filtered probability space \( (\Omega, \mathcal{F}, \mathbb{F}^0, \mathbb{P}) \), the control \( \alpha \) lies in \( \mathcal{A} \), the set of \( \mathbb{F}^0 \)-predictable process valued in \( A \), and define the value function for the control problem:

\[
w(t, x) := \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \int_t^T f(X_s^{t,x,\alpha}, \alpha_s)ds + g(X_T^{t,x,\alpha}, \alpha_T) \right], \quad (t, x) \in [0, T] \times \mathbb{R}^d,
\]

where \( \{X_s^{t,x,\alpha}, t \leq s \leq T\} \) denotes the solution to (3.1.12) starting from \( x \) at \( s = t \), given a control \( \alpha \in \mathcal{A} \). It is well-known (see e.g. [37] or [32]) that the value function \( w \) is characterized as the unique viscosity solution to the dynamic programming HJB equation (3.1.6)-(3.1.7), and therefore by Theorem 3.1.3, \( w = v \). In other words, we have provided a representation of fully nonlinear stochastic control problem, including especially control in the diffusion term, possibly degenerate, in terms of minimal solution to the BSDE (3.1.4)-(3.1.5).

2. Combining the BSDE representation of Theorem 3.1.3 together with the dual representation in Theorem 2.3.2, we obtain an original representation for the value function of stochastic control problem:

\[
\sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \int_0^T f(X_t^\alpha, \alpha_t)dt + g(X_T^\alpha, \alpha_T) \right] = \sup_{\nu \in \mathcal{V}} \mathbb{E}^{\nu} \left[ \int_0^T f(X_t, I_t)dt + g(X_T, I_T) \right]
\]

The r.h.s. in the above relation may be viewed as a weak formulation of the stochastic control problem. Indeed, it is well-known that when there is only control on the drift, the value function may be represented in terms of control on change of equivalent probability measures via Girsanov’s theorem for Brownian motion. Such representation is called weak formulation for stochastic control problem, see [14]. In the general case, when there is control on the diffusion coefficient, such “Brownian” Girsanov’s transformation can not be applied, and the idea here is to introduce an exogenous process \( I \) valued in the control set \( A \), independent of \( W \) and \( \vartheta \) governing the controlled state process \( X^\alpha \), and then to control the change of equivalent probability measures through a Girsanov’s transformation on this auxiliary process.

3. Non Markovian extension. An interesting issue is to extend our BSDE representation of stochastic control problem to a non Markovian context, that is when the coefficients \( b, \sigma \) and \( \beta \) of the controlled process are path-dependent. In this case, we know from the recent works by Ekren, Touzi, and Zhang [13] that the value function to the path-dependent stochastic control is a viscosity solution to a path-dependent fully nonlinear HJB equation. One possible approach for getting a BSDE representation to path-dependent stochastic control, would be to prove that our minimal solution to the BSDE with nonpositive jumps is a viscosity solution to the path-dependent fully nonlinear HJB equation, and then to
conclude with a uniqueness result for path-dependent nonlinear PDE. However, to the best of our knowledge, there is not yet such comparison result for viscosity supersolution and subsolution of path-dependent nonlinear PDEs. Instead, we recently proved in [20] by purely probabilistic arguments that the minimal solution to the BSDE with nonpositive jumps is equal to the value function of a path-dependent stochastic control problem, and our approach circumvents the delicate issue of dynamic programming principle and viscosity solution in the non Markovian context. Our result is also obtained without assuming that \( \sigma \) is non degenerate, in contrast with [13] (see their Assumption 4.7).

The rest of this paper is devoted to the proof of Theorem 3.1.3.

### 3.2 Viscosity property of the penalized BSDE

Let us consider the Markov penalized BSDE associated to (3.1.4)-(3.1.5):

\[
Y_t^n = g(X_T, I_T) + \int_t^T f(X_s, I_s, Y_s^n, Z_s^n)ds + n \int_t^T [U_s^n(a)]^+ \lambda(da)ds \\
- \int_t^T Z_s^n dW_s - \int_t^T \int_A U_s^n(a) \tilde{\mu}(ds, da),
\]

(3.2.13)

and denote by \( \{ (Y_{s,t,x,a}^{n,t,x,a}, Z_{s,t,x,a}^{n,t,x,a}, U_{s,t,x,a}^{n,t,x,a}) \}, t \leq s \leq T \) the unique solution to (3.2.13) when \((X, I) = \{(X_{s,t,x,a}^{t,x,a}, I_{s,t,a}^{t,x,a})\}, t \leq s \leq T\) for any initial condition \((t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q\).

From the Markov property of the jump-diffusion process \((X, I)\), we recall from [3] that

\[
Y_{s,t,x,a}^{n,t,x,a} = v_n(s, X_{s,t,x,a}^{t,x,a}, I_{s,t,a}^{t,x,a}), t \leq s \leq T,
\]

where \( v_n \) is the deterministic function defined on \([0, T] \times \mathbb{R}^d \times \mathbb{R}^q\).

\[
v_n(t, x, a) := Y_{s,t,x,a}^{n,t,x,a}, \quad (t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q.
\]

(3.2.14)

From the convergence result (Theorem 2.2.1) of the penalized solution, we deduce that the minimal solution of the constrained BSDE is actually in the form

\[
Y_{s,t,x,a} = v(s, X_{s,t,x,a}^{t,x,a}, I_{s,t,a}^{t,x,a}), t \leq s \leq T,
\]

with a deterministic function \( v \) defined in (3.1.10).

Moreover, from the uniform estimate (2.2.5) and Lemma 3.1.6 there exists some positive constant \( C \) s.t. for all \( n \),

\[
\left| v_n(t, x, a) \right|^2 \leq C \left( \mathbb{E} |g(X_{T,t}^{t,x,a}, I_{T,t}^{t,x,a})|^2 + \mathbb{E} \left[ \int_t^T |f(X_{s,t}^{t,x,a}, I_{s,t}^{t,x,a}, 0, 0)|^2 ds \right] \right) \\
+ \mathbb{E} \left[ \sup_{t \leq s \leq T} |\bar{v}(s, X_{s,t}^{t,x,a})|^2 \right],
\]

for all \((t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q\). From (HBC) we get that \( g \) and \( f \) satisfy a linear growth condition. Using (3.1.8) for \( \bar{v} \) and the estimate (3.1.3) for \((X, I)\), we obtain that \( v_n \), and thus also \( v \) by passing to the limit, satisfy a polynomial growth condition: there exists some positive constant \( C_v \) such that for all \( n \):

\[
\left| v_n(t, x, a) \right| + \left| v(t, x, a) \right| \leq C_v \left( 1 + |x|^2 + |a|^2 \right), \quad \forall (t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q.
\]

(3.2.15)
We now consider the parabolic semi-linear penalized IPDE for any $n$:

$$- \frac{\partial v_n}{\partial t}(t, x, a) - L^a v_n(t, x, a) = f(x, a, v_n, \sigma^\tau(x, a)D_xv_n)$$  \hspace{1cm} (3.2.16)

$$- \int_A [v_n(t, x, a') - v_n(t, x, a)]\lambda(da')$$

$$- n \int_A [v_n(t, x, a') - v_n(t, x, a)]^+\lambda(da') = 0, \text{ on } [0, T) \times \mathbb{R}^d \times \mathbb{R}^q,$$

$$v_n(T, , , ) = g, \text{ on } \mathbb{R}^d \times \mathbb{R}^q. \hspace{1cm} (3.2.17)$$

From Theorem 3.4 in Barles et al. \[3\], we have the well-known property that the penalized BSDE with jumps \[2.2.3\] provides a viscosity solution to the penalized IPDE \(3.2.16)-(3.2.17)\). More precisely, we have the following result.

**Proposition 3.2.2** Let Assumptions (HFC) and (HBC) hold. The function $v_n$ in (3.2.14) is a continuous viscosity solution to (3.2.16)-(3.2.17), i.e. it is continuous on $[0, T) \times \mathbb{R}^d \times \mathbb{R}^q$, a viscosity supersolution (resp. subsolution) to (3.2.17):

$$v_n(T, x, a) \geq (\text{resp. } \leq) g(x, a),$$

for any $(x, a) \in \mathbb{R}^d \times \mathbb{R}^q$, and a viscosity supersolution (resp. subsolution) to (3.2.16):

$$- \frac{\partial \varphi}{\partial t}(t, x, a) - L^a \varphi(t, x, a)$$  \hspace{1cm} (3.2.18)

$$- f(x, a, v_n(t, x, a), \sigma^\tau(x, a)D_x\varphi(t, x, a))$$

$$- \int_A [\varphi(t, x, a') - \varphi(t, x, a)]\lambda(da') - n \int_A [\varphi(t, x, a') - \varphi(t, x, a)]^+\lambda(da') \geq (\text{resp. } \leq) 0,$$

for any $(t, x, a) \in [0, T) \times \mathbb{R}^d \times \mathbb{R}^q$ and any $\varphi \in C^{1,2}([0, T) \times (\mathbb{R}^d \times \mathbb{R}^q))$ such that

$$(v_n - \varphi)(t, x, a) = \min_{[0,T] \times \mathbb{R}^d \times \mathbb{R}^q} (v_n - \varphi) \text{ (resp. } \max_{[0,T] \times \mathbb{R}^d \times \mathbb{R}^q} (v_n - \varphi)), \hspace{1cm} (3.2.19)$$

In contrast to local PDEs with no integro-differential terms, we cannot restrict in general the global minimum (resp. maximum) condition on the test functions for the definition of viscosity supersolution (resp. subsolution) to local minimum (resp. maximum) condition. In our IPDE case, the nonlocal terms appearing in (3.2.16) involve the values w.r.t. the variable $a$ only on the set $A$. Therefore, we are able to restrict the global extremum condition on the test functions to extremum on $[0, T) \times \mathbb{R}^d \times A$. More precisely, we have the following equivalent definition of viscosity solutions, which will be used later.

**Lemma 3.2.7** Assume that (H\lambda), (HFC), and (HBC) hold. In the definition of $v_n$ being a viscosity supersolution (resp. subsolution) to (3.2.16) at a point $(t, x, a) \in [0, T) \times \mathbb{R}^d \times A$, we can replace condition (3.2.19) by:

$$0 = (v_n - \varphi)(t, x, a) = \min_{[0,T] \times \mathbb{R}^d \times A} (v_n - \varphi) \text{ (resp. } \max_{[0,T] \times \mathbb{R}^d \times A} (v_n - \varphi),$$

and suppose that the test function $\varphi$ is in $C^{1,2,0}([0, T) \times \mathbb{R}^d \times \mathbb{R}^q)$.  

26
Proof. We treat only the supersolution case as the subsolution case is proved by the same arguments, and proceed in two steps.

Step 1. Fix \((t,x,a) \in [0,T) \times \mathbb{R}^d \times \mathbb{R}^q\), and let us show that the viscosity supersolution inequality (3.2.18) also holds for any test function \(\varphi\) in \(C^{1,2,0}([0,T] \times \mathbb{R}^d \times \mathbb{R}^q)\) s.t.

\[
(v_n - \varphi)(t,x,a) = \min_{[0,T] \times \mathbb{R}^d \times \mathbb{R}^q} (v_n - \varphi).
\]  

(3.2.20)

We may assume w.l.o.g. that the minimum for such test function \(\varphi\) is zero, and let us define for \(r > 0\) the function \(\varphi^r\) by

\[
\varphi^r(t',x',a') = \varphi(t',x',a') \left(1 - \Phi \left( \frac{|x'|^2 + |a'|^2}{r^2} \right) \right) - C_r \Phi \left( \frac{|x'|^2 + |a'|^2}{r^2} \right) (1 + |x'|^p + |a'|^p),
\]

where \(C_r > 0\) and \(p \geq 2\) are the constant and degree appearing in the polynomial growth condition (3.2.15) for \(v_n\), \(\Phi : \mathbb{R}_+ \to [0,1]\) is a function in \(C^\infty(\mathbb{R}_+)\) such that \(\Phi|_{[0,1]} \equiv 0\) and \(\Phi|_{[2,\infty)} \equiv 1\). Notice that \(\varphi^r \in C^{1,2,0}([0,T] \times \mathbb{R}^d \times \mathbb{R}^q)\),

\[
(\varphi^r, D_x \varphi^r, D_x^2 \varphi^r) \to (\varphi, D_x \varphi, D_x^2 \varphi) \quad \text{as} \quad r \to \infty
\]  

(3.2.21)

locally uniformly on \([0,T] \times \mathbb{R}^d \times \mathbb{R}^q\), and that there exists a constant \(C_r > 0\) such that

\[
|\varphi^r(t',x',a')| \leq C_r (1 + |x'|^p + |a'|^p)
\]  

(3.2.22)

for all \((t',x',a') \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^q\). Since \(\Phi\) is valued in \([0,1]\), we deduce from the polynomial growth conditions (3.2.15) and (3.2.22) satisfied by \(\varphi^r\) on \([0,T] \times \mathbb{R}^d \times \mathbb{R}^q\) for all \(r > 0\). Moreover, we have \(\varphi^r(t,x,a) = \varphi(t,x,a)\) (\(= v_n(t,x,a)\)) for \(r\) large enough. Therefore we get

\[
(v_n - \varphi^r)(t,x,a) = \min_{[0,T] \times \mathbb{R}^d \times \mathbb{R}^q} (v_n - \varphi^r),
\]  

(3.2.23)

for \(r\) large enough, and we may assume w.l.o.g. that this minimum is strict. Let \((\varphi^r_k)\) be a sequence of function in \(C^{1,2,2}([0,T] \times (\mathbb{R}^d \times \mathbb{R}^q))\) satisfying (3.2.22) and such that

\[
(\varphi^r_k, D_x \varphi^r_k, D_x^2 \varphi^r_k) \to (\varphi^r, D_x \varphi^r, D_x^2 \varphi^r) \quad \text{as} \quad k \to \infty,
\]  

(3.2.24)

locally uniformly on \([0,T] \times \mathbb{R}^d \times \mathbb{R}^q\). From the growth conditions (3.2.15) and (3.2.22) on the continuous functions \(v_n\) and \(\varphi^r_k\), we can assume w.l.o.g. (up to an usual negative perturbation of the function \(\varphi^r_k\) for large \((x',a')\)), that there exists a bounded sequence \((t_k,x_k,a_k)\) in \([0,T] \times \mathbb{R}^d \times \mathbb{R}^q\) such that

\[
(v_n - \varphi^r_k)(t_k,x_k,a_k) = \min_{[0,T] \times \mathbb{R}^d \times \mathbb{R}^q} (v_n - \varphi^r_k).
\]  

(3.2.25)

The sequence \((t_k,x_k,a_k)\) converges up to a subsequence, and thus, by (3.2.23), (3.2.24) and (3.2.25), we have

\[
(t_k,x_k,a_k) \to (t,x,a), \quad \text{as} \quad k \to \infty.
\]  

(3.2.26)
Now, from the viscosity supersolution property of \( v_n \) at \((t_k, x_k, a_k)\) with the test function \( \varphi_k^e \), we have

\[
- \frac{\partial \varphi_k^e}{\partial t}(t_k, x_k, a_k) - \mathcal{L}^{a_k} \varphi_k^e(t_k, x_k, a_k)
- f(x_k, a_k, v_n(t_k, x_k, a_k), \sigma^T(x_k, a_k)D\varphi_k^e(t_k, x_k, a_k))
- \int_{\mathcal{A}} [\varphi_k^e(t, x_k, a') - \varphi_k^e(t_k, x_k, a_k)] \lambda(da') \\
- n \int_{\mathcal{A}} [\varphi_k^e(t_k, x_k, a') - \varphi_k^e(t_k, x_k, a_k)]^+ \lambda(da') \geq 0,
\]

Sending \( k \) and \( r \) to infinity, and using (3.2.21), (3.2.24) and (3.2.26), we obtain the viscosity supersolution inequality at \((t, x, a)\) with the test function \( \varphi \).

**Step 2.** Fix \((t, x, a) \in [0, T] \times \mathbb{R}^d \times \hat{A}\), and let \( \varphi \) be a test function in \( C^{1,2}([0, T] \times (\mathbb{R}^d \times \mathbb{R}^q)) \) such that

\[
0 = (v_n - \varphi)(t, x, a) = \min_{[0,T] \times \mathbb{R}^d \times \hat{A}} (v_n - \varphi).
\]  

(3.2.27)

By the same arguments as in (3.2.22), we can assume w.l.o.g. that \( \varphi \) satisfies the polynomial growth condition:

\[
|\varphi(t', x', a')| \leq C(1 + |x'|^p + |a'|^p), \quad (t', x', a') \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q,
\]

for some positive constant \( C \). Together with (3.2.15), and since \( A \) is compact, we have

\[
(v_n - \varphi)(t', x', a') \geq -C(1 + |x'|^p + |d_A(a')|^p),
\]  

(3.2.28)

for all \((t', x', a') \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q\), where \( d_A(a') \) is the distance from \( a' \) to \( A \). Fix \( \varepsilon > 0 \) and define the function \( \varphi_{\varepsilon} \in C^{1,2,0}([0, T] \times \mathbb{R}^d \times \mathbb{R}^q) \) by

\[
\varphi_{\varepsilon}(t', x', a') = \varphi(t', x', a') - \Phi \left( \frac{d_{A_{\varepsilon}}(a')}{\varepsilon} \right) C(1 + |x'|^p + |d_A(a')|^p)
\]

for all \((t', x', a') \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q\), where

\[
A_{\varepsilon} = \{ a' \in A : d_{\partial A}(a') \geq \varepsilon \},
\]  

(3.2.29)

and \( \Phi : \mathbb{R}_+ \to [0, 1] \) is a function in \( C^\infty(\mathbb{R}_+) \) such that \( \Phi_{\varepsilon}|_{[0, \frac{1}{2}]} \equiv 0 \) and \( \Phi_{\varepsilon}|_{[1, +\infty)} \equiv 1 \). Notice that

\[
(\varphi_{\varepsilon}, D_x \varphi_{\varepsilon}, D_x^2 \varphi_{\varepsilon}) \longrightarrow (\varphi, D_x \varphi, D_x^2 \varphi) \quad \text{as} \quad \varepsilon \to 0,
\]  

(3.2.30)

locally uniformly on \([0, T] \times \mathbb{R}^d \times \hat{A}\). We notice from (3.2.28) and the definition of \( \varphi_{\varepsilon} \) that \( \varphi_{\varepsilon} \leq v_n \) on \([0, T] \times \mathbb{R}^d \times A_{\varepsilon} \). Moreover, since \( \varphi_{\varepsilon} \leq \varphi \) on \([0, T] \times \mathbb{R}^d \times \mathbb{R}^q\), \( \varphi_{\varepsilon} = \varphi \) on \([0, T] \times \mathbb{R}^d \times A_{\varepsilon}\) and \( a \in \hat{A} \), we get by (3.2.27) for \( \varepsilon \) small enough

\[
0 = (v_n - \varphi_{\varepsilon})(t, x, a) = \min_{[0,T] \times \mathbb{R}^d \times \mathbb{R}^q} (v_n - \varphi_{\varepsilon}).
\]
From Step 1, we then have
\[
-\frac{\partial \varphi_\varepsilon}{\partial t}(t, x, a) - \mathcal{L}_a \varphi_\varepsilon(t, x, a) - f(x, a, v_\varepsilon(t, x, a), \sigma^T(x, a)D_x \varphi_\varepsilon(t, x, a))
\]
\[
- \int_A [\varphi_\varepsilon(t, x, a') - \varphi_\varepsilon(t, x, a)]\lambda(da') - n \int_A [\varphi_\varepsilon(t, x, a') - \varphi_\varepsilon(t, x, a)]^+\lambda(da') \geq 0.
\]
By sending \(\varepsilon\) to zero with (3.2.30), and using \(a \in \hat{A}\) with \((H\lambda)\)(ii), we get the required viscosity subsolution inequality at \((t, x, a)\) for the test function \(\varphi\).

### 3.3 The non dependence of the function \(v\) in the variable \(a\)

In this subsection, we aim to prove that the function \(v(t, x, a)\) does not depend on \(a\). From the relation defining the Markov BSDE (3.1.4), and since for the minimal solution \((Y^{t,x,a}, Z^{t,x,a}, U^{t,x,a}, K^{t,x,a})\) to (3.1.4)-(3.1.5), the process \(K^{t,x,a}\) is predictable, we observe that the \(A\)-jump component \(U^{t,x,a}\) is expressed in terms of \(Y^{t,x,a} = v(\cdot, X^{t,x,a}, I^{t,x,a})\) as:
\[
U^{t,x,a}_{s^-}(a') = v(s, X^{t,x,a}_{s^-}, a') - v(s, X^{t,x,a}_{s^-}, I^{t,x,a}_{s^-}), \quad t \leq s \leq T, \ a' \in A,
\]
for all \((t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q\). From the \(A\)-nonpositive constraint (3.1.5), this yields
\[
\mathbb{E} \left[ \int_t^{t+h} \int_A [v(s, X^{t,x,a}_{s^-}, a') - v(s, X^{t,x,a}_{s^-}, I^{t,x,a}_{s^-})]^+ \lambda(da')ds \right] = 0,
\]
for any \(h > 0\). If we knew a priori that the function \(v\) was continuous on \([0, T] \times \mathbb{R}^d \times A\), we could obtain by sending \(h\) to zero in the above equality divided by \(h\) (and by dominated convergence theorem), and from the mean-value theorem:
\[
\int_A [v(t, x, a') - v(t, x, a)]^+ \lambda(da') = 0.
\]
Under condition \((H\lambda)\)(i), this would prove that \(v(t, x, a) \geq v(t, x, a')\) for any \(a, a' \in A\), and thus the function \(v\) would not depend on \(a\) in \(A\).

Unfortunately, we are not able to prove directly the continuity of \(v\) from its very definition (3.1.10), and instead, we shall rely on viscosity solutions approach to derive the non dependence of \(v(t, x, a)\) in \(a \in \hat{A}\). To this end, let us introduce the following first-order PDE:
\[
-|D_a v(t, x, a)| = 0, \quad (t, x, a) \in [0, T] \times \mathbb{R}^d \times \hat{A}.
\]

**Lemma 3.3.8** Let assumptions \((H\lambda)\), \((HFC)\) and \((HBC)\) hold. The function \(v\) is a viscosity supersolution to (3.3.31): for any \((t, x, a) \in [0, T] \times \mathbb{R}^d \times \hat{A}\) and any function \(\varphi \in C^{1,2}([0, T] \times (\mathbb{R}^d \times \mathbb{R}^q))\) such that \((v - \varphi)(t, x, a) = \min_{[0, T] \times \mathbb{R}^d \times \mathbb{R}^q}(v - \varphi)\), we have
\[
-|D_a \varphi(t, x, a)| \geq 0, \ i.e. \ D_a \varphi(t, x, a) = 0.
\]
Proof. We know that $v$ is the pointwise limit of the nondecreasing sequence of functions $(v_n)$. By continuity of $v_n$, the function $v$ is lsc and we have (see e.g. [2] p. 91):

$$v = v_* = \liminf_{n \to \infty} v_n, \quad (3.3.32)$$

where

$$\liminf_{n \to \infty} v_n(t, x, a) := \liminf_{n \to \infty} v_n(t', x', a'), \quad (t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q.$$ 

Let $(t, x, a) \in [0, T] \times \mathbb{R}^d \times \hat{A}$, and $\varphi \in C^{1,2}([0, T] \times (\mathbb{R}^d \times \mathbb{R}^q))$, such that $(v - \varphi)(t, x, a) = \min_{[0, T] \times \mathbb{R}^d \times \mathbb{R}^q} (v - \varphi)$. We may assume w.l.o.g. that this minimum is strict:

$$(v - \varphi)(t, x, a) = \text{strict min}_{[0, T] \times \mathbb{R}^d \times \mathbb{R}^q} (v - \varphi). \quad (3.3.33)$$

Up to a suitable negative perturbation of $\varphi$ for large $(x, a)$, we can assume w.l.o.g. that there exists a bounded sequence $(t_n, x_n, a_n)_n$ in $[0, T] \times \mathbb{R}^d \times \mathbb{R}^q$ such that

$$(v_n - \varphi)(t_n, x_n, a_n) = \min_{[0, T] \times \mathbb{R}^d \times \mathbb{R}^q} (v_n - \varphi). \quad (3.3.34)$$

From (4.2.22), (3.3.33), and (3.3.34), we then have, up to a subsequence:

$$(t_n, x_n, a_n, v_n(t_n, x_n, a_n)) \to (t, x, a, v(t, x, a)) \quad \text{as } n \to \infty. \quad (3.3.35)$$

Now, from the viscosity supersolution property of $v_n$ at $(t_n, x_n, a_n)$ with the test function $\varphi$, we have by (3.3.34):

$$-\frac{\partial \varphi}{\partial t}(t_n, x_n, a_n) - \mathcal{L}^a_n \varphi(t_n, x_n, a_n)$$

$$- f(x_n, a_n, v_n(t_n, x_n, a_n), \sigma^T(x_n, a_n)D_x \varphi(t_n, x_n, a_n))$$

$$- \int_A [\varphi(t_n, x_n, a') - \varphi(t_n, x_n, a_n)] \lambda(da')$$

$$- n \int_A [\varphi(t_n, x_n, a') - \varphi(t_n, x_n, a_n)]^+ \lambda(da') \geq 0,$$

which implies

$$\int_A [\varphi(t_n, x_n, a') - \varphi(t_n, x_n, a_n)]^+ \lambda(da')$$

$$\leq \frac{1}{n} \left[ -\frac{\partial \varphi}{\partial t}(t_n, x_n, a_n) - \mathcal{L}^a_n \varphi(t_n, x_n, a_n) 
- f(x_n, a_n, v_n(t_n, x_n, a_n), \sigma^T(x_n, a_n)D_x \varphi(t_n, x_n, a_n)) 
- \int_A [\varphi(t_n, x_n, a') - \varphi(t_n, x_n, a_n)] \lambda(da') \right].$$

Sending $n$ to infinity, we get from (3.3.35), the continuity of coefficients $b, \sigma, \beta$ and $f$, and the dominated convergence theorem:

$$\int_A [\varphi(t, x, a') - \varphi(t, x, a)]^+ \lambda(da') = 0.$$
Under (Hλ), this means that \( \varphi(t, x, a) = \max_{a' \in A} \varphi(t, x, a') \). Since \( a \in \hat{A} \), we deduce that \( D_a \varphi(t, x, a) = 0 \). \( \square \)

We notice that the PDE (3.3.31) involves only differential terms in the variable \( a \). Therefore, we can freeze the terms \( (t, x) \in [0, T) \times \mathbb{R}^d \) in the PDE (3.3.31), i.e. we can take test functions not depending on the variables \( (t, x) \) in the definition of the viscosity solution, as shown in the following Lemma.

**Lemma 3.3.9** Let assumptions (Hλ), (HFC) and (HBC) hold. For any \( (t, x) \in [0, T) \times \mathbb{R}^d \), the function \( v(t, x, .) \) is a viscosity supersolution to

\[
\begin{align*}
-|D_a v(t, x, a)| &= 0, \quad a \in \hat{A},
\end{align*}
\]

i.e. for any \( a \in \hat{A} \) and any function \( \varphi \in C^2(\mathbb{R}^q) \) such that \( (v(t, x, .) - \varphi)(a) = \min_{\mathbb{R}^q} (v(t, x, .) - \varphi) \), we have: \( -|D_a \varphi(a)| \geq 0 \) (and so = 0).

**Proof.** Fix \( (t, x) \in [0, T) \times \mathbb{R}^d \), \( a \in \hat{A} \) and \( \varphi \in C^2(\mathbb{R}^q) \) such that

\[
(v(t, x, .) - \varphi)(a) = \min_{\mathbb{R}^q} (v(t, x, .) - \varphi).
\]

As usual, we may assume w.l.o.g. that this minimum is strict and that \( \varphi \) satisfies the growth condition \( \sup_{a' \in \mathbb{R}^q} \frac{|\varphi(a')|}{1 + |a'|^2} < \infty \). Let us then define for \( n \geq 1 \), the function \( \varphi^n \in C^{1,2}([0, T] \times (\mathbb{R}^d \times \mathbb{R}^q)) \) by

\[
\varphi^n(t', x', a') = \varphi(a') - n(|t' - t|^2 + |x' - x|^4) - |a' - a|^4
\]

for all \( (t', x', a') \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q \). From the growth condition (3.2.15) on the lsc function \( v \), and the growth condition on the continuous function \( \varphi \), one can find for any \( n \geq 1 \) an element \( (t_n, x_n, a_n) \) of \( [0, T] \times \mathbb{R}^d \times \mathbb{R}^q \) such that

\[
(v - \varphi^n)(t_n, x_n, a_n) = \min_{[0,T] \times \mathbb{R}^d \times \mathbb{R}^q} (v - \varphi^n).
\]

In particular, we have

\[
\begin{align*}
v(t, x, a) - \varphi(a) &= (v - \varphi^n)(t, x, a) \geq (v - \varphi^n)(t_n, x_n, a_n) \tag{3.3.37} \\
&= v(t_n, x_n, a_n) - \varphi(a_n) + n(|t_n - t|^2 + |x_n - x|^4) + |a_n - a|^4 \\
&\geq v(t_n, x_n, a_n) - v(t, x, a_n) + v(t, x, a) - \varphi(a) \\
&\quad + n(|t_n - t|^2 + |x_n - x|^4) + |a_n - a|^4
\end{align*}
\]

by (3.3.36), which implies from the growth condition (3.2.15) on \( v \):

\[
n(|t_n - t|^2 + |x_n - x|^4) + |a_n - a|^4 \leq C(1 + |x_n - x|^2 + |a_n - a|^2).
\]

Therefore, the sequences \( \{n(|t_n - t|^2 + |x_n - x|^4)\}_n \) and \( \{|a_n - a|^2\}_n \) are bounded and (up to a subsequence) we have: \( (t_n, x_n, a_n) \rightarrow (t, x, a_\infty) \) as \( n \) goes to infinity, for some \( a_\infty \in \mathbb{R}^q \). Actually, since \( v(t, x, a) - \varphi(a) \geq v(t_n, x_n, a_n) - \varphi(a_n) \) by (3.3.37), we obtain by sending \( n \) to infinity and since the minimum in (3.3.36) is strict, that \( a_\infty = a \), and so:

\[
(t_n, x_n, a_n) \rightarrow (t, x, a) \quad \text{as} \quad n \rightarrow \infty.
\]

31
On the other hand, from Lemma $3.3.8$ applied to $(t_n, x_n, a_n)$ with the test function $\varphi^n$, we have

$$0 = D_a \varphi^n(t_n, x_n, a_n) = D_a \varphi(a_n) - 4(a_n - a)|a_n - a|^3,$$

for all $n \geq 1$. Sending $n$ to infinity we get the required result: $D_a \varphi(a) = 0$. \hfill $\square$

We are now able to state the main result of this subsection.

**Proposition 3.3.3** Let assumptions $(HA)$, $(H\lambda)$, $(HFC)$ and $(HBC)$ hold. The function $v$ does not depend on the variable $a$ on $[0, T) \times \mathbb{R}^d \times \mathcal{A}$:

$$v(t, x, a) = v(t, x, a'), \quad a, a' \in \mathcal{A},$$

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

**Proof.** We proceed in four steps.

**Step 1. Approximation by inf-convolution.**

We introduce the family of functions $(u_n)_n$ defined by

$$u_n(t, x, a) = \inf_{a' \in \mathcal{A}} \left[ v(t, x, a') + n|a - a'|^4 \right], \quad (t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathcal{A}.$$

It is clear that the sequence $(u_n)_n$ is nondecreasing and upper-bounded by $v$. Moreover, since $v$ is lsc, we have the pointwise convergence of $u_n$ to $v$ on $[0, T] \times \mathbb{R}^d \times \mathcal{A}$. Indeed, fix some $(t, x, a) \in [0, T) \times \mathbb{R}^d \times \mathcal{A}$. Since $v$ is lsc, there exists a sequence $(a_n)_n$ valued in $\mathcal{A}$ such that

$$u_n(t, x, a) = v(t, x, a_n) + n|a - a_n|^4,$$

for all $n \geq 1$. Since $\mathcal{A}$ is compact, the sequence $(a_n)_n$ converges, up to a subsequence, to some $a_\infty \in \mathcal{A}$. Moreover, since $u_n$ is upper-bounded by $v$ and $v$ is lsc, we see that $a_\infty = a$ and

$$u_n(t, x, a) \longrightarrow v(t, x, a) \quad \text{as} \quad n \to \infty. \quad (3.3.38)$$

**Step 2. A test function for $u_n$ seen as a test function for $v$.**

For $r, \delta > 0$ let us define the integer $N(r, \delta)$ by

$$N(r, \delta) = \min \left\{ n \in \mathbb{N} : n \geq \frac{2C_v(1 + 2^{-1} + r^p + 2\max_{a \in \mathcal{A}} |a|^2)}{\left(\frac{\delta}{2}\right)^4} + C_v \right\},$$

where $C_v$ is the constant in the growth condition $3.2.15$, and define the set $\mathcal{A}_\delta$ by

$$\mathcal{A}_\delta = \left\{ a \in \mathcal{A} : d(a, \partial \mathcal{A}) := \min_{a' \in \partial \mathcal{A}} |a - a'| > \delta \right\}.$$

Fix $(t, x) \in [0, T) \times \mathbb{R}^d$. We now prove that for any $\delta > 0$, $n \geq N(|x|, \delta)$, $a \in \mathcal{A}_\delta$ and $\varphi \in C^2(\mathbb{R}^d)$ such that

$$0 = (u_n(t, x, .) - \varphi)(a) = \min_{\mathbb{R}^d}(u_n(t, x, .) - \varphi), \quad (3.3.39)$$

32
there exists \( a_n \in \hat{A} \) and \( \psi \in C^2(\mathbb{R}^q) \) such that

\[
0 = (v(t, x, \cdot) - \psi)(a_n) = \min_{\mathbb{R}^n} (v(t, x, \cdot) - \psi),
\]

and

\[
D_a\psi(a_n) = D_a\varphi(a).
\]

To this end we proceed in two substeps.

Substep 2.1. We prove that for any \( \delta > 0 \), \( (t, x, a) \in [0, T) \times \mathbb{R}^d \times \hat{A}_\delta \), and any \( n \geq N(|x|, \delta) \):

\[
\arg\min_{a' \in A} \{v(t, x, a') + n|a' - a|^4\} \subset \hat{A}.
\]

Fix \( (t, x, a) \in [0, T) \times \mathbb{R}^d \times \hat{A}_\delta \) and let \( a_n \in A \) such that

\[
v(t, x, a_n) + n|a_n - a|^4 = \min_{a' \in A} [v(t, x, a') + n|a' - a|^4].
\]

Then we have

\[
v(t, x, a_n) + n|a_n - a|^4 \leq v(t, x, a),
\]

and by (3.2.15), this gives

\[
-C_v(1 + |x|^p + 2\max_{a \in A} |a|^2 + 2|a_n - a|^2) + n|a_n - a|^4 \leq C_v(1 + |x|^2 + |a|^2).
\]

Then using the inequality \( 2\alpha \beta \leq \alpha^2 + \beta^2 \) to the product \( 2\alpha \beta = 2^{p-1}|a_n - a|^p \), we get:

\[
(n - C_v)|a_n - a|^4 \leq 2C_v(1 + 2^{-1} + |x|^2 + 2\max_{a \in A} |a|^2).
\]

For \( n \geq N(|x|, \delta) \), we get from the definition of \( N(r, \delta) \):

\[
|a_n - a| \leq \frac{\delta}{2},
\]

which shows that \( a_n \in \hat{A} \) since \( a \in \hat{A}_\delta \).

Substep 2.2. Fix \( \delta > 0 \), \( (t, x, a) \in [0, T) \times \mathbb{R}^d \times \hat{A}_\delta \), and \( \varphi \in C^2(\mathbb{R}^q) \) satisfying (3.3.39). Let us then choose \( a_n \in \arg\min \{v(t, x, a') + n|a' - a|^4 : a' \in A\} \), and define \( \psi \in C^2(\mathbb{R}^q) \) by:

\[
\psi(a') = \varphi(a + a' - a_n) - n|a_n - a'|^4, \quad a' \in \mathbb{R}^q.
\]

It is clear that \( \psi \) satisfies (3.3.41). Moreover, we have by (3.3.39) and the inf-convolution definition of \( u_n \):

\[
\psi(a') \leq u_n(t, x, a + a' - a_n) - n|a_n - a'|^4 \leq v(t, x, a'), \quad a' \in \mathbb{R}^q.
\]

Moreover, since \( a_n \in \hat{A} \) attains the infimum in the inf-convolution definition of \( u_n(t, x, a) \), we have

\[
\psi(a_n) = \varphi(a) - n|a_n - a|^4 = u_n(t, x, a) - n|a_n - a|^4 = v(t, x, a_n),
\]

33
which shows \((3.3.40)\).

**Step 3.** The function \(u_n\) does not depend locally on the variable \(a\). From Step 2 and Lemma \(3.3.9\), we obtain that for any fixed \((t,x)\in [0,T) \times \mathbb{R}^d\), the function \(u_n(t,x,.)\) inherits from \(v(t,x,.)\) the viscosity supersolution to

\[
-|D_a u_n(t,x,a)| = 0, \quad a \in \mathring{A}_\delta,
\]

for any \(\delta > 0, \ n \geq N(|x|, \delta)\). Let us then show that \(u_n(t,x,.)\) is locally constant in the sense that for all \(a \in \mathring{A}_\delta\):

\[
u_n(t,x,a) = u_n(t,x,a'), \quad \forall a' \in B(a, \eta), \tag{3.3.43}
\]

for all \(\eta > 0\) such that \(B(a, \eta) \subset \mathring{A}_\delta\). We first notice from the inf-convolution definition that \(u_n(t,x,.)\) is semi concave on \(\mathring{A}_\delta\). By Rademacher theorem, this implies that \(u_n(t,x,.)\) is differentiable almost everywhere on \(\mathring{A}_\delta\). Therefore, by Corollary 2 (ii) in [2], and the viscosity supersolution property \((3.3.42)\), we get that this relation \((3.3.42)\) holds actually in the classical sense for almost all \(a' \in \mathring{A}_\delta\). In other words, \(u_n(t,x,.)\) is a locally Lipschitz continuous function with derivatives equal to zero almost everywhere on \(\mathring{A}_\delta\). This means that it is locally constant (easy exercise in analysis left to the reader).

**Step 4.** From the convergence \((3.3.38)\) of \(u_n\) to \(v\), and the relation \((3.3.43)\), we get by sending \(n\) to infinity that for any \(\delta > 0\) the function \(v\) satisfies for any \((t,x,a)\in [0,T) \times \mathbb{R}^d \times \mathring{A}_\delta\)

\[
v(t,x,a) = v(t,x,a')
\]

for all \(\eta > 0\) such that \(B(a, \eta) \subset \mathring{A}_\delta\) and all \(a' \in B(a, \eta)\). Then by sending \(\delta\) to zero we obtain that \(v\) does not depend on the variable \(a\) locally on \([0,T) \times \mathbb{R}^d \times \mathring{A}\). Since \(\mathring{A}\) is assumed to be connex, we obtain that \(v\) does not depend on the variable \(a\) on \([0,T) \times \mathbb{R}^d \times \mathring{A}\).

\[
\square
\]

### 3.4 Viscosity properties of the minimal solution to the constrained BSDE

From Proposition \(3.3.3\) we can define by misuse of notation the function \(v\) on \([0,T) \times \mathbb{R}^d\) by

\[
v(t,x) = v(t,x,a), \ (t,x) \in [0,T) \times \mathbb{R}^d, \tag{3.4.44}
\]

for any \(a \in \mathring{A}\). Moreover, by the growth condition \((3.2.15)\), we have

\[
\sup_{(t,x)\in[0,T] \times \mathbb{R}^d} \frac{|v(t,x)|}{1 + |x|^2} < \infty. \tag{3.4.45}
\]

The aim of this section is to prove that the function \(v\) is a viscosity solution to \((3.1.6)-(3.1.7)\).
Proof of the viscosity supersolution property to (3.1.6). We first notice from (4.2.22) and (3.4.44) that $v$ is lsc and

$$v(t, x) = v_*(t, x) = \liminf_{n \to \infty} v_n(t, x, a) \quad (3.4.46)$$

for all $(t, x, a) \in [0, T] \times \mathbb{R}^d \times \hat{A}$. Let $(t, x)$ be a point in $[0, T] \times \mathbb{R}^d$, and $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^d)$, such that

$$(v - \varphi)(t, x) = \min_{[0, T] \times \mathbb{R}^d} (v - \varphi). \quad (3.4.47)$$

We may assume w.l.o.g. that $\varphi$ satisfies $\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \frac{|\varphi(t,x)|}{1+|x|^p} < \infty$. Fix some $a \in \hat{A}$, and define for $\varepsilon > 0$, the test function

$$\varphi^\varepsilon(t', x', a') = \varphi(t', x') - \varepsilon(|t' - t|^2 + |x' - x|^4 + |a' - a|^4),$$

for all $(t', x', a') \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q$. Since $\varphi^\varepsilon(t, x, a) = \varphi(t, x)$, and $\varphi^\varepsilon \leq \varphi$ with equality iff $(t', x', a') = (t, x, a)$, we then have

$$(v - \varphi^\varepsilon)(t, x, a) = \text{strict min}_{[0, T] \times \mathbb{R}^d \times \mathbb{R}^q} (v - \varphi^\varepsilon). \quad (3.4.47)$$

From the growth conditions on the continuous functions $v_n$ and $\varphi$, there exists a bounded sequence $(t_n, x_n, a_n)_n$ (we omit the dependence in $\varepsilon$) in $[0, T] \times \mathbb{R}^d \times \mathbb{R}^q$ such that

$$(v_n - \varphi^\varepsilon)(t_n, x_n, a_n) = \min_{[0, T] \times \mathbb{R}^d \times \mathbb{R}^q} (v_n - \varphi^\varepsilon). \quad (3.4.48)$$

From (3.4.46) and (3.4.47), we obtain by standard arguments that up to a subsequence:

$$(t_n, x_n, a_n, v_n(t_n, x_n, a_n)) \to (t, x, a, v(t, x)), \quad \text{as } n \text{ goes to infinity.}$$

Now from the viscosity supersolution property of $v_n$ at $(t_n, x_n, a_n)$ with the test function $\varphi^\varepsilon$, we have

$$-\frac{\partial \varphi^\varepsilon}{\partial t}(t_n, x_n, a_n) - \mathcal{L}^a_n \varphi^\varepsilon(t_n, x_n, a_n) - f(x_n, a_n, v_n(t_n, x_n, a_n), \sigma^f(x_n, a_n)D_x \varphi^\varepsilon(t_n, x_n, a_n))$$

$$- \int_A [\varphi^\varepsilon(t_n, x_n, a') - \varphi^\varepsilon(t_n, x_n, a_n)]\lambda(da')$$

$$- n \int_A [\varphi^\varepsilon(t_n, x_n, a') - \varphi^\varepsilon(t_n, x_n, a_n)]^+\lambda(da') \geq 0.$$

Sending $n$ to infinity in the above inequality, we get from the definition of $\varphi^\varepsilon$ and the dominated convergence Theorem:

$$-\frac{\partial \varphi^\varepsilon}{\partial t}(t, x, a) - \mathcal{L}^a \varphi^\varepsilon(t, x, a)$$

$$- f(x, a, v(t, x), \sigma^f(x, a)D_x \varphi^\varepsilon(t, x, a))$$

$$+ \varepsilon \int_A |a' - a|^4 \lambda(da') \geq 0. \quad (3.4.49)$$
Since $a$ is arbitrarily chosen in $\hat{A}$, we get from (HA) and the continuity of the coefficients $b, \sigma, \gamma$ and $f$ in the variable $a$

$$\frac{\partial \varphi}{\partial t}(t, x) - \sup_{a \in \hat{A}} \left[ L^a \varphi(t, x) + f(x, a, v(t, x), \sigma^t(x, a) D_x \varphi(t, x)) \right] \geq 0,$$

which is the viscosity supersolution property.

**Proof of the viscosity subsolution property to (3.1.6).** Since $v$ is the pointwise limit of the nondecreasing sequence of continuous functions $(v_n)$, and recalling (3.4.44), we have by [2] p. 91:

$$v^*(t, x) = \limsup_{n \to \infty} v_n(t, x, a) \tag{3.4.50}$$

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

$$\limsup_{n \to \infty} v_n(t, x, a) := \limsup_{n \to \infty} v_n(t', x', a').$$

Fix $(t, x) \in [0, T] \times \mathbb{R}^d$ and $\varphi \in C^{1,2}([0, T] \times \mathbb{R}^d)$ such that

$$(v^* - \varphi)(t, x) = \max_{[0, T] \times \mathbb{R}^d} (v^* - \varphi). \tag{3.4.51}$$

We may assume w.l.o.g. that this maximum is strict and that $\varphi$ satisfies

$$\sup_{(t, x) \in [0, T] \times \mathbb{R}^d} \frac{1}{1 + |x|^p} < \infty. \tag{3.4.52}$$

Fix $a \in \hat{A}$ and consider a sequence $(t_n, x_n, a_n)_n$ in $[0, T) \times \mathbb{R}^d \times \hat{A}$ such that

$$(t_n, x_n, a_n, v_n(t_n, x_n, a_n)) \to (t, x, a, v^*(t, x)) \text{ as } n \to \infty. \tag{3.4.53}$$

Let us define for $n \geq 1$ the function $\varphi_n \in C^{1,2,0}([0, T] \times \mathbb{R}^d \times \mathbb{R}^d)$ by

$$\varphi_n(t', x', a') = \varphi(t', x') + n \left( \frac{d \Lambda_n(a')}{\eta_n} \wedge 1 + |t' - t_n|^2 + |x' - x_n|^2 \right)$$

where $\Lambda_{n}$ is defined by (3.3.29) for $\varepsilon = \eta_n$ and $(\eta_n)_n$ is a positive sequence converging to 0 s.t. (such sequence exists by (HL)(ii)):

$$n^2 \lambda(A \setminus \Lambda_{n}) \to 0 \text{ as } n \to \infty. \tag{3.4.54}$$

From the growth conditions (3.4.45) and (3.4.52) on $v$ and $\varphi$, we can find a sequence $(\hat{t}_n, \hat{x}_n, \hat{a}_n)$ in $[0, T] \times \mathbb{R}^d \times \hat{A}$ such that

$$(v_n - \varphi_n)(\hat{t}_n, \hat{x}_n, \hat{a}_n) = \max_{[0, T] \times \mathbb{R}^d \times \hat{A}} (v_n - \varphi_n), \quad n \geq 1. \tag{3.4.55}$$
Using (3.4.50) and (3.4.51), we obtain by standard arguments that up to a subsequence
\[
n\left(\frac{1}{\eta_n} d_{A_n}(\bar{a}_n) + |\bar{t}_n - t_n|^2 + |\bar{x}_n - x_n|^4\right) \to 0 \quad \text{as } n \to \infty, \tag{3.4.56}
\]
and
\[
v_n(\bar{t}_n, \bar{x}_n, \bar{a}_n) \to v^*(t, x) \quad \text{as } n \to \infty.
\]
We deduce from (3.4.56) and (3.4.53) that, up to a subsequence:
\[
(\bar{t}_n, \bar{x}_n, \bar{a}_n) \to (t, x, \bar{a}), \quad \text{as } n \to \infty. \tag{3.4.57}
\]
for some \(\bar{a} \in A\). Moreover, for \(n\) large enough, we have \(\bar{a}_n \in \check{A}\). Indeed, suppose that, up to a subsequence, \(\bar{a}_n \in \partial A\) for \(n \geq 1\). Then we have \(\frac{1}{\eta_n} d_{A_n}(\bar{a}_n) \geq 1\), which contradicts (3.4.56).

Now, from the viscosity subsolution property of \(v_n\) at \((\bar{t}_n, \bar{x}_n, \bar{a}_n)\) with the test function \(\varphi_n\) satisfying (3.4.55), Lemma 3.2.7, and since \(\bar{a}_n \in \check{A}\), we have:
\[
- \frac{\partial \varphi_n}{\partial t}(\bar{t}_n, \bar{x}_n, \bar{a}_n) - \mathcal{L}^{\bar{a}_n} \varphi_n(\bar{t}_n, \bar{x}_n, \bar{a}_n) - f(\bar{x}_n, \bar{a}_n, v_n(\bar{t}_n, \bar{x}_n, \bar{a}_n), \sigma^\top(\bar{x}_n, \bar{a}_n)D_x \varphi(\bar{t}_n, \bar{x}_n))
\]
\[
-(n + 1)n \int_A \left(\frac{d_{A_n}(a')}{\eta_n} \wedge 1\right) \lambda(da') \leq 0, \tag{3.4.58}
\]
for all \(n \geq 1\). From (3.4.54) we get
\[
(n + 1)n \int_A \left(\frac{d_{A_n}(a')}{\eta_n} \wedge 1\right) \lambda(da') \to 0 \quad \text{as } n \to \infty \tag{3.4.59}
\]
Sending \(n\) to infinity into (3.4.58), and using (3.4.50), (3.4.57) and (3.4.59), we get
\[
- \frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}^a \varphi(t, x) - f(x, a, v^*(t, x), \sigma^\top(x, a)D_x \varphi(t, x)) \leq 0.
\]
Since \(\bar{a} \in A\), this gives
\[
- \frac{\partial \varphi}{\partial t}(t, x) - \sup_{a \in A} \left[\mathcal{L}^a \varphi(t, x) + f(x, a, v^*(t, x), \sigma^\top(x, a)D_x \varphi(t, x))\right] \leq 0,
\]
which is the viscosity subsolution property.

\textbf{Proof of the viscosity supersolution property to (3.1.7).} Let \((x, a) \in \mathbb{R}^d \times \check{A}\). From (3.4.46), we can find a sequence \((t_n, x_n, a_n)_n\) valued in \([0, T] \times \mathbb{R}^d \times \mathbb{R}^q\) such that
\[
(t_n, x_n, a_n, v_n(t_n, x_n, a_n)) \to (T, x, a, v_*(T, x)) \quad \text{as } n \to \infty.
\]
The sequence of continuous functions \((v_n)_n\) being nondecreasing and \(v_n(T, \cdot) = g\) we have
\[
v_*(T, x) \geq \lim_{n \to \infty} v_1(t_n, x_n, a_n) = g(x, a).
\]
Since \(a\) is arbitrarily chosen in \(\check{A}\), we deduce that \(v_*(T, x) \geq \sup_{a \in A} g(x, a) = \sup_{a \in A} g(x, a)\) by (HA) and continuity of \(g\) in \(a\). \(\Box\)
Proof of the viscosity subsolution property to $\text{(3.1.7)}$. Let $x \in \mathbb{R}^d$. Then we can find by $(3.4.50)$ a sequence $(t_n, x_n, a_n)_n$ in $[0, T) \times \mathbb{R}^d \times A$ such that

$$(t_n, x_n, v_n(t_n, x_n, a_n)) \to (T, x, v^*(T, x)),$$

as $n \to \infty$. (3.4.60)

Define the function $h : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ by

$$h(t, x) = \sqrt{T-t} + \sup_{a \in A} g(x, a)$$

for all $(t, x) \in [0, T) \times \mathbb{R}^d$. From (HFC) and (HBC), we see that $h$ is a continuous viscosity supersolution to $(3.2.16)$ - $(3.2.17)$, on $[T - \eta, T] \times \overline{B}(x, \eta)$ for $\eta$ small enough. We can then apply Theorem 3.5 in [3] which gives that

$$v_n \leq h \quad \text{on} \quad [T - \eta, T] \times \overline{B}(x, \eta) \times A$$

for all $n \geq 0$. By applying the above inequality at $(t_n, x_n, a_n)$, and sending $n$ to infinity, together with $(3.4.60)$, we get the required result. $\square$
Part II

Discretization of fully nonlinear HJB equations via BSDEs with nonpositive jumps
Chapter 4

Discretization of the nonpositive jump constraint

In this chapter, we present an approximation of the constraint imposed to the jump component of the solution to the BSDE. We first express this approximated constraint as a constraint on the component $Y$ operating only at the times on a fixed grid. We then show that the solution satisfying the approximating constraint converges as soon as the time mesh of the grid goes to zero.

4.1 Discretely jump-constrained BSDE

We introduce in this section discretely jump-constrained BSDE. The nonpositive jump constraint operates only at the times of the grid $\pi = \{ t_0 = 0 < t_1 < \ldots < t_n = T \}$ of $[0,T]$, and we look for a quadruple $(Y_\pi, Y_\pi, Z_\pi, U_\pi) \in S^2 \times S^2 \times L^2(W) \times L^2(\tilde{\mu})$ satisfying:

$$Y_T^\pi = Y^\pi_T = g(X_T)$$

(4.1.1)

and

$$Y_t^\pi = Y_{t_{k+1}}^\pi + \int_t^{t_{k+1}} f(X_s, I_s, Y_s^\pi, Z_s^\pi) ds$$

(4.1.2)

$$- \int_t^{t_{k+1}} Z_s^\pi dW_s - \int_t^{t_{k+1}} \int_A U_s^\pi(a) \tilde{\mu}(ds, da),$$

$$Y_t^\pi = Y_t^\pi 1_{(t_k, t_{k+1})}(t) + \text{ess sup}_{a \in A} \mathbb{E}\left[ Y_t^\pi | X_t, I_t = a \right] 1_{(t_k, t_{k+1})}(t),$$

(4.1.3)

for all $t \in [t_k, t_{k+1})$ and all $0 \leq k \leq n - 1$.

Notice that at each time $t_k$ of the grid, the condition is not known a priori to be square integrable since it involves a supremum over $A$, and the well-posedness of the BSDE (4.1.1)-(4.1.2)-(4.1.3) is not a direct and standard issue. We shall use a PDE approach for proving the existence and uniqueness of a solution. Let us consider the system of integro-partial differential equations (IPDEs) for the functions $v^\pi$ and $\vartheta^\pi$ defined recursively on $[0,T] \times \mathbb{R}^d \times A$ by:
• A terminal condition for $v^\pi$ and $\vartheta^\pi$:
\[ v^\pi(T, x, a) = \vartheta^\pi(T, x, a) = g(x), \quad (x, a) \in \mathbb{R}^d \times A, \quad (4.1.4) \]

• A sequence of IPDEs for $\vartheta^\pi$
\[
\left\{
\begin{aligned}
-L^a \vartheta^\pi - f(x, a, \vartheta^\pi, \sigma^\pi(x, a)D_x \vartheta^\pi) \\
-\int_A (\vartheta^\pi(t, x, a') - \vartheta^\pi(t, x, a))\lambda(da') = 0, \\
\vartheta^\pi(t_{k+1}, x, a) = \sup_{a' \in A} \vartheta^\pi(t_k, x, a') \\
(t, x, a) \in [t_k, t_{k+1}) \times \mathbb{R}^d \times A, \\
(x, a) \in \mathbb{R}^d \times A
\end{aligned}
\right. 
\quad (4.1.5)
\]

for $k = 0, \ldots, n - 1$,

• the relation between $v^\pi$ and $\vartheta^\pi$:
\[ v^\pi(t, x, a) = \vartheta^\pi(t, x, a)\mathbb{1}_{(t_k, t_{k+1})}(t) + \sup_{a' \in A} \vartheta^\pi(t, x, a')\mathbb{1}_{\{t_k\}}(t), \quad (4.1.6) \]

for all $t \in [t_k, t_{k+1})$ and $k = 0, \ldots, n - 1$. The rest of this section is devoted to the proof of existence and uniqueness of a solution to $\vartheta^\pi$ for $\vartheta^\pi$ and $\vartheta^\pi$ together with some uniform Lipschitz properties, and its connection to the discretely jump-constrained BSDE $\vartheta^\pi$.

For any $L$-Lipschitz continuous function $\varphi$ on $\mathbb{R}^d \times A$, and $k \leq n - 1$, we denote:
\[ \mathbb{T}_\pi^k[\varphi](t, x, a) := w(t, x, a), \quad (t, x, a) \in [t_k, t_{k+1}) \times \mathbb{R}^d \times A, \quad (4.1.7) \]

where $w$ is the unique continuous viscosity solution on $[t_k, t_{k+1}) \times \mathbb{R}^d \times A$ with linear growth condition in $x$ to the integro partial differential equation (IPDE):
\[
\left\{
\begin{aligned}
-L^a w - f(x, a, w, \sigma^\pi D_x w) \\
-\int_A (w(t, x, a') - w(t, x, a))\lambda(da') = 0, \\
w(t_{k+1}, x, a) = \varphi(x, a), \\
(t, x, a) \in [t_k, t_{k+1}) \times \mathbb{R}^d \times A, \\
(x, a) \in \mathbb{R}^d \times A
\end{aligned}
\right. \quad (4.1.8)
\]

and we extend by continuity $\mathbb{T}_\pi^k[\varphi](t_{k+1}, x, a) = \varphi(x, a)$. The existence and uniqueness of such a solution $w$ to the semi linear IPDE $\vartheta^\pi$, and its nonlinear Feynman-Kac representation in terms of BSDE with jumps, is obtained e.g. from Theorems 3.4 and 3.5 in [3].

**Lemma 4.1.1** There exists a constant $C$ such that for any $L$-Lipschitz continuous function $\varphi$ on $\mathbb{R}^d \times A$, and $k \leq n - 1$, we have
\[ |\mathbb{T}_\pi^k[\varphi](t, x, a) - \mathbb{T}_\pi^k[\varphi](t, x', a')| \leq \max(L, 1)\sqrt{1 + |\pi|^C}(|x - x'| + |a - a'|), \]
for all $t \in [t_k, t_{k+1})$, and $(x, a), (x', a') \in \mathbb{R}^d \times A$.

**Proof.** Fix $t \in [t_k, t_{k+1})$, $k \leq n - 1$, $(x, a), (x', a') \in \mathbb{R}^d \times A$, and $\varphi$ an $L$-Lipschitz continuous function on $\mathbb{R}^d \times A$. Let $(Y^\varphi, Z^\varphi, U^\varphi)$ and $(Y^{\varphi'}, Z^{\varphi'}, U^{\varphi'})$ be the solutions on $[t, t_{k+1}]$ to
the BSDEs
\[
Y_s^\varphi = \varphi(X_{t_k+1}^s, I_{t_k+1}^s) + \int_s^{t_{k+1}} f(X_r^s, I_r^s, Y_r^\varphi, Z_r^\varphi) dr - \int_s^{t_{k+1}} Z_r^\varphi dW_r - \int_s^{t_{k+1}} \int_A U_r^\varphi(a) \tilde{\mu}(dr, da), \quad t \leq s \leq t_{k+1},
\]
\[
Y_s^{\varphi'} = \varphi(X_{t_k+1}^s, I_{t_k+1}^s) + \int_s^{t_{k+1}} f(X_r^s, I_r^s, Y_r^{\varphi'}, Z_r^{\varphi'}) dr - \int_s^{t_{k+1}} Z_r^{\varphi'} dW_r - \int_s^{t_{k+1}} \int_A U_r^{\varphi'}(a) \tilde{\mu}(dr, da), \quad t \leq s \leq t_{k+1}
\]

From Theorems 3.4 and 3.5 in [3], we have the identification:
\[
Y_t^\varphi = \mathbb{T}_t^k[\varphi](t, x, a) \quad \text{and} \quad Y_t^{\varphi'} = \mathbb{T}_t^k[\varphi](t, x', a').
\]

We now estimate the difference between the processes \(Y^\varphi\) and \(Y^{\varphi'}\), and set \(\delta Y^\varphi = Y^\varphi - Y^{\varphi'}\), \(\delta Z^\varphi = Z^\varphi - Z^{\varphi'}\), \(\delta X = X^t,x,a - X^{t,x',a'}\), \(\delta I = I^t,a - I^{t,a'}\). By Itô’s formula, the Lipschitz condition of \(f\) and \(\varphi\), and Young inequality, we have
\[
\mathbb{E}\left[|\delta Y_s^\varphi|^2\right] + \mathbb{E}\left[\int_s^{t_{k+1}} |\delta Z_r^\varphi|^2 dr\right] \leq L^2 \mathbb{E}\left[|\delta X_t|^2 + |\delta I_t|^2\right] + C \int_s^{t_{k+1}} \mathbb{E}\left[|\delta Y_r^\varphi|^2\right] dr + \frac{1}{2} \mathbb{E}\left[\int_s^{t_{k+1}} (|\delta X_r|^2 + |\delta I_r|^2 + |\delta Z_r^\varphi|^2) dr\right],
\]
for any \(s \in [t, t_{k+1}]\). Now, from classical estimates on jump-diffusion processes we have
\[
\mathbb{E}\left[|\delta X_r|^2 + |\delta I_r|^2\right] \leq e^{C|r|}(|x - x'|^2 + |a - a'|^2),
\]
and thus:
\[
\mathbb{E}\left[|\delta Y_s^\varphi|^2\right] \leq (L^2 + |\pi|) e^{C|r|}(|x - x'|^2 + |a - a'|^2) + C \int_s^{t_{k+1}} \mathbb{E}\left[|\delta Y_r^\varphi|^2\right] dr,
\]
for all \(s \in [t, t_{k+1}]\). By Gronwall’s Lemma, this yields
\[
\sup_{s \in [t, t_{k+1}]} \mathbb{E}\left[|\delta Y_s^\varphi|^2\right] \leq (L^2 + |\pi|) e^{2C|r|}(|x - x'|^2 + |a - a'|^2),
\]
which proves the required result from the identification (4.1.9):
\[
|\mathbb{T}_t^k[\varphi](t, x, a) - \mathbb{T}_t^k[\varphi](t, x', a')| \leq \sqrt{L^2 + |\pi|} e^{C|r|}(|x - x'| + |a - a'|) \leq \max(L, 1) \sqrt{1 + |\pi|} e^{C|r|}(|x - x'| + |a - a'|).
\]

\[
\text{Proposition 4.1.1} \quad \text{There exists a unique viscosity solution } \vartheta^\pi \text{ with linear growth condition to the IPDE } (4.1.4) - (4.1.5), \text{ and this solution satisfies:}
\]
\[
|\vartheta^\pi(t, x, a) - \vartheta^\pi(t, x', a')| \leq \max(L, 1) \sqrt{\left(e^{2C|r|} (1 + |\pi|)\right)^{n-k} (|x - x'| + |a - a'|)}, \quad (4.1.10)
\]
for all \(k = 0, \ldots, n - 1\), \(t \in [t_k, t_{k+1}]\), \((x, a), (x', a') \in \mathbb{R}^d \times A\).
\textbf{Proof.} We prove by a backward induction on $k$ that the IPDE (4.1.4)-(4.1.5) admits a unique solution on $[t_k, T] \times \mathbb{R}^d \times A$, which satisfies (4.1.10).

- For $k = n - 1$, we directly get the existence and uniqueness of $\vartheta^\pi$ on $[t_{n-1}, T] \times \mathbb{R}^d \times A$ from Theorems 3.4 and 3.5 in [3], and we have $\vartheta^\pi = \mathbb{T}_n^{-1}[g]$ on $[t_{n-1}, T) \times \mathbb{R}^d \times A$. Moreover, we also get by Lemma 4.1.1

\begin{equation*}
|\vartheta^\pi(t, x, a) - \vartheta^\pi(t, x', a')| \leq \max(L_2, 1) e^{2C|\pi|(1 + |\pi|)}|x - x'| + |a - a'| \quad \text{for all } t \in [t_{n-1}, t_n), (x, a), (x', a') \in \mathbb{R}^d \times A.
\end{equation*}

- Suppose that the result holds true at step $k + 1$ i.e. there exists a unique function $\vartheta^\pi$ on $[t_{k+1}, T] \times \mathbb{R}^d \times A$ with linear growth and satisfying (4.1.4)-(4.1.5) and (4.1.10). It remains to prove that $\vartheta^\pi$ is uniquely determined by (4.1.5) on $[t_k, t_{k+1}) \times \mathbb{R}^d \times A$ and that it satisfies (4.1.10) on $[t_k, t_{k+1}) \times \mathbb{R}^d \times A$. Since $\vartheta^\pi$ satisfies (4.1.10) at time $t_{k+1}$, we deduce that the function

\begin{equation*}
\psi_{k+1}(x) := \sup_{a \in A} \vartheta^\pi(t_{k+1}, x, a), \quad x \in \mathbb{R}^d,
\end{equation*}

is also Lipschitz continuous, and satisfies by the induction hypothesis:

\begin{equation*}
|\psi_{k+1}(x) - \psi_{k+1}(x')| \leq \max(L_2, 1) e^{2C|\pi|(1 + |\pi|)}|x - x'|, \quad (4.1.11)
\end{equation*}

for all $x, x' \in \mathbb{R}^d$. Under (HFC) and (HBC), we can apply Theorems 3.4 and 3.5 in [3], and we get that $\vartheta^\pi$ is the unique viscosity solution with linear growth to (4.1.5) on $[t_k, t_{k+1}) \times \mathbb{R}^d \times A$, with $\vartheta^\pi = \mathbb{T}_k^{t_{k+1}}[\psi_{k+1}]$. Thus it exists and is unique on $[t_k, T] \times \mathbb{R}^d \times A$. From Lemma 4.1.4 and (4.1.11), we then get

\begin{equation*}
|\vartheta^\pi(t, x, a) - \vartheta^\pi(t, x', a')| = |\mathbb{T}_k^t[\vartheta^\pi](t, x, a) - \mathbb{T}_k^t[\vartheta^\pi](t, x', a')| \leq \max(L_2, 1) e^{2C|\pi|(1 + |\pi|)}|x - x'| + |a - a'| \quad \text{for all } t \in [t_k, t_{k+1}) \text{ and } (x, a), (x', a') \in \mathbb{R}^d \times A,
\end{equation*}

which proves the required induction inequality at step $k$. \hfill \Box

\textbf{Remark 4.1.1} The function $\vartheta^\pi(t, x, \cdot)$ is continuous on $A$, for each $(t, x)$, and so the function $\nu^\pi$ is well-defined by (4.1.6). Moreover, the function $\vartheta^\pi$ may be written recursively as:

\begin{equation*}
\left\{
\begin{array}{ll}
\vartheta^\pi(T, \cdot, \cdot) & = g, & \text{on } \mathbb{R}^d \times A, \\
\vartheta^\pi & = \mathbb{T}_k^t[v^\pi(t_{k+1}, \cdot)], & \text{on } [t_k, t_{k+1}) \times \mathbb{R}^d \times A,
\end{array}
\right.
\end{equation*}

for $k = 0, \ldots, n - 1$. In particular, $\vartheta^\pi$ is continuous on $(t_k, t_{k+1}) \times \mathbb{R}^d \times A, k \leq n - 1$. \hfill \Box

As a consequence of the above proposition, we obtain the uniform Lipschitz property of $\vartheta^\pi$ and $\nu^\pi$, with a Lipschitz constant independent of $\pi$. 43
Corollary 4.1.1 There exists a constant $C$ (independent of $|\pi|$) such that

$$|\vartheta^\pi(t, x, a) - \vartheta^\pi(t, x', a')| + |v^\pi(t, x, a) - v^\pi(t, x', a')| \leq C(|x - x'| + |a - a'|),$$

for all $t \in [0, T]$, $x, x' \in \mathbb{R}^d$, $a, a' \in \mathbb{R}^d$.

Proof. Recalling that $n |\pi|$ is bounded, we see that the sequence appearing in (4.1.10):

$$(e^{2C|\pi| (1 + |\pi|)^{n-k}})_{0 \leq k \leq n-1}$$

is bounded uniformly in $|\pi|$ (or $n$), which shows the required Lipschitz property of $\vartheta^\pi$. Since $A$ is assumed to be compact, this shows in particular that the function $v^\pi$ defined by the relation (4.1.6) is well-defined and finite. Moreover, by noting that

$$\sup_{a \in A} \vartheta^\pi(t, x, a) - \vartheta^\pi(t, x', a) \leq \sup_{a \in A} |\vartheta^\pi(t, x, a) - \vartheta^\pi(t, x', a)|$$

for all $(t, x) \in [0, T] \times \mathbb{R}^d$, we also obtain the required Lipschitz property for $v^\pi$. \qed

We now turn to the existence of a solution to the discretely jump-constrained BSDE.

Proposition 4.1.2 The BSDE (4.1.1), (4.1.2), (4.1.3) admits a unique solution $(Y^\pi, Y^\pi, Z^\pi, U^\pi)$ in $S^2 \times S^2 \times L^2(W) \times L^2(\tilde{\mu})$. Moreover we have

$$Y^\pi_t = \vartheta^\pi(t, X_t, I_t), \text{ and } Y^\pi_t = v^\pi(t, X_t, I_t) \quad (4.1.13)$$

for all $t \in [0, T]$.

Proof. We prove by backward induction on $k$ that $(Y^\pi, Y^\pi, Z^\pi, U^\pi)$ is well defined and satisfies (4.1.13) on $[t_k, T]$.

- Suppose that $k = n - 1$. From Corollary 2.3 in [3], we know that $(Y^\pi, Z^\pi, U^\pi)$, exists and is unique on $[t_{n-1}, T]$. Moreover, from Theorems 3.4 and 3.5 in [3], we get $Y^\pi = T^k_{\pi}[g](t, X_t, I_t) = \vartheta^\pi(t, X_t, I_t)$ on $[t_{n-1}, T]$. By (4.1.3), we then have for all $t \in [t_{n-1}, T]$:

$$Y^\pi_t = \mathbb{1}_t(t_{n-1}, T) \vartheta^\pi(t, X_t, I_t) + \mathbb{1}_{t_{n-1}}(t) \sup_{a \in A} \vartheta^\pi(t, X_t, a)$$

$$= \mathbb{1}_t(t_{n-1}, T) \vartheta^\pi(t, X_t, I_t) + \mathbb{1}_{t_{n-1}}(t) \sup_{a \in A} \vartheta^\pi(t, X_t, a) = v^\pi(t, X_t, I_t),$$

since the essential supremum and supremum coincide by continuity of $a \mapsto \vartheta^\pi(t, X_t, a)$ on the compact set $A$.

- Suppose that the result holds true for some $k \leq n - 1$. Then, we see that $(Y^\pi, Z^\pi, U^\pi)$ is defined on $[t_{k-1}, t_k)$ as the solution to a BSDE driven by $W$ and $\tilde{\mu}$ with a terminal condition $v^\pi(t_k, X_{t_k})$. Since $v^\pi$ satisfies a linear growth condition, we know again by Corollary 2.3 in [3] that $(Y^\pi, Z^\pi, U^\pi)$, thus also $Y^\pi$, exists and is unique on $[t_{k-1}, t_k)$. Moreover, using again Theorems 3.4 and 3.5 in [3], we get (4.1.13) on $[t_{k-1}, t_k)$.

We end this section with a conditional regularity result for the discretely jump-constrained BSDE.
Proposition 4.1.3  There exists some constant $C$ such that

$$\sup_{t \in [t_k, t_{k+1}]} \mathbb{E}_{t_k} [ |Y_{t_k}^\pi - Y_{t_k}^\sigma|^2 ] + \sup_{t \in [t_k, t_{k+1}]} \mathbb{E}_{t_k} [ |Y_{t}^\pi - Y_{t+1}^{\pi, \sigma}|^2 ] \leq C(1 + |X_{t_k}|^2)|\pi|,$$

for all $k = 0, \ldots, n - 1$.

Proof. Fix $k \leq n - 1$. By Itô's formula, we have for all $t \in [t_k, t_{k+1}]:$

$$\mathbb{E}_{t_k} [ |Y_{t_k}^\pi - Y_{t_k}^\sigma|^2 ] = 2\mathbb{E}_{t_k} \left[ \int_{t_k}^t f(X_s, I_s, Y_s^\pi, Z_s^\pi) (Y_{t_k}^\pi - Y_s^\sigma) ds \right] + \mathbb{E}_{t_k} \left[ \int_{t_k}^t |Z_s^\pi|^2 ds \right] + \mathbb{E}_{t_k} \left[ \int_{t_k}^t \int_A |U_s^\pi(a)|^2 \lambda(da) ds \right],$$

$$\leq \mathbb{E}_{t_k} \left[ \int_{t_k}^t |Y_{t_k}^\pi - Y_s^\sigma|^2 ds \right] + C|\pi| \mathbb{E}_{t_k} \left[ \sup_{s \in [t_k, t_{k+1}]} |X_s|^2 \right] + C|\pi| \mathbb{E}_{t_k} \left[ \sup_{s \in [t_k, t_{k+1}]} \left( |Y_{t_k}^\sigma|^2 + |Z_s^\sigma|^2 + \int_A |U_s^\sigma(a)|^2 \lambda(da) \right) \right],$$

by the linear growth condition on $f$ (recall also that $A$ is compact), and Young inequality. Now, by standard estimate for $X$ under growth linear condition on $b$ and $\sigma$, we have:

$$\mathbb{E}_{t_k} \left[ \sup_{s \in [t_k, t_{k+1}]} |X_s|^2 \right] \leq C(1 + |X_{t_k}|^2). \quad (4.1.14)$$

We also know from Proposition 4.2 in [8], under (H1) and (H2), that there exists a constant $C$ depending only on the Lipschitz constants of $b, \sigma f$ and $\nu^\pi(t_{k+1}, \cdot)$ (which does not depend on $\pi$ by Corollary 4.1.1), such that

$$\mathbb{E}_{t_k} \left[ \sup_{s \in [t_k, t_{k+1}]} \left( |Y_{t_k}^\sigma|^2 + |Z_s^\sigma|^2 + \int_A |U_s^\sigma(a)|^2 \lambda(da) \right) \right] \leq C(1 + |X_{t_k}|^2). \quad (4.1.15)$$

We deduce that

$$\mathbb{E}_{t_k} [ |Y_{t_k}^\pi - Y_{t_k}^\sigma|^2 ] \leq \mathbb{E}_{t_k} \left[ \int_{t_k}^t |Y_{t_k}^\pi - Y_s^\sigma|^2 ds \right] + C|\pi|(1 + |X_{t_k}|^2),$$

and we conclude for the regularity of $Y^\pi$ by Gronwall’s lemma. Finally, from the definition (4.1.2)-(4.1.3) of $Y^\pi$ and $Y^\sigma$, Itô isometry for stochastic integrals, and growth linear condition on $f$, we have for all $t \in (t_k, t_{k+1}):

$$\mathbb{E}_{t_k} [ |Y_{t}^\pi - Y_{t+1}^\pi|^2 ] = \mathbb{E}_{t_k} [ |Y_{t_k}^\pi - Y_{t_k+1}^\pi|^2 ] \leq 3\mathbb{E}_{t_k} \left[ \int_{t_k}^{t_{k+1}} \left( |f(X_s, I_s, Y_s^\pi, Z_s^\pi)|^2 + |Z_s^\pi|^2 + \int_A |U_s^\pi(a)|^2 \lambda(da) \right) ds \right],$$

$$\leq C|\pi| \mathbb{E}_{t_k} \left[ 1 + \sup_{s \in [t_k, t_{k+1}]} \left( |X_s|^2 + |Y_s|^2 + |Z_s|^2 + \int_A |U_s(a)|^2 \lambda(da) \right) \right] \leq C|\pi|(1 + |X_{t_k}|^2),$$

where we used again (4.1.14) and (4.1.15). This ends the proof. □
4.2 Convergence of discretely jump-constrained BSDE

This section is devoted to the convergence of the discretely jump-constrained BSDE towards the minimal solution to the BSDE with nonpositive jump.

Under \( (HFC) \) and \( (HBC) \), we have seen in Chapter 3 the existence and uniqueness of a minimal solution \( (Y,Z,U,K) \in S^2 \times L^2(W) \times L^2(\tilde{\mu}) \times L^2(\tilde{\mu}) \) to

\[
\begin{align*}
Y_t &= g(X_T) + \int_t^T f(X_s, I_s, Y_s, Z_s) ds + K_T - K_t \\
&\quad - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \tilde{\mu}(ds, da), \quad 0 \leq t \leq T. \\
U_t(a) &\leq 0 \quad dP \otimes dt \otimes \lambda(da) \text{ a.e.}
\end{align*}
\] (4.2.16)

Moreover, the minimal solution \( Y \) is in the form

\[
Y_t = v(t, X_t), \quad 0 \leq t \leq T,
\] (4.2.17)

where \( v : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R} \) is a viscosity solution with linear growth to the fully nonlinear HJB type equation:

\[
\begin{align*}
-\sup_{a \in A} \left[ L^a v + f(x, a, \sigma^T(x, a) D_x v) \right] &= 0, \text{ on } [0, T) \times \mathbb{R}^d, \\
v(T, x) &= g, \text{ on } \mathbb{R}^d,
\end{align*}
\] (4.2.18)

where

\[
L^a v = \frac{\partial v}{\partial t} + b(x, a) \cdot D_x v + \frac{1}{2} \text{tr}(\sigma \sigma^T(x, a) D^2_x v).
\]

We shall make the standing assumption that comparison principle holds for (4.2.18).

\( (HC) \) Let \( \bar{w} \) (resp. \( w \)) be a lower-semicontinuous (resp. upper-semicontinuous) viscosity supersolution (resp. subsolution) with linear growth condition to (4.2.18). Then, \( \bar{w} \geq w \).

When \( f \) does not depend on \( y, z \), i.e. (4.2.18) is the usual HJB equation for a stochastic control problem, Assumption (HC) holds true, see [18] or [38]. In the general case, we refer to [12] for sufficient conditions to comparison principles. Under (HC), the function \( v \) in (4.2.17) is the unique viscosity solution to (4.2.18), and is in particular continuous. Actually, we have the standard Hölder and Lipschitz property (see Appendix in [29] or [5]):

\[
|v(t, x) - v(t', x')| \leq C(|t - t'|^{\frac{1}{2}} + |x - x'|), \quad (t, t') \in [0, T], x, x' \in \mathbb{R}^d. \quad (4.2.19)
\]

This implies that the process \( Y \) is continuous, and thus the jump component \( U = 0 \). In the sequel, we shall focus on the approximation of the remaining components \( Y \) and \( Z \) of the minimal solution to (4.2.16).

4.2.1 Convergence result

**Lemma 4.2.2** We have the following assertions:

1) The family \( (\vartheta^\pi)_\pi \) is nondecreasing and upper bounded by \( v \): for any grids \( \pi \) and \( \pi' \) such that \( \pi \subset \pi' \), we have

\[
\vartheta^\pi(t, x, a) \leq \vartheta^{\pi'}(t, x, a) \leq v(t, x), \quad (t, x, a) \in [0, T] \times \mathbb{R}^d \times A.
\]

46
2) The family \((\vartheta^n)_n\) satisfies a uniform linear growth condition: there exists a constant \(C\) such that

\[
|\vartheta^n(t, x, a)| \leq C(1 + |x|),
\]

for any \((t, x, a) \in [0, T] \times \mathbb{R}^d \times A\) and any grid \(\pi\).

**Proof.** 1) Let us first prove that \(\vartheta^n \leq v\). Since \(v\) is a (continuous) viscosity solution to the HJB equation (4.2.18), and \(v\) does not depend on \(a\), we see that \(v\) is a viscosity supersolution to the IPDE in (4.1.5) satisfied by \(\vartheta^n\) on each interval \([t_k, t_{k+1})\). Now, since \(v(T, x) = \vartheta^n(T, x, a)\), we deduce by comparison principle for this IPDE (see e.g. Theorem 3.4 in [3]) on \([t_{n-1}, T] \times \mathbb{R}^d \times A\) that \(v(t, x) \geq \vartheta^n(t, x, a)\) for all \(t \in [t_{n-1}, T]\), \((x, a) \in \mathbb{R}^d \times A\). In particular, \(v(t_{n-1}, x) = v(t_{n-1}, x) \geq \sup_{a \in A} \vartheta^n(t_{n-1}, x, a) = \vartheta^n(t_{n-1}, x, a)\). Again, by comparison principle for the IPDE (4.1.5) on \([t_{n-2}, t_{n-1}) \times \mathbb{R}^d \times A\), it follows that \(v(t, x) \geq \vartheta^n(t, x, a)\) for all \(t \in [t_{n-2}, t_{n-1}), (x, a) \in \mathbb{R}^d \times A\). By backward induction on time, we conclude that \(v \geq \vartheta^n\) on \([0, T] \times \mathbb{R}^d \times A\).

Let us next consider two partitions \(\pi = (t_k)_{0 \leq k \leq n}\) and \(\pi' = (t'_k)_{0 \leq k \leq n'}\) of \([0, T]\) with \(\pi \subset \pi'\), and denote by \(m = \max\{k \leq n' : t'_m \notin \pi\}\). Thus, all the points of the grid \(\pi\) and \(\pi'\) coincide after time \(t'_m\), and since \(\vartheta^n\) and \(\vartheta^m\) are viscosity solution to the same IPDE (4.1.5) starting from the same terminal data \(g\), we deduce by uniqueness that \(\vartheta^n = \vartheta^m\) on \([t'_m, T] \times \mathbb{R}^d \times A\). Then, we have \(\vartheta^n(t'_m, x, a) = \sup_{a \in A} \vartheta^n(t'_m, x, a) = \sup_{a \in A} \vartheta^m(t'_m, x, a) \geq \vartheta^m(t'_m, x, a)\) since \(\vartheta^n\) is continuous outside of the points of the grid \(\pi\) (recall Remark 4.1.1). Now, since \(\vartheta^n\) and \(\vartheta^m\) are viscosity solution to the same IPDE (4.1.5) on \([t'_m-1, t'_m)\), we deduce by comparison principle that \(\vartheta^m \geq \vartheta^n\) on \([t'_m-1, t'_m) \times \mathbb{R}^d \times A\). Proceeding by backward induction, we conclude that \(\vartheta^m \geq \vartheta^n\) on \([0, T] \times \mathbb{R}^d \times A\).

2) Denote by \(\pi_0 = \{t_0 = 0, t_1 = T\}\) the trivial grid of \([0, T]\). Since \(\vartheta^{\pi_0} \leq \vartheta^n \leq v\) and \(\vartheta^{\pi_0}\) and \(v\) satisfy a linear growth condition, we get (recall that \(A\) is compact):

\[
|\vartheta^n(t, x, a)| \leq |\vartheta^{\pi_0}(t, x, a)| + |v(t, x)| \leq C(1 + |x|),
\]

for any \((t, x, a) \in [0, T] \times \mathbb{R}^d \times A\) and any grid \(\pi\). \(\square\)

In the sequel, we denote by \(\vartheta\) the increasing limit of the sequence \((\vartheta^n)_n\) when the grid increases by becoming finer, i.e. its modulus \(|\pi|\) goes to zero. The next result shows that \(\vartheta\) does not depend on the variable \(a\) in \(A\).

**Proposition 4.2.4** The function \(\vartheta\) is l.s.c. and does not depend on the variable \(a \in A\):

\[
\vartheta(t, x, a) = \vartheta(t, x, a'), \quad (t, x) \in [0, T] \times \mathbb{R}^d, \ a, a' \in A.
\]

To prove this result we use the following lemma. Observe by definition (4.1.6) of \(v^n\) that the function \(v^n\) does not depend on \(a\) on the grid times \(\pi\), and we shall denote by misuse of notation: \(v^n(t_k, x)\), for \(k \leq n, x \in \mathbb{R}^d\).

**Lemma 4.2.3** There exists a constant \(C\) (not depending on \(\pi\)) such that

\[
|\vartheta^n(t, x, a) - v^n(t_{k+1}, x)| \leq C(1 + |x|)|\pi|^{\frac{1}{2}}
\]

for all \(k = 0, \ldots, n - 1, t \in [t_k, t_{k+1}), (x, a) \in \mathbb{R}^d \times A\).
Proof. Fix \( k = 0, \ldots, n - 1 \), \( t \in [t_k, t_{k+1}) \) and \( (x, a) \in \mathbb{R}^d \times A \). Let \( (\tilde{Y}, \tilde{Z}, \tilde{U}) \) be the solution to the BSDE

\[
\tilde{Y}_s = v^\pi(t_{k+1}, X_{t_{k+1}}^t, a) + \int_s^{t_{k+1}} f(X_{s}^{t,x,a}, I_{s}^{t,a}, \tilde{Y}_s, \tilde{Z}_s) \, ds - \int_s^{t_{k+1}} \tilde{Z}_s \, dW_s - \int_s^{t_{k+1}} \int_A \tilde{U}_s(a') \tilde{\mu}(ds, da'), \quad s \in [t, t_{k+1}].
\]

From Proposition \( 4.1.2 \), Markov property and uniqueness of a solution to the BSDE \( (4.1.1)-(4.1.3) \) we have: \( \tilde{Y}_s = \vartheta^\pi(s, X_{t_{k+1}}^t, a) \), for \( s \in [t, t_{k+1}] \), and so:

\[
|\vartheta^\pi(t, x, a) - \vartheta^\pi(t_{k+1}, x)| = |\tilde{Y}_t - \vartheta^\pi(t_{k+1}, x)| \\
\leq \mathbb{E}|\vartheta^\pi(t_{k+1}, X_{t_{k+1}}^t, a) - \vartheta^\pi(t_{k+1}, x)| \\
\quad + \mathbb{E}\left[ \int_t^{t_{k+1}} |f(X_{s}^{t,x,a}, I_{s}^{t,a}, \tilde{Y}_s, \tilde{Z}_s)| \, ds \right]. \quad (4.2.20)
\]

From Corollary \( 4.1.1 \) we have

\[
\mathbb{E}|\vartheta^\pi(t_{k+1}, X_{t_{k+1}}^t, a) - \vartheta^\pi(t_{k+1}, x)| \leq C \sqrt{\mathbb{E}(|X_{t_{k+1}}^t - x|^2)} \leq C \sqrt{|x|}. \quad (4.2.21)
\]

Moreover, by the growth linear condition on \( f \) in \( (H2) \), and on \( \vartheta^\pi \) in Lemma \( 4.2.2 \) we have

\[
\mathbb{E}\left[ \int_t^{t_{k+1}} |f(X_{s}^{t,x,a}, I_{s}^{t,a}, \tilde{Y}_s, \tilde{Z}_s)| \, ds \right] \leq CE\left[ \int_t^{t_{k+1}} \left( 1 + |X_{s}^{t,x,a}| + |\tilde{Z}_s| \right) \, ds \right].
\]

By classical estimates, we have

\[
\sup_{s \in [t, T]} \mathbb{E}|X_{s}^{t,x,a}|^2 \leq C(1 + |x|^2).
\]

Moreover, under \( (H1) \) and \( (H2) \), we know from Proposition 4.2 in [3] that there exists a constant \( C \) depending only on the Lipschitz constants of \( b, \sigma \) and \( v^\pi(t_{k+1}, ) \) such that

\[
\mathbb{E}\left[ \sup_{s \in [t_k, t_{k+1}]} |\tilde{Z}_s|^2 \right] \leq C(1 + |x|^2).
\]

This proves that

\[
\mathbb{E}\left[ \int_t^{t_{k+1}} |f(X_{s}^{t,x,a}, I_{s}^{t,a}, \tilde{Y}_s, \tilde{Z}_s)| \, ds \right] \leq C(1 + |x|)|x|.
\]

Combining this last estimate with \( (4.2.20) \) and \( (4.2.21) \), we get the result \( \square \)

Proof of Proposition \( 4.2.4 \) The function \( \vartheta \) is l.s.c. as the supremum of the l.s.c. functions \( \vartheta^\pi \). Fix \( (t, x) \in [0, T) \times \mathbb{R}^d \) and \( a, a' \in A \). Let \( (\pi^p)_p \) be a sequence of subdivisions of \( [0, T] \) such that \( |\pi^p| \downarrow 0 \) as \( p \uparrow \infty \). We define the sequence \( (t_p)_p \) of \( [0, T] \) by

\[
t_p = \min \{ s \in \pi^p : s > t \}, \quad p \geq 0.
\]
Since $|\pi_p| \to 0$ as $p \to \infty$ we get $t_p \to t$ as $p \to +\infty$. We then have from the previous lemma:

$$|\varphi^{\pi_p}(t, x, a) - \varphi^{\pi_p}(t, x, a')| \leq |\varphi^{\pi_p}(t, x, a) - v^{\pi_p}(t_p, x)| + |v^{\pi_p}(t_p, x) - \varphi^{\pi_p}(t, x, a')|$$

$$\leq 2C|\pi_p|^\frac{1}{2}.$$

Sending $p$ to $\infty$ we obtain that $\vartheta(t, x, a) = \vartheta(t, x, a')$.

\[ \square \]

**Corollary 4.2.2** We have the identification: $\vartheta = v$, and the sequence $(v^\pi)_\pi$ also converges to $v$.

**Proof.** We proceed in two steps.

**Step 1.** The function $\vartheta$ is a supersolution to (4.2.18). Since $\varphi^{\pi_k}(T, \cdot) = g$ for all $k \geq 1$, we first notice that $\vartheta(T, \cdot) = g$. Next, since $\vartheta$ does not depend on the variable $a$, we have

$$\varphi^\pi(t, x, a) \uparrow \vartheta(t, x) \text{ as } |\pi| \downarrow 0$$

for any $(t, x, a) \in [0, T] \times \mathbb{R}^d \times A$. Moreover, since the function $\vartheta$ is l.s.c, we have

$$\vartheta = \vartheta_* = \lim\inf_{|\pi| \to 0} \varphi^\pi,$$

where

$$\lim\inf_{|\pi| \to 0} \varphi^\pi(t, x, a) := \lim\inf_{|\pi| \to 0} \varphi^\pi(t', x', a'), \quad (t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q.$$

Fix now some $(t, x) \in [0, T] \times \mathbb{R}^d$ and $a \in A$ and $(p, q, M) \in J^2 - \vartheta(t, x)$, the limiting parabolic subjet of $\vartheta$ at $(t, x)$ (see definition in [12]). From standard stability results, there exists a sequence $(\pi_k, t_k, x_k, a_k, p_k, q_k, M_k)_k$ such that

$$(p_k, q_k, M_k) \in J^2 - \varphi^{\pi_k}(t_k, x_k, a_k)$$

for all $k \geq 1$ and

$$(t_k, x_k, a_k, \varphi^{\pi_k}(t_k, x_k, a_k)) \to (t, x, a, \vartheta(t, x, a)) \text{ as } k \to \infty, \ |\pi_k| \to 0.$$

From the viscosity supersolution property of $\varphi^{\pi_k}$ to (4.1.5) in terms of subjets, we have

$$-p_k - b(x, a_k)q_k - \frac{1}{2}\text{tr}(\sigma \sigma^T(x, a_k)M_k) - f(x, a_k, \varphi^{\pi_k}(t_k, x, a_k), \sigma^T(x, a_k)q_k)$$

$$\quad - \int_A (\varphi^{\pi_k}(t_k, x_k, a') - \varphi^{\pi_k}(t_k, x_k, a_k)) \lambda(da') \geq 0$$

for all $k \geq 1$. Sending $k$ to infinity and using (4.2.22), we get

$$-p - b(x, a)q - \frac{1}{2}\text{tr}(\sigma \sigma^T(x, a)M) - f(x, a, \vartheta(t, x), \sigma^T(x, a)q) \geq 0.$$
Since $a$ is arbitrary in $A$, this shows
\[ -p - \sup_{a \in A} \left[ b(x,a)q + \frac{1}{2} \text{tr}(\sigma^T(x,a)M) + f(x,a,\vartheta(t,x),\sigma(x,a)q) \right] \geq 0, \]
i.e. the viscosity supersolution property of $\vartheta$ to (4.2.18).

**Step 2. Comparison.** Since the PDE (4.2.18) satisfies a comparison principle, we have from the previous step $\vartheta \geq v$, and we conclude with Lemma 4.2.2 that $\vartheta = v$. Finally, by definition (4.1.6) of $v^\pi$ and from Lemma 4.2.2 we clearly have $\vartheta^\pi \leq v^\pi \leq v$, which also proves that $(v^\pi)_\pi$ converges to $v$. \(\square\)

In terms of the discretely jump-constrained BSDE, the convergence result is formulated as follows:

**Proposition 4.2.5** We have $\mathcal{Y}^\pi_t \leq Y^\pi_t \leq Y_t$, 0 \leq t \leq T, and
\[ \mathbb{E}\left[ \sup_{t \in [0,T]} |Y_t - \mathcal{Y}^\pi_t|^2 \right] + \mathbb{E}\left[ \sup_{t \in [0,T]} |Y_t - Y^\pi_t|^2 \right] + \mathbb{E}\left[ \int_0^T |\delta Z_t - Z^\pi_t|^2 dt \right] \rightarrow 0, \]
as $|\pi|$ goes to zero.

**Proof.** Recall from (4.2.17) and (4.1.13) that we have the representation:
\[ Y_t = v(t,X_t), \quad Y^\pi_t = \vartheta(t,X_t,I_t), \quad \mathcal{Y}^\pi_t = \vartheta(t,X_t,I_t), \quad (4.2.23) \]
and the first assertion of Lemma (4.2.2), we clearly have: $\mathcal{Y}^\pi_t \leq Y^\pi_t \leq Y_t$, 0 \leq t \leq T. The convergence in $S^2$ for $Y^\pi$ to $Y$ and $Y^\pi$ to $Y$ comes from the above representation (4.2.23), the pointwise convergence of $\vartheta^\pi$ and $v^\pi$ to $v$ in Corollary 4.2.2 and the dominated convergence theorem by recalling that $0 \leq (v - v^\pi)(t,x,a) \leq (v - \vartheta^\pi)(t,x,a) \leq v(t,x) \leq C(1 + |x|)$. Let us now turn to the component $Z$. By definition (4.1.1)-(4.1.2)-(4.1.3) of the discretely jump-constrained BSDE we notice that $Y^\pi$ can be written on $[0,T]$ as:
\[ \mathcal{Y}^\pi_t = g(X_T) + \int_t^T f(X_s,I_s,Y_s^\pi,Z_s^\pi) - \int_t^T Z^\pi_s dW_s - \int_t^T \int_A U^\pi_s(a)\mu(ds,da) + \mathcal{K}^\pi_T - \mathcal{K}^\pi_t, \]
where $\mathcal{K}^\pi$ is the nondecreasing process defined by: $\mathcal{K}^\pi_t = \sum_{k \leq t} (Y^\pi_{t_k} - \mathcal{Y}^\pi_{t_k})$, for $t \in [0,T]$. Denote by $\delta Y = Y - Y^\pi$, $\delta Z = Z - Z^\pi$, $\delta U = U - U^\pi$ and $\delta K = K - \mathcal{K}^\pi$. From Itô’s formula, Young Inequality and (H2), there exists a constant $C$ such that
\[ \mathbb{E}\left[ |\delta Y_t|^2 \right] \leq C \int_t^T \mathbb{E}\left[ |\delta Z_s|^2 ds \right] + \frac{1}{2} \mathbb{E}\left[ \int_t^T |\delta U_s|^2 \lambda(da) ds \right] \]
\[ \leq C \int_t^T \mathbb{E}\left[ |\delta Y_s|^2 ds \right] + \frac{1}{\varepsilon} \mathbb{E}\left[ \sup_{s \in [0,T]} |\delta Y_s|^2 \right] + \varepsilon \mathbb{E}\left[ |\delta K_T - \delta K_t|^2 \right] \quad (4.2.24) \]
for all $t \in [0,T]$, with $\varepsilon$ a constant to be chosen later. From the definition of $\delta K$ we have
\[ \delta K_T - \delta K_t = \delta Y_t - \int_t^T \left( f(X_s,I_s,Y_s,Z_s) - f(X_s,I_s,Y^\pi_s,Z^\pi_s) \right) ds \]
\[ + \int_t^T \delta Z_s dW_s + \int_t^T \int_A \delta U_s(a)\mu(ds,da), \]
\[ \leq \int_t^T \mathbb{E}\left[ |\delta Y_s|^2 ds \right] + \frac{1}{\varepsilon} \mathbb{E}\left[ \sup_{s \in [0,T]} |\delta Y_s|^2 \right] + \varepsilon \mathbb{E}\left[ |\delta K_T - \delta K_t|^2 \right] \quad (4.2.24) \]
for all $t \in [0,T]$, with $\varepsilon$ a constant to be chosen later. From the definition of $\delta K$ we have
\[ \delta K_T - \delta K_t = \delta Y_t - \int_t^T \left( f(X_s,I_s,Y_s,Z_s) - f(X_s,I_s,Y^\pi_s,Z^\pi_s) \right) ds \]
\[ + \int_t^T \delta Z_s dW_s + \int_t^T \int_A \delta U_s(a)\mu(ds,da), \]
\[ \leq \int_t^T \mathbb{E}\left[ |\delta Y_s|^2 ds \right] + \frac{1}{\varepsilon} \mathbb{E}\left[ \sup_{s \in [0,T]} |\delta Y_s|^2 \right] + \varepsilon \mathbb{E}\left[ |\delta K_T - \delta K_t|^2 \right] \quad (4.2.24) \]
for all $t \in [0,T]$, with $\varepsilon$ a constant to be chosen later. From the definition of $\delta K$ we have
\[ \delta K_T - \delta K_t = \delta Y_t - \int_t^T \left( f(X_s,I_s,Y_s,Z_s) - f(X_s,I_s,Y^\pi_s,Z^\pi_s) \right) ds \]
\[ + \int_t^T \delta Z_s dW_s + \int_t^T \int_A \delta U_s(a)\mu(ds,da), \]
Therefore, by (H2), we get the existence of a constant $C'$ such that
\[
\mathbb{E}
\left[
\left|
\delta K_T - \delta K_t \right|^2
\right]
\leq
C' \left( \mathbb{E} \left[ \sup_{s \in [0,T]} \left| \delta Y_s \right|^2 \right] + \mathbb{E} \left[ \int_t^T \left| \delta Z_s \right|^2 ds \right] + \mathbb{E} \left[ \int_t^T \left| \delta U_s(a) \right|^2 \lambda(da) ds \right] \right)
\]
Taking $\varepsilon = \frac{C'}{4}$ and plugging this last inequality in (4.2.24), we get the existence of a constant $C''$ such that
\[
\mathbb{E} \left[ \int_t^T \left| \delta Z_s \right|^2 ds \right] + \mathbb{E} \left[ \int_t^T \left| \delta U_s(a) \right|^2 \lambda(da) ds \right] \leq C'' \left( \mathbb{E} \left[ \sup_{s \in [0,T]} \left| \delta Y_s \right|^2 \right] \right),
\]
which shows the $L^2(W)$ convergence of $Z^\pi$ to $Z$ from the $S^2$ convergence of $Y^\pi$ to $Y$.

**4.2.2 Rate of convergence**

We next provide an error estimate for the convergence of the discretely jump-constrained BSDE. We shall combine BSDE methods and PDE arguments adapted from the shaking coefficients approach of Krylov [29] and switching systems approximation of Barles, Jacobsen [5]. We make further assumptions:

(HFC') The functions $b$ and $\sigma$ are uniformly bounded:
\[
\sup_{x \in \mathbb{R}^d, a \in A} |b(x,a)| + |\sigma(x,a)| < \infty.
\]

(HBC') The function $f$ does not depend on $z$: $f(x,a,y,z) = f(x,a,y)$ for all $(x,a,y,z) \in \mathbb{R}^d \times A \times \mathbb{R} \times \mathbb{R}^d$ and

(i) the functions $f(.,.,0)$ and $g$ are uniformly bounded:
\[
\sup_{x \in \mathbb{R}^d, a \in A} |f(x,a,0)| + |g(x)| < \infty,
\]

(ii) for all $(x,a) \in \mathbb{R}^d \times A$, $y \mapsto f(x,a,y)$ is convex.

Under these assumptions, we obtain the rate of convergence for $v^\pi$ and $\vartheta^\pi$ towards $v$.

**Theorem 4.2.1** Under (HFC') and (HBC'), there exists a constant $C$ such that
\[
0 \leq v(t,x) - v^\pi(t,x,a) \leq v(t,x) - \vartheta^\pi(t,x,a) \leq C|\pi|^{\frac{1}{10}}
\]
for all $(t,x,a) \in [0,T] \times \mathbb{R}^d \times A$ and all grid $\pi$ with $|\pi| \leq 1$. Moreover, when $f(x,a)$ does not depend on $y$, the rate of convergence is improved to $|\pi|^{\frac{1}{5}}$.

Before proving this result, we give as corollary the rate of convergence for the discretely jump-constrained BSDE.

**Corollary 4.2.3** Under (HFC') and (HBC'), there exists a constant $C$ such that
\[
\mathbb{E} \left[ \sup_{t \in [0,T]} |Y_t - Y_t^\pi|^2 \right] + \mathbb{E} \left[ \sup_{t \in [0,T]} |Y_t - Y_t^\pi|^2 \right] + \mathbb{E} \left[ \int_0^T |Z_t - Z_t^\pi|^2 dt \right] \leq C|\pi|^{\frac{1}{5}}.
\]
for all grid $\pi$ with $|\pi| \leq 1$, and the above rate is improved to $|\pi|^{\frac{1}{3}}$ when $f(x,a)$ does not depend on $y$. 

51
Proof. From the representation (4.2.23), and Theorem 4.2.1, we immediately have
\[ E \left[ \sup_{t \in [0,T]} |Y_t - Y^\pi_t|^2 \right] + E \left[ \sup_{t \in [0,T]} |Y_t - Y^\pi^1_t|^2 \right] \leq C|\pi|^{\frac{1}{10}}, \quad (4.2.26) \]
(resp. $|\pi|^{\frac{1}{12}}$ when $f(x,a)$ does not depend on $y$). Finally, the convergence rate for $Z$ follows from the inequality (4.2.25). \square

Remark 4.2.2 The above convergence rate $|\pi|^{\frac{1}{10}}$ is the optimal rate that one can prove in our generalized stochastic control context with fully nonlinear HJB equation by PDE approach and shaking coefficients technique, see [29], [5], [17] or [42]. However, this rate may not be the sharpest one. In the case of continuously reflected BSDEs, i.e. BSDEs with upper or lower constraint on $Y$, it is known that $Y$ can be approximated by discretely reflected BSDEs, i.e. BSDEs where reflection on $Y$ operates a finite set of times on the grid $\pi$, with a rate $|\pi|^{\frac{1}{2}}$ (see [1]). The standard arguments for proving this rate is based on the representation of the continuously (resp. discretely) reflected BSDE as optimal stopping problems where stopping is possible over the whole interval time (resp. only on the grid times). In our jump-constrained case, we know from Chapter [3] that, when $f(x,a)$ does not depend on $y$ and $z$, the minimal solution to the BSDE with nonpositive jumps has the stochastic control representation
\[ v(t,x) = \sup_{\alpha} \mathbb{E} \left[ \int_t^T f(X^\alpha_s, \alpha_s) ds + g(X^\alpha_T) \Big| X^\alpha_t = x \right], \quad (4.2.27) \]
with controlled diffusion in $\mathbb{R}^d$:
\[ dX^\alpha_t = b(X^\alpha_t, \alpha_t) dt + \sigma(X^\alpha_t, \alpha_t) dW_t, \]
and where $\alpha$ is an adapted control process valued in $A$. We shall prove an analog representation for discretely jump-constrained BSDEs, and this helps to improve the rate of convergence from $|\pi|^{\frac{1}{10}}$ to $|\pi|^{\frac{1}{5}}$. \square

The rest of this section is devoted to the proof of Theorem 4.2.1. We first consider the special case where $f(x,a)$ does not depend on $y$, and then address the case $f(x,a,y)$.

Proof of Theorem 4.2.1 in the case $f(x,a)$. In the case where $f(x,a)$ does not depend on $y$, by (linear) Feynman-Kac formula for $\varphi^\pi$ solution to (4.1.5), and by definition of $v^\pi$ in (4.1.6), we have:
\[ v^\pi(t_k,x) = \sup_{\alpha \in A} \mathbb{E} \left[ \int_{t_k}^{t_{k+1}} f(X^\alpha_t, \alpha_t, I^\alpha_t) dt + v^\pi(t_{k+1}, X^\alpha_{t_{k+1}}) \right], \quad k \leq n - 1, \quad x \in \mathbb{R}^d. \]
By induction, this dynamic programming relation leads to the following stochastic control problem with discrete time policies:
\[ v^\pi(t_k,x) = \sup_{\alpha \in A^\pi_k} \mathbb{E} \left[ \int_{t_k}^T f(\tilde{X}^\alpha_t, \tilde{I}^\alpha_t) dt + g(\tilde{X}^\alpha_T) \right], \]
where $\mathcal{A}_F^n$ is the set of discrete time processes $\alpha = (\alpha_t)_j \leq n-1$, with $\alpha_t$ $\mathcal{F}_t$-measurable, valued in $A$, and
\[
\hat{X}^{t_k,x,\alpha}_t = x + \int_{t_k}^{t} b(\hat{X}^{s,x,\alpha}_s, \hat{I}^{\alpha}_s)ds + \int_{t_k}^{t} \sigma(\hat{X}^{s,x,\alpha}_s, \hat{I}^{\alpha}_s)dW_s, \quad t_k \leq t \leq T,
\]
\[
\hat{I}^{\alpha}_t = \alpha_{t_j} + \int_{(t_j,t]} \int_{A} (a - \hat{I}^{\alpha}_{s^-})\mu(ds,da), \quad t_j \leq t < t_{j+1}, j \leq n-1.
\]

In other words, $v^\pi(t_k, x)$ corresponds to the value function for a stochastic control problem where the controller can act only at the dates $t_j$ of the grid $\pi$, and then let the regime of the coefficients of the diffusion evolve according to the Poisson random measure $\mu$. Let us introduce the following stochastic control problem with piece-wise constant control policies:
\[
\tilde{v}^\pi(t_k, x) = \sup_{\alpha \in \mathcal{A}_F^n} \mathbb{E}\left[ \int_{t_k}^{T} f(\hat{X}^{s,x,\alpha}_s, \hat{I}^{\alpha}_s)ds + g(\hat{X}^{t_k,x,\alpha}_T) \right],
\]
where for $\alpha = (\alpha_t)_j \leq n-1 \in \mathcal{A}_F^n$:
\[
\hat{X}^{t_k,x,\alpha}_t = x + \int_{t_k}^{t} b(\hat{X}^{s,x,\alpha}_s, \hat{I}^{\alpha}_s)ds + \int_{t_k}^{t} \sigma(\hat{X}^{s,x,\alpha}_s, \hat{I}^{\alpha}_s)dW_s, \quad t_k \leq t \leq T,
\]
\[
\hat{I}^{\alpha}_t = \alpha_{t_j}, \quad t_j \leq t < t_{j+1}, j \leq n-1.
\]

It is shown in [28] that $\tilde{v}^\pi$ approximates the value function $v$ for the controlled diffusion problem (4.2.27), solution to the HJB equation (4.2.18), with a rate $|\pi|^{\frac{1}{2}}$:
\[
0 \leq v(t_k, x) - \tilde{v}^\pi(t_k, x) \leq C|\pi|^{\frac{1}{2}}, \quad (4.2.28)
\]
for all $t_k \in \pi, x \in \mathbb{R}^d$. Now, recalling that $A$ is compact and $\lambda(A) < \infty$, it is clear that there exists some positive constant $C$ such that for all $\alpha \in \mathcal{A}_F^n, j \leq n-1$:
\[
\mathbb{E}\left[ \sup_{t \in [t_j,t_{j+1}]} |\hat{I}^{\alpha}_t - \hat{I}^{\alpha}_t|^2 \right] \leq C|\pi|,
\]
and then by standard arguments under Lipschitz condition on $b, \sigma$:
\[
\mathbb{E}\left[ \sup_{t \in [t_j,t_{j+1}]} |\hat{X}^{t_k,x,\alpha}_t - \hat{X}^{t_k,x,\alpha}_t|^2 \right] \leq C|\pi|, \quad k \leq j \leq n-1, x \in \mathbb{R}^d.
\]

By the Lipschitz conditions on $f$ and $g$, it follows that
\[
|v^\pi(t_k, x) - \tilde{v}^\pi(t_k, x)| \leq C|\pi|^{\frac{1}{2}},
\]
and thus with (4.2.28):
\[
0 \leq \sup_{x \in \mathbb{R}^d} (v - v^\pi)(t_k, x) \leq C|\pi|^{\frac{1}{2}}.
\]

Finally, by combining with the estimate in Lemma [4.2.3] which gives actually under (HBC') (i):
\[
|\vartheta^\pi(t, x, a) - v^\pi(t_{k+1}, x)| \leq C|\pi|^{\frac{1}{2}}, \quad t \in [t_k, t_{k+1}), (x, a) \in \mathbb{R}^d \times A,
\]

53
together with the $1/2$-Hölder property of $v$ in time (see (4.2.19)), we obtain:

$$
\sup_{(t,x,a) \in [0,T] \times \mathbb{R}^d \times A} (v - \vartheta^\pi)(t, x, a) \leq C(|\pi|^{\frac{1}{6}} + |\pi|^{\frac{1}{2}}) \leq C|\pi|^{\frac{1}{6}},
$$

for $|\pi| \leq 1$. This ends the proof.

Let us now turn to the case where $f(x,a,y)$ may also depend on $y$. We cannot rely anymore on the convergence rate result in [28]. Instead, recalling that $A$ is compact and since $\sigma, b$ and $f$ are Lipschitz in $(x,a)$, we shall apply the switching system method of Barles and Jacobsen [5], which is a variation of the shaken coefficients method and smoothing technique used in Krylov [29], in order to obtain approximate smooth subsolution to (4.2.18). By Lemmas 3.3 and 3.4 in [5], one can find a family of smooth functions $(w_\varepsilon)_{0<\varepsilon \leq 1}$ on $[0,T] \times \mathbb{R}^d$ such that:

$$
\sup_{[0,T] \times \mathbb{R}^d} |w_\varepsilon| \leq C, \quad (4.2.29)
$$

$$
\sup_{[0,T] \times \mathbb{R}^d} |w_\varepsilon - w| \leq C\varepsilon^{\frac{1}{2}}, \quad (4.2.30)
$$

$$
\sup_{[0,T] \times \mathbb{R}^d} |\partial^\beta_\tau D^\beta w_\varepsilon| \leq C\varepsilon^{1-2\beta_0-\sum_{i=1}^d \beta_i}, \quad \beta_0 \in \mathbb{N}, \quad \beta = (\beta^1, \ldots, \beta^d) \in \mathbb{N}^d, \quad (4.2.31)
$$

for some positive constant $C$ independent of $\varepsilon$, and by convexity of $f$ in (HBC')(ii), for any $\varepsilon \in (0,1]$, $(t,x) \in [0,T] \times \mathbb{R}^d$, there exists $a_{t,x,\varepsilon} \in A$ such that:

$$
-L^a_{t,x,\varepsilon} w_\varepsilon(t,x) - f(x,a_{t,x,\varepsilon},w_\varepsilon(t,x)) \geq 0. \quad (4.2.32)
$$

Recalling the definition of the operator $T^k_\pi$ in (4.1.7), we define for any function $\varphi$ on $[0,T] \times \mathbb{R}^d \times A$, Lipschitz in $(x,a)$:

$$
T_\pi[\varphi](t,x,a) := T^k_\pi[\varphi(t_{k+1},\ldots)](t,x,a), \quad t \in [t_k,t_{k+1}), (x,a) \in \mathbb{R}^d \times A,
$$

for $k = 0,\ldots, n-1$, and

$$
S_\pi[\varphi](t,x,a) := \frac{1}{|\pi|} \left[ \varphi(t,x) - T_\pi[\varphi](t,x,a) + (t_{k+1} - t)(L^a_\varphi(t,x) + f(x,a,\varphi(t,x))) \right],
$$

for $(t,x,a) \in [t_k,t_{k+1}) \times \mathbb{R}^d \times A$, $k \leq n-1$.

We have the following key error bound on $S_\pi$.

**Lemma 4.2.4** Let (HFC') and (HBC')(i) hold. There exists a constant $C$ such that

$$
|S_\pi[\varphi_\varepsilon](t,x,a)| \leq C \left( |\pi|^{\frac{1}{6}}(1 + \varepsilon^{-1}) + |\pi|\varepsilon^{-3} \right), \quad (t,x,a) \in [0,T] \times \mathbb{R}^d \times A,
$$

for any family $(\varphi_\varepsilon)_\varepsilon$ of smooth functions on $[0,T] \times \mathbb{R}^d$ satisfying (4.2.29) and (4.2.31).
From Theorems 3.4 and 3.5 in [3], we have

\[ Y_t \in (4.2.33), \]  we thus get:

Define \( Z_t \) and we study each term by the Lipschitz continuity of \( t < T \) and fix \( \epsilon \) such that \( t \in [t_k, t_{k+1}) \). Given a smooth function \( \varphi_\epsilon \) satisfying (4.2.29) and (4.2.31), we split:

\[ E \left[ \varphi_\epsilon(t) \right] - \varphi_\epsilon(t) = (t_k+1-t) \mathcal{C}_a \varphi_\epsilon(t), \]

and so:

\[ \mathcal{E} \left[ \varphi_\epsilon(t) \right] - \varphi_\epsilon(t) = (t_k+1-t) \mathcal{E} \varphi_\epsilon(t), \]

and we study each term \( \mathcal{E}_\epsilon \) and \( B_\epsilon \) separately.

1. **Estimate on \( A_\epsilon(t, x, a) \).**

Define \( (Y^\varphi_\epsilon, Z^\varphi_\epsilon, U^\varphi_\epsilon) \) as the solution to the BSDE on \( [t, t_{k+1}) \):

\[
Y^\varphi_\epsilon_s = \varphi_\epsilon(t_{k+1}, X^{t,x,a}_{t_{k+1}}) + \int_s^{t_{k+1}} f(X^t,x,a, I^t,a, Y^\varphi_\epsilon) \, dr \\
- \int_s^{t_{k+1}} Z^\varphi_\epsilon \, dW_r - \int_s^{t_{k+1}} \int_A U^\varphi_\epsilon(a) \mu(da, dr), \quad s \in [t, t_{k+1}].
\]

From Theorems 3.4 and 3.5 in [3], we have \( Y_t^\varphi_\epsilon = \mathbb{E}_\pi[\varphi_\epsilon(t, x, a)] \), and by taking expectation in (4.2.33), we thus get:

\[
Y^\varphi_\epsilon_t = \mathbb{E}_\pi[\varphi_\epsilon(t, x, a)] = \mathbb{E}[\varphi_\epsilon(t_{k+1}, X^{t,x,a}_{t_{k+1}})] + \mathbb{E} \left[ \int_t^{t_{k+1}} f(X^t,x,a, I^t,a, Y^\varphi_\epsilon) \, ds \right]
\]

and so:

\[
A_\epsilon(t, x, a) \leq \frac{1}{|\pi|} \mathbb{E} \left[ \int_t^{t_{k+1}} |f(X^t,x,a, I^t,a, Y^\varphi_\epsilon) - f(x, a, \varphi_\epsilon(t))| \, ds \right]
\]

\[
\leq C \left( \mathbb{E} \left[ \sup_{s \in [t, t_{k+1}]} |X^t,x,a - x| + |I^t,a - a| \right] + \mathbb{E} \left[ \sup_{s \in [t, t_{k+1}]} |Y^\varphi_\epsilon_s - \varphi_\epsilon(t, x)| \right] \right),
\]

by the Lipschitz continuity of \( f \). From standard estimate for SDE, we have (recall that the coefficients \( b \) and \( \sigma \) are bounded under (HFC') and \( A \) is compact):

\[
\mathbb{E} \left[ \sup_{s \in [t, t_{k+1}]} |X^t,x,a - x| + |I^t,a - a| \right] \leq C |\pi|^\frac{1}{2}.
\]

Moreover, by (4.2.33), the boundedness condition in (HBC')(ii) together with the Lipschitz condition of \( f \), and Burkholder-Davis-Gundy inequality, we have:

\[
\mathbb{E} \left[ \sup_{s \in [t, t_{k+1}]} |Y^\varphi_\epsilon_s - \varphi_\epsilon(t, x)| \right] \leq \mathbb{E} \left[ |\varphi_\epsilon(t_{k+1}, X^{t,x,a}_{t_{k+1}}) - \varphi_\epsilon(t, x)| \right]
\]

\[
+ C |\pi| \mathbb{E} \left[ 1 + \sup_{s \in [t, t_{k+1}]} |Y^\varphi_\epsilon_s| \right]
\]

\[
+ C |\pi| \left( \mathbb{E} \left[ \sup_{s \in [t, t_{k+1}]} |Z^\varphi_\epsilon|^2 \right] + \mathbb{E} \left[ \sup_{s \in [t, t_{k+1}]} \int_A |U^\varphi_\epsilon(a)|^2 \lambda(da) \right] \right).
\]

55
From standard estimate for the BSDE (4.2.33), we have:

\[ E\left[ \sup_{s \in [t,t_{k+1}]} |Y_s^{\varphi_\varepsilon}|^2 \right] \leq C, \]

for some positive constant \( C \) depending only on the Lipschitz constant of \( f \), the upper bound of \(|f(x, a, 0, 0)|\) in (HBC\(^*\))(i), and the upper bound of \(|\varphi_\varepsilon|\) in (4.2.29). Moreover, from the estimate in Proposition 4.2 in \([8]\) about the coefficients \( Z^{\varphi_\varepsilon} \) and \( U^{\varphi_\varepsilon} \) of the BSDE with jumps (4.2.33), there exists some constant \( C \) depending only on the Lipschitz constant of \( b, \sigma, f \), and of the Lipschitz constant of \( \varphi_\varepsilon(t_{k+1}, \cdot) \) (which does not depend on \( \varepsilon \) by (4.2.31)), such that:

\[ E\left[ \sup_{s \in [t,t_{k+1}]} |Z_s^{\varphi_\varepsilon}|^2 \right] + E\left[ \sup_{s \in [t,t_{k+1}]} \int_A |U_s^{\varphi_\varepsilon}(a)|^2 \lambda(da) \right] \leq C. \]

From (4.2.31), we then have:

\[ E\left[ \sup_{s \in [t,t_{k+1}]} |Y_s^{\varphi_\varepsilon} - \varphi_\varepsilon(t,x)| \right] \leq C(|t_{k+1} - t|\varepsilon^{-1} + E[|X_{t_{k+1}}^{t,x,a} - x|] + |\pi|) \]

\[ \leq C|\pi|^{\frac{1}{2}}(1 + \varepsilon^{-1}), \]

by (4.2.34). This leads to the error bound for \( A_\varepsilon(t,x,a) \):

\[ A_\varepsilon(t,x,a) \leq C|\pi|^{\frac{1}{2}}(1 + \varepsilon^{-1}). \]

2. \textit{Estimate on } \( B_\varepsilon(t,x,a) \).

From Itô’s formula we have

\[ B_\varepsilon(t,x,a) = \frac{1}{|\pi|} E\left[ \int_t^{t_{k+1}} (L_s^{t,x,a} \varphi_\varepsilon(s, X_s^{t,x,a}) - L_s^{t,x,a}(t,x)) ds \right] \]

\[ \leq B_1^\varepsilon(t,x,a) + B_2^\varepsilon(t,x,a) \]

where

\[ B_1^\varepsilon(t,x,a) = \frac{1}{|\pi|} E\left[ \int_t^{t_{k+1}} \left( b(X_s^{t,x,a}, I_s^{t,a}) - b(x,a)) D_x \varphi_\varepsilon(s, X_s^{t,x,a}) \right. \]

\[ + \frac{1}{2} \text{tr}\left[ \left( \sigma \sigma^\top(X_s^{t,x,a}, I_s^{t,a}) - \sigma \sigma^\top(x,a) \right) D_x^2 \varphi_\varepsilon(t,x) \right] ds \]

and

\[ B_2^\varepsilon(t,x,a) = \frac{1}{|\pi|} E\left[ \int_t^{t_{k+1}} \left| \tilde{L}_{t,x,a} \varphi_\varepsilon(t,x,a) - \tilde{L}_{t,x,a} \varphi_\varepsilon(t,x) \right| ds \right], \]

with \( \tilde{L}_{t,x,a} \) defined by

\[ \tilde{L}_{t,x,a} \varphi_\varepsilon(t',x') = \frac{\partial \varphi_\varepsilon}{\partial t}(t',x') + b(x,a) D_x \varphi_\varepsilon(t',x') + \frac{1}{2} \text{tr}\left( \sigma \sigma^\top(x,a) D_x^2 \varphi_\varepsilon(t',x') \right). \]

Under (HFC), (HFC\(^*\)), and by (4.2.31), we have

\[ B_1^\varepsilon(t,x,a) \leq C(1 + \varepsilon^{-1}) E\left[ \sup_{s \in [t,t_{k+1}]} |X_s^{t,x,a} - x| + |I_s^{t,a} - a| \right] \]

\[ \leq C(1 + \varepsilon^{-1})|\pi|^{\frac{1}{2}}, \]
where we used again (4.2.34). On the other hand, since \( \varphi_\varepsilon \) is smooth, we have from Itô’s formula

\[
B_\varepsilon^2(t, x, a) = \frac{1}{|\pi|} \mathbb{E} \left[ \int_t^{t_k+1} \left| \int_t^s \mathcal{L} \mathcal{L} \phi(r, X^r_{t,x,a}) dr \right| ds \right].
\]

Under (HFC') and by (4.2.31), we then see that

\[
B_\varepsilon^2(t, x, a) \leq C|\pi|\varepsilon^{-3},
\]

and so:

\[
B_\varepsilon(t, x, a) \leq C \left( |\pi|^{\frac{1}{2}} (1 + \varepsilon^{-1}) + |\pi|\varepsilon^{-3} \right).
\]

Together with the estimate for \( A_\varepsilon(t, x, a) \), this proves the error bound for \( |S_\pi[\varphi_\varepsilon](t, x, a)| \).

We next state a maximum principle type result for the operator \( \mathbb{T}_\pi \).

**Lemma 4.2.5** Let \( \varphi \) and \( \psi \) be two functions on \([0, T] \times \mathbb{R}^d \times A\), Lipschitz in \((x, a)\). Then, there exists some positive constant \( C \) independent of \( \pi \) such that

\[
\sup_{(t, x, a) \in [t_k, t_{k+1}] \times \mathbb{R}^d \times A} (\mathbb{T}_\pi[\varphi] - \mathbb{T}_\pi[\psi])(t, x, a) \leq C|\pi| \sup_{(x, a) \in \mathbb{R}^d \times A} (\varphi - \psi)(t_{k+1}, x, a),
\]

for all \( k = 0, \ldots, n-1 \).

**Proof.** Fix \( k \leq n-1 \), and set

\[
M := \sup_{(x, a) \in \mathbb{R}^d \times A} (\varphi - \psi)(t_{k+1}, x, a).
\]

We can assume w.l.o.g. that \( M < \infty \) since otherwise the required inequality is trivial. Let us denote by \( \Delta v \) the function

\[
\Delta v(t, x, a) = \mathbb{T}_\pi[\varphi](t, x, a) - \mathbb{T}_\pi[\psi](t, x, a),
\]

for all \((t, x, a) \in [t_k, t_{k+1}] \times \mathbb{R}^d \times A\). By definition of \( \mathbb{T}_\pi \), and from the Lipschitz condition of \( f \), we see that \( \Delta v \) is a viscosity subsolution to

\[
\begin{cases}
-\mathcal{L} \Delta v(t, x, a) - C(\Delta v(t, x, a) + |D \Delta v(t, x, a)|) \\
- \int_A (∆v(t, x, a') - ∆v(t, x, a)) λ(da') = 0, & \text{for } (t, x, a) \in [t_k, t_{k+1}] \times \mathbb{R}^d \times A,
\end{cases}
\]

\[
\Delta v(t_{k+1}, x, a) \leq M, \quad \text{for } (x, a) \in \mathbb{R}^d \times A.
\]

(4.2.35)

Then, we easily check that the function \( \Phi \) defined by

\[
\Phi(t, x, a) = Me^{C(t_{k+1} - t)}, \quad (t, x, a) \in [t_k, t_{k+1}] \times \mathbb{R}^d \times A,
\]

is a solution to

\[
\begin{cases}
-\mathcal{L} \Phi(t, x, a) - C(\Phi(t, x, a) + |D \Phi(t, x, a)|) \\
- \int_A (∆\Phi(t, x, a') - ∆\Phi(t, x, a)) λ(da') = 0, & \text{for } (t, x, a) \in [t_k, t_{k+1}] \times \mathbb{R}^d \times A,
\end{cases}
\]

\[
\Phi(t_{k+1}, x, a) = M, \quad \text{for } (x, a) \in \mathbb{R}^d \times A.
\]

(4.2.36)
From the comparison theorem in \cite{3} for viscosity solutions of semi-linear IPDEs, we get that $\Delta v \leq \Phi$ on $[t_k, t_{k+1}] \times \mathbb{R}^d \times A$, which proves the required inequality.

\textbf{Proof of Theorem 4.2.1}. By (4.1.6) and (4.1.12), we observe that $v^\pi$ is a fixed point of $T_\pi$, i.e.

$$T_\pi[v^\pi] = v^\pi.$$ 

On the other hand, by (4.2.32), and the estimate of Lemma 4.2.4 applied to $w_\pi$, we have:

$$w_\pi(t, x) - T_\pi[w_\pi](t, x, a_t, x, \varepsilon) \leq |\pi| S_\pi[w_\pi](t, x, a_t, x, \varepsilon) \leq C|\pi|\bar{S}(\pi, \varepsilon)$$

where we set: $\bar{S}(\pi, \varepsilon) = (|\pi|^\frac{3}{2}(1 + \varepsilon^{-1}) + |\pi|^2\varepsilon^{-3})$. Fix $k \leq n - 1$. By Lemma 4.2.5 we then have for all $t \in [t_k, t_{k+1}]$, $x \in \mathbb{R}^d$:

$$w_\pi(t, x) - v^\pi(t, x, a_t, x, \varepsilon) = w_\pi(t, x) - T_\pi[w_\pi](t, x, a_t, x, \varepsilon) + (T_\pi[w_\pi] - T_\pi[v^\pi])(t, x, a_t, x, \varepsilon) \leq C|\pi|\bar{S}(\pi, \varepsilon) + e^{C|\pi|} \sup_{(x, a) \in \mathbb{R}^d \times A} (w_\pi - v^\pi)(t_{k+1}, x, a). \quad (4.2.37)$$

Recalling by its very definition that $v^\pi$ does not depend on $a \in A$ on the grid times of $\pi$, and denoting then $M_k := \sup_{x \in \mathbb{R}^d} (w_\pi - v^\pi)(t_k, x)$, we have by (4.2.37) the relation:

$$M_k \leq C|\pi|\bar{S}(\pi, \varepsilon) + e^{C|\pi|} M_{k+1}.$$ 

By induction, this yields:

$$\sup_{x \in \mathbb{R}^d} (w_\pi - v^\pi)(t_k, x) \leq C e^{C|\pi|} \frac{1}{e^{C|\pi|} - 1} |\pi|\bar{S}(\pi, \varepsilon) + e^{C|\pi|} \sup_{x \in \mathbb{R}^d} (w_\pi - v^\pi)(T, x) \leq C\bar{S}(\pi, \varepsilon) + C \sup_{x \in \mathbb{R}^d} (w_\pi - v)(T, x),$$

since $n|\pi|$ is bounded and $v(T, x) = v^\pi(T, x) = g(x)$. From (4.2.30), we then get:

$$\sup_{x \in \mathbb{R}^d} (v - v^\pi)(t_k, x) \leq C(\varepsilon^\frac{1}{3} + |\pi|\varepsilon^\frac{1}{2}(1 + \varepsilon^{-1}) + |\pi|\varepsilon^{-3}).$$

By minimizing the r.h.s of this estimate with respect to $\varepsilon$, this leads to the error bound when taking $\varepsilon = |\pi|\frac{1}{2\beta} \leq 1$:

$$\sup_{x \in \mathbb{R}^d} (v - v^\pi)(t_k, x) \leq C|\pi|\frac{1}{2\beta}.$$ 

Finally, by combining with the estimate in Lemma 4.2.4 which gives actually under (HBC)\textsuperscript{'}(i):

$$|\vartheta^\pi(t, x, a) - v^\pi(t_{k+1}, x)| \leq C|\pi|^{\frac{1}{2}}, \quad t \in [t_k, t_{k+1}), (x, a) \in \mathbb{R}^d \times A,$$

together with the $1/2$-Hölder property of $v$ in time (see (4.2.19)), we obtain:

$$\sup_{(t,x,a) \in [0,T] \times \mathbb{R}^d \times A} (v - \vartheta^\pi)(t, x, a) \leq C(|\pi|^{\frac{1}{2\beta}} + |\pi|^{\frac{1}{2}}) \leq C|\pi|^{\frac{1}{2\beta}}.$$ 

This ends the proof. \hfill $\square$
Chapter 5

Approximation scheme for jump-constrained BSDE and stochastic control problem

In this chapter, we study the discrete time approximation of the solution to the discretely constrained BSDE. We first deal with the approximation of the forward process. We then provide a discrete-time scheme for the discretely constrained BSDE. We finally show that the scheme converges to the constrained solution as soon as the discrete-time mesh goes to zero.

5.1 The forward regime switching process

In this section, we consider the discrete time approximation of the forward process \((X,I)\) on \([0,T]\). Recall that it is defined by

\[
\begin{aligned}
X_t &= X_0 + \int_0^t b(X_s, I_s)ds + \int_0^t \sigma(X_s, I_s)dW_s \\
I_t &= I_0 + \int_{(0,t]} A(a - I_s^-)\mu(ds, da),
\end{aligned}
\]

for all \(t \in [0,T]\).

In the sequel, we shall denote by \(C_1\) a generic positive constant which depends only on the Lipschitz constant of \(b\) and \(\sigma\), \(T\), \((X_0,I_0)\) and \(\lambda(A) < \infty\), and may vary from lines to lines. Under (HFC), we have the existence and uniqueness of a solution to (5.1.1), and in the sequel, we shall denote by \((X^{t,x,a}, I^{t,a})\) the solution to (5.1.1) starting from \((x,a)\) at time \(t\).

**Remark 5.1.3** We do not make any ellipticity assumption on \(\sigma\). In particular, some lines and columns of \(\sigma\) may be equal to zero, and so there is no loss of generality by considering that the dimension \(d\) of \(X\) and \(W\) are equal.

Denoting by \((T_n,\epsilon_n)\) the jump times and marks associated to \(\mu\), we observe that \(I\) is
explicitly written as:

\[ I_t = I_0 \mathbb{1}_{[0,T_1]}(t) + \sum_{n \geq 1} t_n \mathbb{1}_{[T_n,T_{n+1})}(t), \quad 0 \leq t \leq T, \]

where the jump times \((T_n)_n\) evolve according to a Poisson distribution of parameter \(\lambda := \int_A \lambda(da) < \infty\), and the i.i.d. marks \((\tau_n)_n\) follow a probability distribution \(\bar{\lambda}(da) := \lambda(da)/\lambda\). Assuming that one can simulate the probability distribution \(\bar{\lambda}\), we then see that the pure jump process \(I\) is perfectly simulated. Given a partition \(\pi = \{t_0 = 0 < \ldots < t_k < \ldots t_n = T\}\) of \([0, T]\), we shall use the natural Euler scheme \(\bar{X}^\pi\) for \(X\), defined by:

\[
\bar{X}_{0}^\pi = X_0 \\
\bar{X}_{t_{k+1}}^\pi = \bar{X}_{t_k}^\pi + b(\bar{X}_{t_k}^\pi, I_{t_k})(t_{k+1} - t_k) + \sigma(\bar{X}_{t_k}^\pi, I_{t_k})(W_{t_{k+1}} - W_{t_k}),
\]

for \(k = 0, \ldots, n - 1\). We denote as usual by \(|\pi| = \max_{k \leq n-1} (t_{k+1} - t_k)\) the modulus of \(\pi\), and assume that \(|\pi|\) is bounded by a constant independent of \(n\), which holds for instance when the grid is regular, i.e. \((t_{k+1} - t_k) = |\pi|\) for all \(k \leq n - 1\). We also define the continuous-time version of \(\bar{X}^\pi\) by setting:

\[
\bar{X}_t^\pi = \bar{X}_{t_k}^\pi + b(\bar{X}_{t_k}^\pi, I_{t_k})(t - t_k) + \sigma(\bar{X}_{t_k}^\pi, I_{t_k})(W_t - W_{t_k}), \quad t \in [t_k, t_{k+1}), \quad k < n.
\]

By standard arguments, see e.g. [27], one can obtain under (HFC) the \(L^2\)-error estimate for the above Euler scheme:

\[
\mathbb{E}\left[ \sup_{t \in [t_k, t_{k+1}]} |X_t - \bar{X}_{t_k}^\pi|^2 \right] \leq C_1 |\pi|, \quad k < n.
\]

For our purpose, we shall need a stronger result, and introduce the following error control for the Euler scheme:

\[
\mathcal{E}_k^\pi(X) := \mathbb{E}\left[ \mathrm{ess sup}_{a \in A} \mathbb{E}_{t_1,a} \ldots \mathrm{ess sup}_{a \in A} \mathbb{E}_{t_k,a} \left[ \sup_{t \in [t_k, t_{k+1}]} |X_t - \bar{X}_{t_k}^\pi|^2 \right] \ldots \right], \quad (5.1.2)
\]

where \(\mathbb{E}_{t_k,a}[\cdot]\) denotes the conditional expectation \(\mathbb{E}[\cdot | \mathcal{F}_{t_k}, I_{t_k} = a]\). We also denote by \(\mathbb{E}_{t_k}[\cdot]\) the conditional expectation \(\mathbb{E}[\cdot | \mathcal{F}_{t_k}]\). Since \(I_{t_k}\) is \(\mathcal{F}_{t_k}\)-measurable, and by the law of iterated conditional expectations, we notice that

\[
\mathbb{E}\left[ \sup_{t \in [t_k, t_{k+1}]} |X_t - \bar{X}_{t_k}^\pi|^2 \right] \leq \mathcal{E}_k^\pi(X), \quad k < n.
\]

**Lemma 5.1.6** Under (HFC), we have

\[
\max_{k < n} \mathcal{E}_k^\pi(X) \leq C_1 |\pi|.
\]

**Proof.** From the definition of the Euler scheme, and under the growth linear condition in (HFC), we easily see that

\[
\mathbb{E}_{t_k}\left[ |\bar{X}_{t_{k+1}}^\pi|^2 \right] \leq C_1 (1 + |\bar{X}_{t_k}^\pi|^2), \quad k < n. \quad (5.1.3)
\]
From the definition of the continuous-time Euler scheme, and by Burkholder-Davis-Gundy
inequality, it is also clear that
\[ \mathbb{E}_t \left[ \sup_{t \in [t_k, t_{k+1}]} |\bar{X}_t^\pi - \bar{X}_t^\pi|^2 \right] \leq C_1 \left( 1 + |\bar{X}_t^\pi|^2 \right)|\pi|, \quad k < n. \quad (5.1.4) \]

We also have the standard estimate for the pure jump process \( I \) (recall that \( A \) is assumed
to be compact and \( \lambda(A) < \infty \)):
\[ \mathbb{E}_t \left[ \sup_{t \in [t_k, t_{k+1}]} |I_s - I_{t_k}|^2 \right] \leq C_1 |\pi|. \quad (5.1.5) \]

Let us denote by \( \Delta X_t = X_t - \bar{X}_t^\pi \), and apply Itô’s formula to \(|\Delta X_t|^2\) so that for all \( t \in [t_k, t_{k+1}]\):
\[
|\Delta X_t|^2 = |\Delta X_{t_k}|^2 + \int_{t_k}^t 2(b(X_s, I_s) - b(\bar{X}_{t_k}^\pi, I_{t_k}))(\Delta X_s + |\sigma(X_s, I_s) - \sigma(\bar{X}_{t_k}^\pi, I_{t_k})|^2)ds
+ 2 \int_{t_k}^t (\Delta X_s)'(\sigma(X_s, I_s) - \sigma(\bar{X}_{t_k}^\pi, I_{t_k}))dW_s
\leq |\Delta X_{t_k}|^2 + C_1 \int_{t_k}^t |\Delta X_s|^2 + |\bar{X}_s^\pi - \bar{X}_{t_k}^\pi|^2 + |I_s - I_{t_k}|^2ds
+ 2 \int_{t_k}^t (\Delta X_s)'(\sigma(X_s, I_s) - \sigma(\bar{X}_{t_k}^\pi, I_{t_k}))dW_s,
\]
from the Lipschitz condition on \( b, \sigma \) in (HFC). By taking conditional expectation in the
above inequality, we then get:
\[
\mathbb{E}_t[|\Delta X_t|^2] \leq |\Delta X_{t_k}|^2 + C_1 \int_{t_k}^t \mathbb{E}_t[|\Delta X_s|^2 + |\bar{X}_s^\pi - \bar{X}_{t_k}^\pi|^2 + |I_s - I_{t_k}|^2]ds
\leq |\Delta X_{t_k}|^2 + C_1 (1 + |\bar{X}_{t_k}^\pi|^2)|\pi|^2 + C_1 \int_{t_k}^t \mathbb{E}_t[|\Delta X_s|^2]ds, \quad t \in [t_k, t_{k+1}],
\]
by (5.1.4)-(5.1.5). From Gronwall’s lemma, we thus deduce that
\[ \mathbb{E}_t[|\Delta X_{t+1,k}|^2] \leq e^{C_1|\pi|}|\Delta X_{t,k}|^2 + C_1 (1 + |\bar{X}_{t_k}^\pi|^2)|\pi|^2, \quad k < n. \quad (5.1.6) \]

Since the right hand side of (5.1.6) does not depend on \( I_{t_k} \), this shows that
\[ \text{ess sup}_{a \in A} \mathbb{E}_{t_k,a}[|\Delta X_{t+1,k}|^2] \leq e^{C_1|\pi|}|\Delta X_{t,k}|^2 + C_1 (1 + |\bar{X}_{t_k}^\pi|^2)|\pi|^2. \]

By taking conditional expectation w.r.t. \( F_{t_k-1} \) in the above inequality, using again estimate
(5.1.6) together with (5.1.3) at step \( k - 1 \), and iterating this backward procedure until the
initial time \( t_0 = 0 \), we obtain:
\[
\mathbb{E} \left[ \text{ess sup}_{a \in A} \mathbb{E}_{t_1,a} \left[ \ldots \text{ess sup}_{a \in A} \mathbb{E}_{t_k,a} \left[ |\Delta X_{t+1,k+1}|^2 \right] \ldots \right] \right]
\leq e^{C_1|\pi|}|\Delta X_0|^2 + C_1 (1 + |X_0|^2)|\pi|^2 \frac{e^{C_1|\pi|} - 1}{e^{C_1|\pi|} - 1}
\leq C_1 |\pi|, \quad (5.1.7)
\]
since $\Delta X_0 = 0$ and $n|\pi|$ is bounded.

Moreover, the process $X$ satisfies the standard conditional estimate similarly as for the Euler scheme:

$$\mathbb{E}_{t_k}\left[|X_{t_{k+1}}|^2\right] \leq C_1(1 + |X_{t_k}|^2),$$

$$\mathbb{E}_{t_k}\left[\sup_{t \in [t_k, t_{k+1}]}|X_t - X_{t_k}|^2\right] \leq C_1(1 + |X_{t_k}|^2)|\pi|, \quad k < n,$$

from which we deduce by backward induction on the conditional expectations:

$$\mathbb{E}_{t_k}\left[\sup_{a \in A} \mathbb{E}_{t_1,a}\left[\ldots \sup_{a \in A} \mathbb{E}_{t_{k+1},a}\left[\sup_{t \in [t_k, t_{k+1}]}|X_t - X_{t_k}|^2\right] \ldots \right]\right] \leq C_1|\pi|. \quad (5.1.8)$$

Finally, by writing that $\sup_{t \in [t_k, t_{k+1}]}|X_t - X_{t_k}|^2 \leq 2 \sup_{t \in [t_k, t_{k+1}]}|X_t - X_{t_k}|^2 + 2\Delta X_{t_k}$, taking successive condition expectations w.r.t to $F_{t_\ell}$ and essential supremum over $I_{t_\ell} = a$, for $\ell$ going recursively from $k$ to 0, we get:

$$\mathbb{E}_{t_k}\left[\sup_{t \in [t_k, t_{k+1}]}|X_t - X_{t_k}|^2\right] \leq 2\mathbb{E}_{t_k}\left[\sup_{a \in A} \mathbb{E}_{t_1,a}\left[\ldots \sup_{a \in A} \mathbb{E}_{t_{k+1},a}\left[\sup_{t \in [t_k, t_{k+1}]}|X_t - X_{t_k}|^2\right] \ldots \right]\right]$$

$$+ 2\mathbb{E}_{t_k}\left[\sup_{a \in A} \mathbb{E}_{t_1,a}\left[\ldots \sup_{a \in A} \mathbb{E}_{t_{k-1},a}\left[|\Delta X_{t_k}|^2\right] \ldots \right]\right] \leq C_1|\pi|,$$

by $(5.1.7)$–$(5.1.8)$, which ends the proof. \qed

### 5.2 BSDE scheme

We consider the discrete time approximation of the discretely jump-constrained BSDE in the case where $f(x,a,y)$ does not depend on $z$, and define the scheme $(\tilde{Y}^\pi, \tilde{\gamma}^\pi, \tilde{Z}^\pi)$ by induction on the grid $\pi = \{t_0 = 0 < \ldots < t_k < \ldots < t_n = T\}$ by:

$$\begin{cases}
\tilde{Y}^\pi_T = \tilde{Y}^\pi_T = g(X^\pi_T) \\
\tilde{Y}^\pi_{t_k} = \mathbb{E}_{t_k}\left[\tilde{Y}^\pi_{t_{k+1}}\right] + f(X^\pi_{t_k}, I_{t_k}, \tilde{\gamma}^\pi_{t_k})\Delta t_k \\
\tilde{Y}^\pi_{t_k} = \sup_{a \in A} \mathbb{E}_{t_k,a}\left[\tilde{Y}^\pi_{t_k}\right], \quad k = 0, \ldots, n-1,
\end{cases} \quad (5.2.9)$$

where $\Delta t_k = t_{k+1} - t_k$, $\mathbb{E}_{t_k}[\cdot]$ stands for $\mathbb{E}[\cdot|\mathcal{F}_{t_k}]$, and $\mathbb{E}_{t_k,a}[\cdot]$ for $\mathbb{E}[\cdot|\mathcal{F}_{t_k}, I_{t_k} = a]$.

By induction argument, we easily see that $\tilde{Y}^\pi_{t_k}$ is a deterministic function of $(X^\pi_{t_k}, I_{t_k})$, while $\tilde{\gamma}^\pi_{t_k}$ is a deterministic function of $X^\pi_{t_k}$, for $k = 0, \ldots, n$, and by the Markov property of the process $(X^\pi, I)$, the conditional expectations in (5.2.9) can be replaced by the corresponding regressions:

$$\mathbb{E}_{t_k}\left[\tilde{Y}^\pi_{t_{k+1}}\right] = \mathbb{E}\left[\tilde{Y}^\pi_{t_{k+1}}|X^\pi_{t_k}, I_{t_k}\right] \quad \text{and} \quad \mathbb{E}_{t_k,a}\left[\tilde{Y}^\pi_{t_k}\right] = \mathbb{E}\left[\tilde{Y}^\pi_{t_k}|X^\pi_{t_k}, I_{t_k} = a\right].$$

We then have:

$$\tilde{Y}^\pi_{t_k} = \tilde{\gamma}^\pi_{t_k}(X^\pi_{t_k}, I_{t_k}), \quad Y^\pi_{t_k} = \tilde{\gamma}^\pi_{t_k}(X^\pi_{t_k}),$$

62
for some sequence of functions $(\bar{\varphi}_k^\pi)_k$ and $(\bar{v}_k^\pi)_k$ defined respectively on $\mathbb{R}^d \times A$ and $\mathbb{R}^d$ by backward induction:

$$
\begin{aligned}
\bar{v}_n^\pi(x,a) &= \bar{\varphi}_n^\pi(x) = g(x) \\
\bar{\varphi}_k^\pi(x,a) &= \mathbb{E}\left[\bar{v}_{k+1}^\pi(X_{t_{k+1}}, I_{t_{k+1}}) \mid (X_{t_k}, I_k) = (x,a)\right] + f(x,a, \bar{\varphi}_k^\pi(x,a)) \Delta t_k \\
\bar{v}_k^\pi(x) &= \sup_{a \in A} \bar{\varphi}_k^\pi(x,a), \quad k = 0, \ldots, n - 1.
\end{aligned}
$$

The above convergence rate $|\pi|^{\frac{1}{2}}$ in the $L^2$-norm for the discretization of the discretely jump-constrained BSDE is the same as for standard BSDE, see [9], [44]. By combining with the convergence result in Section 4, we finally obtain an estimate on the error due to the discrete time approximation of the minimal solution $Y$ to the BSDE with nonpositive jumps. We split the error between the positive and negative parts:

$$
\text{Err}_+^\pi(Y) := \max_{k \leq n-1} \left(\mathbb{E}\left[(Y_{t_k} - \bar{Y}_{t_k}^\pi)_+^2\right] + \sup_{t \in [t_k, t_{k+1}]} \mathbb{E}\left[(Y_t - \bar{Y}_{t_{k+1}}^\pi)_+^2\right]\right)^{\frac{1}{2}}
$$

$$
\text{Err}_-^\pi(Y) := \max_{k \leq n-1} \left(\mathbb{E}\left[(Y_{t_k} - \bar{Y}_{t_k}^\pi)_-^2\right] + \sup_{t \in [t_k, t_{k+1}]} \mathbb{E}\left[(Y_t - \bar{Y}_{t_{k+1}}^\pi)_-^2\right]\right)^{\frac{1}{2}}.
$$

**Corollary 5.3.4** We have:

$$
\text{Err}^\pi(Y) \leq C|\pi|^{\frac{1}{2}}.
$$
Moreover, under \((HFC^\prime)\) and \((HBC^\prime)\),

\[
\text{Err}_n^\pi(Y) \leq C|\pi|^\frac{1}{\hat{b}},
\]

and when \(f(x,a)\) does not depend on \(y\):

\[
\text{Err}_n^\pi(Y) \leq C|\pi|^\frac{1}{\hat{b}}.
\]

**Proof.** Recall from Proposition 4.2.5 that \(Y_t^\pi \leq Y_t \leq t \leq T\). Then, we have:

\[
(Y_{tk} - Y_{tk}^\pi)_- \leq \left|Y_{tk}^\pi - Y_{tk}\right|, \quad (Y_t - Y_{tk}^\pi)_- \leq \left|Y_t^\pi - Y_{tk}^\pi\right|,
\]

and \(Y_{tk} - Y_{tk}^\pi \leq \left|Y_t - Y_{tk}^\pi\right|\), for all \(k \leq n - 1\), and \(t \in [0,T]\). The error bound on \(\text{Err}_n^\pi(Y)\) follows then from the estimation in Theorem 5.3.2. The error bound on \(\text{Err}_n^\pi(Y)\) follows from Corollary 4.2.3 and Theorem 5.3.2.

**Remark 5.3.4** In the particular case where \(f\) depends only on \((x,a)\), our discrete time approximation scheme is a probabilistic scheme for the fully nonlinear HJB equation associated to the stochastic control problem (4.2.27). As in [29], [5] or [17], we have non asymptotic bounds on the rate of convergence. For instance, in [17], the authors obtained a convergence rate \(|\pi|^\frac{1}{2}\) on one side and \(|\pi|^\frac{1}{10}\) on the other side, while we improve the rate to \(|\pi|^\frac{1}{2}\) for one side, and \(|\pi|^\frac{1}{4}\) on the other side. This induces a global error \(\text{Err}^\pi(Y) = \text{Err}^\pi_n(Y) + \text{Err}^\pi_n(Y)\) of order \(|\pi|^\frac{1}{4}\), which is derived without any non degeneracy condition on the controlled diffusion coefficient.

**Proof of Theorem 5.3.2**

Let us introduce the continuous time version of (5.2.9). By the martingale representation theorem, there exists \(\tilde{Z}^\pi \in \mathcal{L}^2(W)\) and \(\tilde{U}^\pi \in \mathcal{L}^2(\tilde{\mu})\) such that

\[
Y^\pi_{tk+1} = \mathbb{E}_t\left[Y^\pi_{tk+1}\right] + \int_{tk}^{tk+1} \tilde{Z}^\pi_t dW_t + \int_{tk}^{tk+1} \int_A \tilde{U}^\pi_t(a) \tilde{\mu}(dt, da), \quad k < n,
\]

and we can then define the continuous-time processes \(\bar{Y}^\pi\) and \(\tilde{Y}^\pi\) by:

\[
\bar{Y}^\pi_t = \bar{Y}^\pi_{tk+1} + (tk+1 - t)f(X^\pi_{tk}, I_{tk}, \bar{Y}^\pi_{tk}) - \int_t^{tk+1} \tilde{Z}^\pi_t dW_t - \int_t^{tk+1} \int_A \tilde{U}^\pi_t(a) \tilde{\mu}(dt, da), \quad t \in [tk, tk+1),
\]

\[
\tilde{Y}^\pi_t = \tilde{Y}^\pi_{tk+1} + (tk+1 - t)f(X^\pi_{tk}, I_{tk}, \bar{Y}^\pi_{tk}) - \int_t^{tk+1} \tilde{Z}^\pi_t dW_t - \int_t^{tk+1} \int_A \tilde{U}^\pi_t(a) \tilde{\mu}(dt, da), \quad t \in (tk, tk+1],
\]

for \(k = 0, \ldots, n - 1\). Denote by \(\delta Y_t^\pi = Y_t^\pi - Y_t\), \(\delta Y_t^\pi = Y_t^\pi - \bar{Y}_t^\pi\), \(\delta Z^\pi_t = Z^\pi_t - \tilde{Z}^\pi_t\), \(\delta U^\pi_t = U^\pi_t - \tilde{U}^\pi_t\) and \(\delta f_t = f(X_t, I_t, Y_t^\pi) - f(X^\pi_{tk}, I_{tk}, \bar{Y}^\pi_{tk})\) for \(t \in [tk, tk+1)\). Recalling (4.1.2) and (5.3.11), we have by Itô’s formula:

\[
\Delta_t := \mathbb{E}_t \left[|\delta Y_t^\pi|^2 + \int_t^{tk+1} |\delta Z^\pi_s|^2 ds + \int_t^{tk+1} \int_A |\delta U^\pi_s(a)|^2 \lambda(da) ds\right]
\]

\[
= \mathbb{E}_t \left[|\delta Y_{tk+1}^\pi|^2\right] + \mathbb{E}_t \left[\int_t^{tk+1} 2\delta Y^\pi_s \delta f_s ds\right]
\]

64
for all $t \in [t_k, t_{k+1})$. By the Lipschitz continuity of $f$ in (H2) and Young inequality, we then have:

$$
\Delta_t \leq E_{t_k}[|\delta Y_{t_k+1}^\pi|^2] + E_{t_k} \left[ \int_t^{t_{k+1}} \eta |\delta Y_s^\pi|^2 ds + \frac{C}{\eta} |\delta Y_{t_k}^\pi|^2 \right] + \frac{C}{\eta} E_{t_k} \left[ \int_t^{t_{k+1}} (|X_s - \bar{X}_{t_k}^\pi|^2 + |I_s - I_{t_k}|^2 + |Y_s^\pi - \bar{Y}_{t_k}^\pi|^2) ds \right].
$$

From Gronwall’s lemma, and by taking $\eta$ large enough, this yields for all $k \leq n - 1$:

$$
E_{t_k}[|\delta Y_{t_{k+1}}^\pi|^2] \leq e^{C|\pi|} E_{t_k}[|\delta Y_{t_{k+1}}^\pi|^2] + C B_k
$$

where

$$
B_k = E_{t_k} \left[ \int_t^{t_{k+1}} (|X_s - \bar{X}_{t_k}^\pi|^2 + |I_s - I_{t_k}|^2 + |Y_s^\pi - \bar{Y}_{t_k}^\pi|^2) ds \right] \leq C|\pi| \left( E_{t_k} \left[ \sup_{s \geq [t_k, t_{k+1}]} |X_s - \bar{X}_{t_k}^\pi|^2 \right] + |\pi| (1 + |X_{t_k}|) \right),
$$

by (5.1.5) and Proposition 4.1.3. Now, by definition of $Y_{t_{k+1}}^\pi$ and $\bar{Y}_{t_{k+1}}^\pi$, we have

$$
|\delta Y_{t_{k+1}}^\pi|^2 \leq \esssup_{a \in A} E_{t_{k+1},a}[|\delta Y_{t_{k+1}}^\pi|^2].
$$

By plugging (5.3.14), (5.3.15) into (5.3.13), taking conditional expectation with respect to $I_{t_k} = a$, and taking essential supremum over $a$, we obtain:

$$
\esssup_{a \in A} E_{t_k,a}[|\delta Y_{t_k}^\pi|^2] \leq e^{C|\pi|} \esssup_{a \in A} E_{t_k,a}[\esssup_{a \in A} E_{t_{k+1},a}[|\delta Y_{t_{k+1}}^\pi|^2]] + C|\pi| \left( \esssup_{a \in A} E_{t_k,a}[\sup_{s \geq [t_k, t_{k+1}]} |X_s - \bar{X}_{t_k}^\pi|^2] + |\pi| (1 + |X_{t_k}|) \right).
$$

By taking conditional expectation with respect to $F_{t_{k-1}}$, and $I_{t_{k-1}} = a$, taking essential supremum over $a$ in the above inequality, and iterating this backward procedure until time $t_0 = 0$, we obtain:

$$
E_k^\pi(\mathcal{Y}) \leq e^{C|\pi|} E_{k+1}^\pi(\mathcal{Y}) + C|\pi| \left( E_{k+1}^\pi(X) + |\pi| (1 + E[|X_{t_k}|]) \right) \leq e^{C|\pi|} E_{k+1}^\pi(\mathcal{Y}) + C|\pi|^2, \quad k \leq n - 1,
$$

where we recall the auxiliary error control $E_k^\pi(X)$ on $X$ in (5.1.2) and its estimate in Lemma 5.1.6, and set:

$$
E_k^\pi(\mathcal{Y}) := E \left[ \esssup_{a \in A} E_{t_1,a}[\ldots \esssup_{a \in A} E_{t_k,a}[|\delta Y_{t_k}^\pi|^2] \ldots] \right].
$$

By a direct induction on (5.3.16), and recalling that $n|\pi|$ is bounded, we get

$$
E_k^\pi(\mathcal{Y}) \leq C(E_k^\pi(\mathcal{Y}) + |\pi|) \leq C(E_k^\pi(X) + |\pi|) \leq C|\pi|,
$$

65
since \(g\) is Lipschitz, and using again the estimate in Lemma \ref{5.1.6} Observing that \(\mathbb{E}[|\delta Y_{t_k}^\pi|^2] \leq \mathcal{E}_k^\pi(Y)\), we get the estimate:

\[
\max_{k \leq n} \mathbb{E}[|Y_{t_k}^\pi - \tilde{Y}_{t_k}^\pi|^2] + \mathbb{E}[|Y_{t_{k+1}}^\pi - \tilde{Y}_{t_{k+1}}^\pi|^2] \leq C|\pi|.
\]

Moreover, by Proposition \ref{4.1.3} we have

\[
\sup_{t \in [t_k, t_{k+1}]} \mathbb{E}[|Y_t^\pi - Y_{t_k}^\pi|^2] + \sup_{t \in (t_k, t_{k+1}]} \mathbb{E}[|Y_t^\pi - Y_{t_{k+1}}^\pi|^2] \leq C(1 + \mathbb{E}[|X_{t_k}|]|\pi|) \leq C(1 + |X_0|)|\pi|.
\]

This implies finally that:

\[
\sup_{s \in [t_k, t_{k+1}]} \mathbb{E}[|Y_s^\pi - \tilde{Y}_{t_k}^\pi|^2] \leq 2 \sup_{s \in [t_k, t_{k+1}]} \mathbb{E}[|Y_s^\pi - Y_{t_k}^\pi|^2] + 2\mathbb{E}[|Y_{t_k}^\pi - \tilde{Y}_{t_k}^\pi|^2] \leq C|\pi|,
\]

as well as

\[
\sup_{s \in [t_k, t_{k+1}]} \mathbb{E}[|Y_s^\pi - Y_{t_{k+1}}^\pi|^2] \leq 2 \sup_{s \in [t_k, t_{k+1}]} \mathbb{E}[|Y_s^\pi - Y_{t_k}^\pi|^2] + 2\mathbb{E}[|Y_{t_k}^\pi - \tilde{Y}_{t_{k+1}}^\pi|^2] \leq C|\pi|.
\]

\[\square\]

### 5.4 Approximate optimal control

We now consider the special case where \(f(x, a)\) does not depend on \(y\), so that the discrete time scheme \((1.0.15)\) is an approximation for the value function of the stochastic control problem:

\[
V_0 := \sup_{\alpha \in A} J(\alpha) = Y_0, \quad (5.4.17)
\]

\[
J(\alpha) = \mathbb{E}\left[\int_0^T f(X_t^\alpha, \alpha_t)dt + g(X_T^\alpha)\right],
\]

where \(A\) is the set of \(\mathcal{G}\)-adapted control processes \(\alpha\) valued in \(A\), and \(X^\alpha\) is the controlled diffusion in \(\mathbb{R}^d\):

\[
X_t^\alpha = X_0 + \int_0^t b(X_s^\alpha, \alpha_s)ds + \int_0^t \sigma(X_s^\alpha, \alpha_s)dW_s, \quad 0 \leq t \leq T.
\]

(Here \(\mathcal{G} = (\mathcal{G}_t)_{0 \leq t \leq T}\) denotes some filtration under which \(W\) is a standard Brownian motion). Let us now define the discrete time version of \((5.4.17)\). We introduce the set \(A^\pi\) of discrete time processes \(\alpha = (\alpha_t)_{k}\) with \(\alpha_t \mathcal{G}_{t_k}\)-measurable, and valued in \(A\). For each \(\alpha \in A^\pi\), we consider the controlled discrete time process \((X_{t_k}^\pi,\alpha)_{k}\) of Euler type defined by:

\[
X_{t_k}^\pi,\alpha = X_0 + \sum_{j=0}^{k-1} b(X_{t_j}^\pi,\alpha_t)\Delta t_j + \sum_{j=0}^{k-1} \sigma(X_{t_j}^\pi,\alpha_t)\Delta W_t, \quad k \leq n,
\]

where \(\Delta t_j = t_{j+1} - t_j\) and \(\Delta W_t = W_{t+1} - W_t\) for \(0 \leq t \leq T\).
Let us define the process \( \hat{a}^{\sup} \) for the stochastic control problem (5.4.17) by taking
\[
\hat{a}^{\sup} = \sup_{\alpha} \{ f(X_{t_k}^{\pi,\alpha}, \alpha t_k) \Delta t_k + g(X_{t_k}^{\pi,\alpha}) \}.
\]

Given any \( \alpha \in \mathcal{A}^\pi \), we define its continuous time piecewise-constant interpolation \( \alpha \in \mathcal{A} \) by setting: \( \alpha_t = \alpha t_k \), for \( t \in [t_k, t_{k+1}) \) (by misuse of notation, we keep the same notation \( \alpha \) for the discrete time and continuous time interpolation). By standard arguments similar to those for Euler scheme of SDE, there exists some positive constant \( C \) such that for all \( \alpha \in \mathcal{A}^\pi \), \( k \leq n-1 \):
\[
\mathbb{E} \left[ \sup_{t \in [t_k, t_{k+1}]} \left| X_t^\alpha - X_{t_k}^{\pi,\alpha} \right|^2 \right] \leq C |\pi|, 
\]
from which we easily deduce by Lipschitz property of \( f \) and \( g \):
\[
|J(\alpha) - J^{\pi}(\alpha)| \leq C |\pi|^{\frac{1}{2}}, \quad \forall \alpha \in \mathcal{A}^\pi. \quad (5.4.18)
\]

Let us now consider at each time step \( k \leq n-1 \), the function \( \hat{a}_k(x) \) which attains the supremum over \( \alpha \in \mathcal{A} \) in the scheme (5.2.10), so that:
\[
\overline{v}_k^\pi(x) = \overline{\hat{v}}_k(x, \hat{a}_k(x)), \quad k = 0, \ldots, n-1.
\]

Let us define the process \( \hat{X}_k^\pi \) by: \( \hat{X}_0^\pi = X_0 \),
\[
\hat{X}_{t_k+1}^\pi = \hat{X}_{t_k}^\pi + b(\hat{X}_{t_k}^\pi, \hat{a}_k(\hat{X}_{t_k}^\pi)) \Delta t_k + \sigma(\hat{X}_{t_k}^\pi, \hat{a}_k(\hat{X}_{t_k}^\pi)) \Delta W_{t_k}, \quad k \leq n-1,
\]
and notice that \( \hat{X}_k^\pi = X_k^{\pi,\hat{a}} \), where \( \hat{a} \in \mathcal{A}^\pi \) is a feedback control defined by:
\[
\hat{a}_k = \hat{a}_k(\hat{X}_{t_k}^\pi) = \hat{a}_k(X_{t_k}^{\pi,\hat{a}}), \quad k = 0, \ldots, n.
\]

Next, we observe that the conditional law of \( \hat{X}_{t_{k+1}}^\pi \) given \( \hat{X}_{t_k}^\pi = x, I_{t_k} = \hat{a}_k(\hat{X}_{t_k}^\pi) = \hat{a}_k(x) \) is the same than the conditional law of \( X_{t_{k+1}}^{\pi,\hat{a}} \) given \( X_{t_k}^{\pi,\hat{a}} = x, \) for \( k \leq n-1 \), and thus the induction step in the scheme (5.2.9) or (5.2.10) reads as:
\[
\overline{v}_k^\pi(X_{t_k}^{\pi,\hat{a}}) = \mathbb{E} \left[ \overline{v}_{k+1}^\pi(X_{t_{k+1}}^{\pi,\hat{a}}) | X_{t_k}^{\pi,\hat{a}} \right] + f(X_{t_k}^{\pi,\hat{a}}, \hat{a}_k(\hat{X}_{t_k}^\pi)) \Delta t_k, \quad k \leq n-1.
\]

By induction, and law of iterated conditional expectations, we then get:
\[
\hat{Y}_0^\pi = \overline{v}_0^\pi(X_0) = J^\pi(\hat{a}). \quad (5.4.19)
\]

Consider the continuous time piecewise-constant interpolation \( \hat{\alpha} \in \mathcal{A} \) defined by: \( \hat{\alpha}_t = \hat{\alpha}_k \), for \( t \in [t_k, t_{k+1}) \). By (5.4.17), (5.4.18), (5.4.19), and Corollary 5.3.4, we finally obtain:
\[
0 \leq V_0 - J(\hat{\alpha}) = Y_0 - \hat{Y}_0^\pi + J^\pi(\hat{\alpha}) - J(\hat{\alpha}) \leq C |\pi|^{\frac{1}{2}} + C |\pi|^{\frac{1}{2}} \leq C |\pi|^{\frac{1}{2}},
\]
for \( |\pi| \leq 1 \). In other words, for any small \( \varepsilon > 0 \), we obtain an \( \varepsilon \)-approximate optimal control \( \hat{\alpha} \) for the stochastic control problem (5.4.17) by taking \( |\pi| \) of order \( \varepsilon^6 \).


