Feynman-Kac representation for Hamilton-Jacobi-Bellman IPDE*

Idris KHARROUBI[1)], Huyên PHAM[2)]

December 11, 2012
revised version: November 27, 2013

[1)] CEREMADE, CNRS, UMR 7534
Université Paris Dauphine
kharroubi at ceremade.dauphine.fr

[2)] Laboratoire de Probabilités et
Modèles Aléatoires, CNRS, UMR 7599
Université Paris 7 Diderot,
and CREST-ENSAE
pham at math.univ-paris-diderot.fr

Abstract

We aim to provide a Feynman-Kac type representation for Hamilton-Jacobi-Bellman equation, in terms of Forward Backward Stochastic Differential Equation (FBSDE) with a simulatable forward process. For this purpose, we introduce a class of BSDE where the jumps component of the solution is subject to a partial nonpositive constraint. Existence and approximation of a unique minimal solution is proved by a penalization method under mild assumptions. We then show how minimal solution to this BSDE class provides a new probabilistic representation for nonlinear integro-partial differential equations (IPDEs) of Hamilton-Jacobi-Bellman (HJB) type, when considering a regime switching forward SDE in a Markovian framework, and importantly we do not make any ellipticity condition. Moreover, we state a dual formula of this BSDE minimal solution involving equivalent change of probability measures. This gives in particular an original representation for value functions of stochastic control problems including controlled diffusion coefficient.

Key words: BSDE with jumps, constrained BSDE, regime-switching jump-diffusion, Hamilton-Jacobi-Bellman equation, nonlinear Integral PDE, viscosity solutions, inf-convolution, semiconcave approximation.

MSC Classification: 60H10, 60H30, 35K55, 93E20.

---

*The authors would like to thank Pierre Cardaliaguet for useful discussions.

1 Introduction

The classical Feynman-Kac theorem states that the solution to the linear parabolic partial differential equation (PDE) of second order:

$$\begin{aligned}\frac{\partial v}{\partial t} + b(x).D_x v + \frac{1}{2}\text{tr}(\sigma\sigma^\intercal(x)D_x^2 v) + f(x) &= 0, \qquad (t,x) \in [0,T) \times \mathbb{R}^d, \\ v(T,x) &= g(x), \qquad x \in \mathbb{R}^d,\end{aligned}$$

may be probabilistically represented under some general conditions as (see e.g. [11]):

$$v(t,x) = \mathbb{E}\Big[\int_t^T f(X_s^{t,x})ds + g(X_T^{t,x})\Big], \tag{1.1}$$

where $X^{t,x}$ is the solution to the stochastic differential equation (SDE) driven by a d-dimensional Brownian motion W on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$:

$$dX_s = b(X_s)ds + \sigma(X_s)dW_s,$$

starting from $x \in \mathbb{R}^d$ at $t \in [0,T]$. By considering the process $Y_t = v(t, X_t)$, and from Itô's formula (when v is smooth) or in general from martingale representation theorem w.r.t. the Brownian motion W, the Feynman-Kac formula (1.1) is formulated equivalently in terms of (linear) Backward Stochastic Equation:

$$Y_t = g(X_T) + \int_t^T f(X_s)ds - \int_t^T Z_s dW_s, \quad t \le T,$$

with Z an adapted process, which is identified to: $Z_t = \sigma^\intercal(X_t)D_x v(t, X_t)$ when v is smooth.

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

$$\begin{aligned}\frac{\partial v}{\partial t} + \sup_{a\in A}\big[b(x,a).D_x v + \frac{1}{2}\text{tr}(\sigma\sigma^\intercal(x,a)D_x^2 v) + f(x,a)\big] &= 0, \text{ on } [0,T) \times \mathbb{R}^d, \quad (1.2) \\ v(T,x) &= g(x), \qquad x \in \mathbb{R}^d,\end{aligned}$$

where A is a subset of $\mathbb{R}^q$. It is well-known (see e.g. [24]) that such nonlinear PDE is the dynamic programming equation associated to the stochastic control problem with value function defined by:

$$v(t,x) := \sup_\alpha \mathbb{E}\Big[\int_t^T f(X_s^{t,x,\alpha}, \alpha_s)ds + g(X_T^{t,x,\alpha})\Big], \tag{1.3}$$

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 α valued in A.

Our main goal is to provide a probabilistic representation for the nonlinear HJB equation using Backward Stochastic Differential Equation (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 $\sigma\sigma^{\intercal}(x)$ is of full rank. Denoting by $\theta(x,a) = \sigma^{\intercal}(x)(\sigma\sigma^{\intercal}(x))^{-1}b(x,a)$ a solution to $\sigma(x)\theta(x,a) = b(x,a)$, we notice that the HJB equation reduces into a semi-linear PDE:

$$\frac{\partial v}{\partial t} + \frac{1}{2}\mathrm{tr}(\sigma\sigma^{\intercal}(x)D_x^2 v) + F(x, \sigma^{\intercal}D_x v) \;\; = \;\; 0, \tag{1.4}$$

where $F(x,z) = \sup_{a\in A}[f(x,a) + \theta(x,a).z]$ is the θ-Fenchel-Legendre transform of f. In this case, we know from the seminal works by Pardoux and Peng [19], [20], that the (viscosity) solution v to the semi-linear PDE (1.4) is connected to the BSDE:

$$Y_t \;\; = \;\; g(X_T^0) + \int_t^T F(X_s^0, Z_s)ds - \int_t^T Z_s dW_s, \quad t \leq T, \tag{1.5}$$

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

$$dX_s^0 \;\; = \;\; \sigma(X_s^0)dW_s.$$

This probabilistic representation leads to a probabilistic numerical scheme for the resolution to (1.4) by discretization and simulation of the BSDE (1.5), see [4]. Alternatively, when the function $F(x,z)$ is of polynomial type on z, the semi-linear PDE (1.4) can be numerically solved by a forward Monte-Carlo scheme relying on marked branching diffusion, as recently pointed out in [13]. Moreover, as showed in [9], the solution to the BSDE (1.5) admits a dual representation in terms of equivalent change of probability measures as:

$$Y_t \;\; = \;\; \operatorname*{ess\,sup}_{\alpha} \mathbb{E}^{\mathbb{P}^\alpha}\Big[\int_t^T f(X_s^0, \alpha_s)ds + g(X_T^0)\big|\mathcal{F}_t\Big], \tag{1.6}$$

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

$$dX_s^0 \;\; = \;\; b(X_s^0, \alpha_s)ds + \sigma(X_s^0)dW_s^\alpha,$$

with W^α a $\mathbb{P}^\alpha$-Brownian motion by Girsanov's theorem. In other words, the process X^0 has the same dynamics under $\mathbb{P}^\alpha$ than the controlled process X^α under $\mathbb{P}$, and the representation (1.6) can be viewed as a weak formulation (see [8]) of the stochastic control problem (1.3) 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:

$$dX_s^\alpha \;\; = \;\; \alpha_s dW_s,$$

where the control process α 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:

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

The associated HJB equation takes the form:

$$\frac{\partial v}{\partial t} + G(D_x^2 v) = 0, \quad (t,x) \in [0,T) \times \mathbb{R}, \quad v(T,x) = g(x), \quad x \in \mathbb{R}, \tag{1.7}$$

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 (1.7) is represented in terms of the so-called G-Brownian motion B, and G-expectation $\mathbb{E}_G$, concepts introduced in [22]:

$$v(t,x) = \mathbb{E}_G\big[g(x + B_{T-t})\big].$$

Moreover, G-expectation is closely related to second order BSDE studied in [27], namely the process $Y_t = v(t, B_t)$ satisfies a 2BSDE, which is formulated under a nondominated family of singular probability measures given by the law of X^α 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 nondominated singular measures), and under which the forward process can be simulated. The idea, used in [16] for quasi variational inequalities arising in impulse control problems, is the following. We introduce a Poisson random measure $\mu_A(dt,da)$ on $\mathbb{R}_+ \times A$ with finite intensity measure $\lambda_A(da)dt$ associated to the marked point process $(\tau_i,\zeta_i)_i$, independent of W, and consider the pure jump process $(I_t)_t$ equal to the mark ζ_i valued in A between two jump times τ_i and τ_{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 $\mathbb{F} = \mathbb{F}^{W,\mu_A}$:

$$Y_t = g(X_T) + \int_t^T f(X_s, I_s)ds - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a)\tilde{\mu}_A(ds,da), \tag{1.8}$$

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

$$\frac{\partial v}{\partial t} + b(x,a).D_x v + \frac{1}{2}\mathrm{tr}(\sigma\sigma^{\intercal}(x,a)D_x^2 v) \tag{1.9}$$

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

$$v(T,x,a) = g(x), \quad (x,a) \in \mathbb{R}^d \times A, \tag{1.10}$$

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 (1.2), we constrain the jump component to the BSDE (1.8) to be nonpositive, i.e.

$$U_t(a) \le 0, \quad \forall (t,a). \tag{1.11}$$

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 μ, i.e of I, and assuming that v is continuous, the constraint (1.11) 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 (1.9) 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 (1.9), we then obtain the nonlinear HJB equation (1.2).

Inspired by the above discussion, we now introduce the following general class of BSDE with partially nonpositive jumps, which is a non Markovian extension of (1.8)-(1.11):

$$\begin{aligned} Y_t &= \xi + \int_t^T F(s, \omega, Y_s, Z_s, U_s) ds + K_T - K_t \\ &\quad - \int_t^T Z_s dW_s - \int_t^T \int_E U_s(e) \tilde{\mu}(ds, de), \quad 0 \leq t \leq T, \ a.s. \end{aligned} \tag{1.12}$$

with

$$U_t(e) \leq 0, \quad d\mathbb{P} \otimes dt \otimes \lambda(de) \ \ a.e. \text{ on } \Omega \times [0, T] \times A. \tag{1.13}$$

Here μ is a Poisson random measure on $\mathbb{R}_+ \times E$ with intensity measure $\lambda(de)dt$, A a subset of E, ξ an $\mathcal{F}_T$ measurable random variable, and F a generator function. 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 (1.13) 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 (1.12)-(1.13), we must have $Y \leq \bar{Y}$.

We use a penalization method for constructing an approximating sequence $(Y^n, Z^n, U^n, K^n)_n$ of BSDEs with jumps, and prove that it converges to the minimal solution that we are looking for. The proof relies on comparison results, uniform estimates and monotonic convergence theorem for BSDEs with jumps. Notice that compared to [16], we do not assume that the intensity measure λ of μ is finite on the whole set E, but only on the subset A on which the jump constraint is imposed. Moreover in [16], the process I does not influence directly the coefficients of the process X, which is Markov in itself. In contrast, in this paper, we need to enlarge the state variables by considering the additional state variable I, which makes Markov the forward regime switching jump-diffusion process (X, I). Our main result is then to relate the minimal solution to the BSDE with A-nonpositive jumps to a fully nonlinear IPDE of HJB type:

$$\begin{aligned} &\frac{\partial v}{\partial t} + \sup_{a \in A} \Big[b(x, a).D_x v(t, x) + \frac{1}{2} \mathrm{tr}(\sigma\sigma^\intercal(x, a) D_x^2 v(t, x)) \\ &+ \int_{E \setminus A} \big[v(t, x + \beta(x, a, e)) - v(t, x) - \beta(x, a, e).D_x v(t, x) \big] \lambda(de) \\ &\qquad + f\big(x, a, v, \sigma^\intercal(x, a) D_x v\big) \Big] = 0, \quad \text{on } [0, T) \times \mathbb{R}^d. \end{aligned}$$

This equation clearly extends HJB equation (1.2) by incorporating integral terms, and with a function f depending on v, $D_x v$ (actually, we may also allow f to depend on integral terms). By the Markov property of the forward regime switching jump-diffusion process, we easily see that the minimal solution to the the BSDE with A-nonpositive jumps is a deterministic function v of (t,x,a). The main task is to derive the key property that v does not actually depend on a, as a consequence of the A-nonpositive constrained jumps. This issue is a novelty with respect to the framework of [16] where there is a positive cost at each change of the regime I, while in the current paper, the cost is identically degenerate to zero. The proof relies on sharp arguments from viscosity solutions, inf-convolution and semiconcave approximation, as we don't know a priori any continuity results on v.

In the case where the generator function F or f does not depend on y, z, u, which corresponds to the stochastic control framework, we provide a dual representation of the minimal solution to the BSDE by means of a family of equivalent change of probability measures in the spirit of (1.6). This gives in particular an original representation for value functions of stochastic control problems, and unifies the weak formulation for both uncontrolled and controlled diffusion coefficient.

We conclude this introduction by pointing out that our results are stated without any ellipticity assumption on the diffusion coefficient, and includes the case of control affecting independently drift and diffusion, in contrast with the theory of second order BSDE. Moreover, our probabilistic BSDE representation leads to a new numerical scheme for HJB equation, based on the simulation of the forward process (X,I) and empirical regression methods, hence taking advantage of the high dimensional properties of Monte-Carlo method. Convergence analysis for the discrete time approximation of the BSDE with nonpositive jumps is studied in [14], while numerous numerical tests illustrate the efficiency of the method in [15].

The rest of the paper is organized as follows. In Section 2, we give a detailed formulation of BSDE with partially nonpositive jumps. We develop the penalization approach for studying the existence and the approximation of a unique minimal solution to our BSDE class, and give a dual representation of the minimal BSDE solution in the stochastic control case. We show in Section 3 how the minimal BSDE solution is related by means of viscosity solutions to the nonlinear IPDE of HJB type. Finally, we conclude in Section 4 by indicating extensions to our paper, and discussing probabilistic numerical scheme for the resolution of HJB equations.

2 BSDE with partially nonpositive jumps

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 μ on $\mathbb{R}_+ \times E$, where E is a Borelian subset of $\mathbb{R}^q$, endowed with its Borel σ-field $\mathcal{B}(E)$. We assume that the random measure μ has the intensity measure $\lambda(de)dt$ for some σ-finite measure λ on $(E, \mathcal{B}(E))$ satisfying

$$\int_E \big(1 \wedge |e|^2\big)\lambda(de) \quad < \quad \infty \,.$$

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

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

- $\mathbf{S^2}$ the set of real-valued càdlàg adapted processes $Y=(Y_t)_{0\leq t\leq T}$ such that $\|Y\|_{\mathbf{S^2}} := \Big(\mathbb{E}\Big[\sup_{0\leq t\leq T}|Y_t|^2\Big]\Big)^{\frac{1}{2}} < \infty$.
- $\mathbf{L^p(0,T)}$, $p\geq 1$, the set of real-valued adapted processes $(\phi_t)_{0\leq t\leq T}$ such that $\mathbb{E}\Big[\int_0^T|\phi_t|^p dt\Big] < \infty$.
- $\mathbf{L^p(W)}$, $p\geq 1$, the set of $\mathbb{R}^d$-valued $\mathcal{P}$-measurable processes $Z=(Z_t)_{0\leq t\leq T}$ such that $\|Z\|_{\mathbf{L^p(W)}} := \Big(\mathbb{E}\Big[\int_0^T|Z_t|^p dt\Big]\Big)^{\frac{1}{p}} < \infty$.
- $\mathbf{L^p}(\tilde{\mu})$, $p\geq 1$, the set of $\mathcal{P}\otimes\mathcal{B}(E)$-measurable maps $U:\Omega\times[0,T]\times E\to\mathbb{R}$ such that $\|U\|_{\mathbf{L^p}(\tilde{\mu})} := \Big(\mathbb{E}\big[\int_0^T\big(\int_E|U_t(e)|^2\lambda(de)\big)^{\frac{p}{2}}dt\big]\Big)^{\frac{1}{p}} < \infty$.
- $\mathbf{L^2}(\lambda)$ is the set of $\mathcal{B}(E)$-measurable maps $u:\ E\to\mathbb{R}$ such that $|u|_{\mathbf{L^2}(\lambda)} := \Big(\int_E|u(e)|^2\lambda(de)\Big)^{\frac{1}{2}} < \infty$.
- $\mathbf{K^2}$ the closed subset of $\mathbf{S^2}$ consisting of nondecreasing processes $K=(K_t)_{0\leq t\leq T}$ with $K_0=0$.

We are then given three objects:

1. A terminal condition ξ, which is an $\mathcal{F}_T$-measurable random variable.
2. A generator function $F:\Omega\times[0,T]\times\mathbb{R}\times\mathbb{R}^d\times\mathbf{L^2}(\lambda)\to\mathbb{R}$, which is a $\mathcal{P}\otimes\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(\mathbf{L^2}(\lambda))$-measurable map.
3. A Borelian subset A of E such that $\lambda(A)<\infty$.

We shall impose the following assumption on these objects:

(H0)

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

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

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

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

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

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

$$F(t,y,z,u) - F(t,y,z,u') \quad \leq \quad \int_E \gamma(t,e,y,z,u,u')(u(e)-u'(e))\lambda(de) \; ,$$

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

$$\begin{aligned} Y_t \quad &= \quad \xi + \int_t^T F(s, Y_s, Z_s, U_s)ds + K_T - K_t \\ & \quad - \int_t^T Z_s dW_s - \int_t^T \int_E U_s(e)\tilde{\mu}(ds,de), \quad 0 \leq t \leq T, \ a.s. \end{aligned} \tag{2.1}$$

with

$$U_t(e) \quad \leq \quad 0 \; , \qquad d\mathbb{P} \otimes dt \otimes \lambda(de) \;\; a.e. \text{ on } \Omega \times [0,T] \times A. \tag{2.2}$$

Definition 2.1 *A minimal solution to the BSDE with terminal data/generator* (ξ, F) *and* A*-nonpositive jumps is a quadruple of processes* $(Y,Z,U,K) \in \mathbf{S^2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{K^2}$ *satisfying* (2.1)-(2.2) *such that for any other quadruple* $(\bar{Y}, \bar{Z}, \bar{U}, \bar{K}) \in \mathbf{S^2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{K^2}$ *satisfying* (2.1)-(2.2)*, we have*

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

Remark 2.1 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 μ 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)-(2.2). □

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 $(\bar{Y}, \bar{Z}, \bar{K}, \bar{U}) \in \mathbf{S^2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{K^2}$ satisfying (2.1)-(2.2).

We shall see later in Lemma 3.1 how such condition is satisfied in a Markovian framework.

2.2 Existence and approximation by penalization

In this paragraph, we prove the existence of a minimal solution to (2.1)-(2.2), based on approximation via penalization. For each $n \in \mathbb{N}$, we introduce the penalized BSDE with jumps

$$\begin{aligned} Y_t^n &= \xi + \int_t^T F(s, Y_s^n, Z_s^n, U_s^n) ds + K_T^n - K_t^n \\ &\quad - \int_t^T Z_s^n dW_s - \int_t^T \int_E U_s^n(e) \tilde{\mu}(ds, de), \quad 0 \leq t \leq T, \end{aligned} \tag{2.3}$$

where K^n is the nondecreasing process in $\mathbf{K^2}$ defined by

$$K_t^n = n \int_0^t \int_A [U_s^n(e)]^+ \lambda(de) ds, \quad 0 \leq t \leq T.$$

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

$$Y_t^n = \xi + \int_t^T F_n(s, Y_s^n, Z_s^n, U_s^n) ds - \int_t^T Z_s^n dW_s - \int_t^T \int_E U_s^n(e) \tilde{\mu}(ds, de), \quad 0 \leq t \leq T,$$

where the generator F_n is defined by

$$F_n(t, y, z, u) = F(t, y, z, u) + n \int_A [u(e)]^+ \lambda(de),$$

for all $(t, y, z, u) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d \times \mathbf{L}^2(\lambda)$. Under **(H0)**(ii)-(iii) and since $\lambda(A) < \infty$, 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 [28], that under **(H0)**, BSDE (2.3) admits a unique solution $(Y^n, Z^n, U^n) \in \mathbf{S^2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu})$ for any $n \in \mathbb{N}$.

Lemma 2.1 *Let Assumption* **(H0)** *holds. The sequence* $(Y^n)_n$ *is nondecreasing, i.e.* $Y_t^n \leq Y_t^{n+1}$ *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 $(t, e, y, z, u) \in [0, T] \times E \times \mathbb{R} \times \mathbb{R}^d \times \mathbf{L^2}(\lambda)$. Under Assumption **(H0)**, we can apply the comparison Theorem 2.5 in [26], which shows that $Y_t^n \leq Y_t^{n+1}$, $0 \leq t \leq T$, a.s. □

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

Lemma 2.2 *Let Assumption* **(H0)** *holds. For any quadruple* $(\bar{Y}, \bar{Z}, \bar{U}, \bar{K}) \in \mathbf{S^2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{K^2}$ *satisfying* (2.1)-(2.2), *we have*

$$Y_t^n \leq \bar{Y}_t, \quad 0 \leq t \leq T, \ n \in \mathbb{N}. \tag{2.4}$$

Proof. Fix $n \in \mathbb{N}$, and consider a quadruple $(\bar{Y}, \bar{Z}, \bar{U}, \bar{K}) \in \mathbf{S^2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{K^2}$ solution to (2.1)-(2.2). Then, $\bar{U}$ clearly satisfies $\int_0^t \int_A [\bar{U}_s(e)]^+ \lambda(de)ds = 0$ for all $t \in [0,T]$, and so $(\bar{Y}, \bar{Z}, \bar{U}, \bar{K})$ is a supersolution to the penalized BSDE (2.3), i.e:

$$
\begin{aligned}
\bar{Y}_t &= \xi + \int_t^T F_n(s, \bar{Y}_s, \bar{Z}_s, \bar{U}_s)ds + \bar{K}_T - \bar{K}_t \\
&\quad - \int_t^T \bar{Z}_s dW_s - \int_t^T \int_E \bar{U}_s(e)\tilde{\mu}(ds, de), \quad 0 \le t \le T.
\end{aligned}
$$

By a slight adaptation of the comparison Theorem 2.5 in [26] under **(H0)**, we obtain the required inequality: $Y_t^n \le \bar{Y}_t$, $0 \le t \le T$. □

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

Lemma 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*

$$
\begin{aligned}
& \|Y^n\|^2_{\mathbf{S^2}} + \|Z^n\|^2_{\mathbf{L^2(W)}} + \|U^n\|^2_{\mathbf{L^2}(\tilde{\mu})} + \|K^n\|^2_{\mathbf{S^2}} \\
\le \; & C\Big(\mathbb{E}|\xi|^2 + \mathbb{E}\Big[\int_0^T |F(t,0,0,0)|^2 dt\Big] + \mathbb{E}\big[\sup_{0\le t\le T} |\bar{Y}_t|^2\big]\Big), \qquad \forall n \in \mathbb{N}. \qquad (2.5)
\end{aligned}
$$

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 $|Y_t^n|^2$, and observing that K^n is continuous and $\Delta Y_t^n = \int_E U_t^n(e)\mu(\{t\}, de)$, we have

$$
\begin{aligned}
\mathbb{E}|\xi|^2 &= \mathbb{E}|Y_t^n|^2 - 2\mathbb{E}\int_t^T Y_s^n F(s, Y_s^n, Z_s^n, U_s^n)ds - 2\mathbb{E}\int_t^T Y_s^n dK_s^n + \mathbb{E}\int_t^T |Z_s^n|^2 ds \\
&\quad + \mathbb{E}\int_t^T \int_E \left\{|Y_{s-}^n + U_s^n(e)|^2 - |Y_{s-}^n|^2 - 2Y_{s-}^n U_s^n(e)\right\} \mu(de, ds) \\
&= \mathbb{E}|Y_t^n|^2 + \mathbb{E}\int_t^T |Z_s^n|^2 ds + \mathbb{E}\int_t^T \int_E |U_s^n(e)|^2 \lambda(de)ds \\
&\quad - 2\mathbb{E}\int_t^T Y_s^n F(s, Y_s^n, Z_s^n, U_s^n)ds - 2\mathbb{E}\int_t^T Y_s^n dK_s^n, \quad 0 \le t \le T.
\end{aligned}
$$

From **(H0)**(iii), the inequality $Y_t^n \le \bar{Y}_t$ by Lemma 2.2 under **(H1)**, and the inequality $2ab \le \frac{1}{\alpha}a^2 + \alpha b^2$ for any constant $\alpha > 0$, we have:

$$
\begin{aligned}
& \mathbb{E}|Y_t^n|^2 + \mathbb{E}\int_t^T |Z_s^n|^2 ds + \mathbb{E}\int_t^T \int_E |U_s^n(e)|^2 \lambda(de)ds \\
\le \; & \mathbb{E}|\xi|^2 + C\mathbb{E}\int_t^T |Y_s^n| \Big(|F(s,0,0,0)| + |Y_s^n| + |Z_s^n| + |U_s^n|_{\mathbf{L^2}(\lambda)}\Big) ds \\
& + \frac{1}{\alpha}\mathbb{E}\Big[\sup_{s\in[0,T]} |\bar{Y}_s|^2\Big] + \alpha\mathbb{E}|K_T^n - K_t^n|^2.
\end{aligned}
$$

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

$$
\begin{aligned}
& \mathbb{E}|Y_t^n|^2 + \frac{1}{2}\mathbb{E}\int_t^T |Z_s^n|^2 ds + \frac{1}{2}\mathbb{E}\int_t^T \int_E |U_s^n(e)|^2 \lambda(de) ds \qquad (2.6) \\
\leq \;\; & C\mathbb{E}\int_t^T |Y_s^n|^2 ds + \mathbb{E}|\xi|^2 + \frac{1}{2}\mathbb{E}\int_0^T |F(s,0,0,0)|^2 ds + \frac{1}{\alpha}\mathbb{E}\Big[\sup_{s\in[0,T]} |\bar{Y}_s|^2\Big] + \alpha \mathbb{E}|K_T^n - K_t^n|^2 .
\end{aligned}
$$

Now, from the relation (2.3), we have:

$$
\begin{aligned}
K_T^n - K_t^n &= Y_t^n - \xi - \int_t^T F(s, Y_s^n, Z_s^n, U_s^n) ds \\
& \quad + \int_t^T Z_s^n dW_s + \int_t^T \int_E U_s^n(e) \tilde{\mu}(ds, de).
\end{aligned}
$$

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

$$
\begin{aligned}
\mathbb{E}|K_T^n - K_t^n|^2 &\leq C_1\Big(\mathbb{E}|\xi|^2 + \mathbb{E}\int_0^T |F(s,0,0,0)|^2 ds + \mathbb{E}|Y_t^n|^2 \\
& \quad + \mathbb{E}\int_t^T \big(|Y_s^n|^2 + |Z_s^n|^2 + |U_s^n|^2_{\mathbf{L^2}(\lambda)}\big) ds\Big), \quad 0 \leq t \leq T. \qquad (2.7)
\end{aligned}
$$

Hence, by choosing $\alpha > 0$ s.t. $C_1\alpha \leq \frac{1}{4}$, and plugging into (2.6), we get

$$
\begin{aligned}
& \frac{3}{4}\mathbb{E}|Y_t^n|^2 + \frac{1}{4}\mathbb{E}\int_t^T |Z_s^n|^2 ds + \frac{1}{4}\mathbb{E}\int_t^T \int_E |U_s^n(e)|^2 \lambda(de) ds \\
\leq \;\; & C\mathbb{E}\int_t^T |Y_s^n|^2 ds + \frac{5}{4}\mathbb{E}|\xi|^2 + \frac{1}{4}\mathbb{E}\int_0^T |F(s,0,0,0)|^2 ds + \frac{1}{\alpha}\mathbb{E}\big[\sup_{s\in[0,T]} |\bar{Y}_s|^2\big], \quad 0 \leq t \leq T.
\end{aligned}
$$

Thus application of Gronwall's lemma to $t \mapsto \mathbb{E}|Y_t^n|^2$ yields:

$$
\begin{aligned}
& \sup_{0\leq t\leq T} \mathbb{E}|Y_t^n|^2 + \mathbb{E}\int_0^T |Z_t^n|^2 dt + \mathbb{E}\int_0^T \int_E |U_t^n(e)|^2 \lambda(de) dt \\
\leq \;\; & C\Big(\mathbb{E}|\xi|^2 + \mathbb{E}\int_0^T |F(t,0,0,0)|^2 dt + \mathbb{E}\big[\sup_{t\in[0,T]} |\bar{Y}_t|^2\big]\Big), \qquad (2.8)
\end{aligned}
$$

which gives the required uniform estimates (2.5) for $(Z^n, U^n)_n$ and also $(K^n)_n$ by (2.7). Finally, by writing from (2.3) that

$$
\begin{aligned}
\sup_{0\leq t\leq T} |Y_t^n| &\leq |\xi| + \int_0^T |F(t, Y_t^n, Z_t^n, U_t^n)| dt + K_T^n \\
& \quad + \sup_{0\leq t\leq T} \Big|\int_0^t Z_s^n dW_s\Big| + \sup_{0\leq t\leq T} \Big|\int_0^t \int_E U_s^n(e) \tilde{\mu}(ds, de)\Big|,
\end{aligned}
$$

we obtain the required uniform estimate (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 2.1 *Under* **(H0)** *and* **(H1)**, *there exists a unique minimal solution* (Y, Z, U, K) $\in \mathbf{S^2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu}) \times \mathbf{K^2}$ *with* K *predictable, to* (2.1)-(2.2). Y *is the increasing limit of* $(Y^n)_n$ *and also in* $\mathbf{L^2(0,T)}$, K_t *is the weak limit of* $(K_t^n)_n$ *in* $\mathbf{L^2}(\mathbf{\Omega}, \mathcal{F}_\mathbf{t}, \mathbb{P})$ *for all* $t \in [0,T]$, *and for any* $p \in [1,2)$,

$$\|Z^n - Z\|_{\mathbf{L^p(W)}} + \|U^n - U\|_{\mathbf{L^p}(\tilde{\mu})} \quad \longrightarrow \quad 0,$$

as n *goes to infinity.*

Proof. By the Lemmata 2.1 and 2.2, $(Y^n)_n$ converges increasingly to some adapted process Y, satisfying: $\|Y\|_{\mathbf{S^2}} < \infty$ by the uniform estimate for $(Y^n)_n$ in Lemma 2.3 and Fatou's lemma. Moreover by dominated convergence theorem, the convergence of $(Y^n)_n$ to Y also holds in $\mathbf{L^2(0,T)}$. Next, by the uniform estimates for $(Z^n, U^n, K^n)_n$ in Lemma 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 [21] for BSDE. This provides the existence of $(Z, U) \in \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\mu})$, and K predictable, nondecreasing with $\mathbb{E}[K_T^2] < \infty$, such that the sequence $(Z^n, U^n, K^n)_n$ converges in the sense of Theorem 2.1 to (Z, U, K) satisfying:

$$\begin{aligned} Y_t \quad &= \quad \xi + \int_t^T F(s, Y_s, Z_s, U_s)ds + K_T - K_t \\ &\quad - \int_t^T Z_s dW_s - \int_t^T \int_E U_s(e)\tilde{\mu}(ds, de), \quad 0 \le t \le T. \end{aligned}$$

Thus, the process Y is the difference of a càd-làg process and the nondecreasing process K, and by Lemma 2.2 in [21], this implies that Y and K are also càd-làg, hence respectively in $\mathbf{S^2}$ and $\mathbf{K^2}$. Moreover, from the strong convergence in $\mathbf{L^1}(\tilde{\mu})$ of $(U^n)_n$ to U and since $\lambda(A) < \infty$, we have

$$\mathbb{E}\int_0^T \int_A [U_s^n(e)]^+ \lambda(de)ds \quad \longrightarrow \quad \mathbb{E}\int_0^T \int_A [U_s(e)]^+ \lambda(de)ds,$$

as n goes to infinity. Since $K_T^n = n\int_0^T \int_A [U_s^n(e)]^+ \lambda(de)ds$ is bounded in $\mathbf{L^2}(\Omega, \mathcal{F}_T, \mathbb{P})$, this implies

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

which means that the A-nonpositive constraint (2.2) is satisfied. Hence, (Y, Z, K, U) is a solution to the constrained BSDE (2.1)-(2.2), and by Lemma 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. □

2.3 Dual representation

In this subsection, 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 A-nonpositive jumps in terms of a family of equivalent probability measures.

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

$$
\begin{aligned}
L_t^\nu &:= \mathcal{E}\Big(\int_0^{\cdot} \int_E (\nu_s(e) - 1) \tilde{\mu}(ds, de) \Big)_t \\
&= \exp\Big(\int_0^t \int_E \ln \nu_s(e) \mu(ds, de) - \int_0^t \int_E (\nu_s(e) - 1) \lambda(de) ds \Big), \quad 0 \le t \le T. \qquad (2.9)
\end{aligned}
$$

When L^ν 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:

$$
\frac{d\mathbb{P}^\nu}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = L_t^\nu, \quad 0 \le t \le T, \qquad (2.10)
$$

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 $\lambda(de)dt$ of μ under $\mathbb{P}$ to $\nu_t(e)\lambda(de)dt$ under $\mathbb{P}^\nu$. We denote by $\tilde{\mu}^\nu(dt, de) = \mu(dt, de) - \nu_t(e)\lambda(de)dt$ the compensated martingale measure of μ under $\mathbb{P}^\nu$. We then introduce the subset $\mathcal{V}_A$ of $\mathcal{V}$ by:

$$
\begin{aligned}
\mathcal{V}_A = \Big\{ \nu \in \mathcal{V}, \ \text{valued in } [1, \infty) \ \text{and essentially bounded :} \\
\nu_t(e) = 1, \quad e \in E \setminus A, \ d\mathbb{P} \otimes dt \otimes \lambda(de) \ a.e. \Big\},
\end{aligned}
$$

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

Lemma 2.4 *For any $\nu \in \mathcal{V}_A$, L^ν is a uniformly integrable martingale, and L_T^ν is square integrable.*

Proof. Several sufficient criteria for L^ν to be a uniformly integrable martingale are known. We refer for example to the recent paper [25], which shows that if

$$
S_T^\nu := \exp\Big(\int_0^T \int_E |\nu_t(e) - 1|^2 \lambda(de) dt \Big)
$$

is integrable, then L^ν is uniformly integrable. By definition of $\mathcal{V}_A$, we see that for $\nu \in \mathcal{V}_A$,

$$
S_T^\nu = \exp\Big(\int_0^T \int_A |\nu_t(e) - 1|^2 \lambda(de) dt \Big),
$$

which is essentially bounded since ν is essentially bounded and $\lambda(A) < \infty$. Moreover, from the explicit form (2.9) of L^ν, we have $|L_T^\nu|^2 = L_T^{\nu^2} S_T^\nu$, and so $\mathbb{E}|L_T^\nu|^2 \le \|S_T^\nu\|_\infty$. □

We can then associate to each $\nu \in \mathcal{V}_A$ the probability measure $\mathbb{P}^\nu$ through (2.10). We first provide a dual representation of the penalized BSDEs in terms of such $\mathbb{P}^\nu$. To this end, we need the following Lemma.

Lemma 2.5 *Let $\phi \in \mathbf{L^2(W)}$ and $\psi \in \mathbf{L^2}(\tilde{\mu})$. Then for every $\nu \in \mathcal{V}_A$, the processes $\int_0^{\cdot} \phi_t dW_t$ and $\int_0^{\cdot} \int_E \psi_t(e) \tilde{\mu}^\nu(dt, de)$ are $\mathbb{P}^\nu$-martingales.*

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

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

since L_T^ν is square integrable by Lemma 2.4, and $\phi \in \mathbf{L^2(W)}$. This implies that M^ϕ is $\mathbb{P}^\nu$-uniformly integrable, and hence a true $\mathbb{P}^\nu$-martingale. The proof for $\int_0^\cdot \int_E \phi_t(e)\tilde{\mu}^\nu(dt,de)$ follows exactly the same lines and is therefore omitted. □

Proposition 2.1 *For all $n \in \mathbb{N}$, the solution to the penalized BSDE (2.3) is explicitly represented as*

$$Y_t^n = \operatorname*{ess\,sup}_{\nu\in\mathcal{V}_A^n} \mathbb{E}^\nu\Big[\xi + \int_t^T F(s)ds\Big|\mathcal{F}_t\Big], \quad 0 \leq t \leq T. \tag{2.11}$$

Proof. Fix $n \in \mathbb{N}$. For any $\nu \in \mathcal{V}_A^n$, and by introducing the compensated martingale measure $\tilde{\mu}^\nu(dt,de) = \tilde{\mu}(dt,de) - (\nu_t(e)-1)\lambda(de)dt$ under $\mathbb{P}^\nu$, we see that the solution (Y^n, Z^n, U^n) to the BSDE (2.3) satisfies:

$$\begin{aligned}Y_t^n &= \xi + \int_t^T \Big[F(s) + \int_A \Big(n[U_s^n(e)]^+ - (\nu_s(e)-1)U_s^n(e)\Big)\lambda(de)\Big]ds \\ &\quad - \int_t^T \int_{E\setminus A} (\nu_s(e)-1)U_s^n(e)\lambda(de)ds - \int_t^T Z_s^n dW_s - \int_t^T \int_E U_s^n(e)\tilde{\mu}^\nu(ds,de).\end{aligned} \tag{2.12}$$

By definition of $\mathcal{V}_A$, we have

$$\int_t^T \int_{E\setminus A} (\nu_s(e)-1)U_s^n(e)\lambda(de)ds = 0, \quad 0 \leq t \leq T, \ a.s.$$

By taking expectation in (2.12) under $\mathbb{P}^\nu$ ($\sim \mathbb{P}$), we then get from Lemma 2.5:

$$Y_t^n = \mathbb{E}^\nu\Big[\xi + \int_t^T \Big(F(s) + \int_A \big(n[U_s^n(e)]^+ - (\nu_s(e)-1)U_s^n(e)\big)\lambda(de)\Big)ds\Big|\mathcal{F}_t\Big]. \tag{2.13}$$

Now, observe that for any $\nu \in \mathcal{V}_A^n$, hence valued in $[1, n+1]$, we have

$$n[U_t^n(e)]^+ - (\nu_t(e)-1)U_t^n(e) \geq 0, \quad d\mathbb{P}\otimes dt\otimes\lambda(de) \ a.e.$$

which yields by (2.13):

$$Y_t^n \geq \operatorname*{ess\,sup}_{\nu\in\mathcal{V}_A^n} \mathbb{E}^\nu\Big[\xi + \int_t^T F(s)ds\Big|\mathcal{F}_t\Big]. \tag{2.14}$$

On the other hand, let us consider the process $\nu^* \in \mathcal{V}_A^n$ defined by

$$\nu_t^*(e) = \mathbb{1}_{e\in E\setminus A} + \Big(\mathbb{1}_{U_t(e)\leq 0} + (n+1)\mathbb{1}_{U_t(e)>0}\Big)\mathbb{1}_{e\in A}, \quad 0 \leq t \leq T, e \in E.$$

By construction, we clearly have

$$n[U_t^n(e)]^+ - (\nu_t^*(e)-1)U_t^n(e) \;=\; 0, \quad \forall 0 \le t \le T,\ e \in A,$$

and thus for this choice of $\nu = \nu^*$ in (2.13):

$$Y_t^n \;=\; \mathbb{E}^{\nu^*}\Big[\xi + \int_t^T F(s)ds\Big|\mathcal{F}_t\Big].$$

Together with (2.14), this proves the required representation of Y^n. □

Remark 2.2 Arguments in the proof of Proposition 2.1 shows that the relation (2.11) holds for general generator function F depending on (y,z,u), i.e.

$$Y_t^n \;=\; \operatorname*{ess\,sup}_{\nu \in \mathcal{V}_A^n} \mathbb{E}^\nu\Big[\xi + \int_t^T F(s, Y_s^n, Z_s^n, U_s^n)ds\Big|\mathcal{F}_t\Big],$$

which is in this case an implicit relation for Y^n. Moreover, the essential supremum in this dual representation is attained for some ν^*, 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 2.2 *Under **(H1)**, the minimal solution to* (2.1)-(2.2) *is explicitly represented as*

$$Y_t \;=\; \operatorname*{ess\,sup}_{\nu \in \mathcal{V}_A} \mathbb{E}^\nu\Big[\xi + \int_t^T F(s)ds\Big|\mathcal{F}_t\Big], \quad 0 \le t \le T. \tag{2.15}$$

Proof. Let $(Y,Z,U,K) \in$ be the minimal solution to (2.1)-(2.2). Let us denote by $\tilde{Y}$ the process defined in the r.h.s of (2.15). Since $\mathcal{V}_A^n \subset \mathcal{V}_A$, it is clear from the representation (2.11) that $Y_t^n \le \tilde{Y}_t$, for all n. Recalling from Theorem 2.1 that Y is the pointwise limit of Y^n, we deduce that $Y_t = \lim_{n\to\infty} Y_t^n \le \tilde{Y}_t$, $0 \le t \le T$.

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

$$\begin{aligned} Y_t \;=\; & \xi + \int_t^T \Big[F(s) - \int_A (\nu_s(e)-1)U_s(e)\lambda(de)\Big]ds + K_T - K_t \\ & - \int_t^T \int_{E\setminus A} (\nu_s(e)-1)U_s(e)\lambda(de)ds - \int_t^T Z_s dW_s - \int_t^T \int_E U_s(e)\tilde{\mu}^\nu(ds,de). \end{aligned} \tag{2.16}$$

By definition of $\nu \in \mathcal{V}_A$, we have: $\int_t^T \int_{E\setminus A}(\nu_s(e)-1)U_s(e)\lambda(de)ds = 0$. Thus, by taking expectation in (2.16) under $\mathbb{P}^\nu$ from Lemma 2.5, and recalling that K is nondecreasing, we have:

$$\begin{aligned} Y_t \;\ge\; & \mathbb{E}^\nu\Big[\xi + \int_t^T \Big(F(s) - \int_A (\nu_s(e)-1)U_s(e)\lambda(de)\Big)ds\Big|\mathcal{F}_t\Big] \\ \ge\; & \mathbb{E}^\nu\Big[\xi + \int_t^T F(s)ds\Big|\mathcal{F}_t\Big], \end{aligned}$$

since ν is valued in $[1,\infty)$, and U satisfies the nonpositive constraint (2.2). Since ν is arbitrary in $\mathcal{V}_A$, this proves the inequality $Y_t \ge \tilde{Y}_t$, and finally the required relation $Y = \tilde{Y}$. □

3 Nonlinear IPDE and Feynman-Kac formula

In this section, 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 are given a compact set A of $\mathbb{R}^q$, and a Borelian subset $L \subset \mathbb{R}^l \setminus \{0\}$, equipped with respective Borel σ-fields $\mathcal{B}(A)$ and $\mathcal{B}(L)$. We assume that

(HA) The interior set $\mathring{A}$ of A is connex, and $A = \mathrm{Adh}(\mathring{A})$, the closure of its interior.

We consider the case where $E = L \cup A$ and we may assume w.l.o.g. that $L \cap A = \emptyset$ by identifying A and L respectively with the sets $A \times \{0\}$ and $\{0\} \times L$ in $\mathbb{R}^q \times \mathbb{R}^l$. We consider two independent Poisson random measures ϑ and π defined respectively on $\mathbb{R}_+ \times L$ and $\mathbb{R}_+ \times A$. We suppose that ϑ and π have respective intensity measures $\lambda_\vartheta(d\ell)dt$ and $\lambda_\pi(da)dt$ where λ_ϑ and λ_π are two σ-finite measures with respective supports L and A, and satisfying

$$\int_L (1 \wedge |\ell|^2)\lambda_\vartheta(d\ell) \ < \ \infty \qquad \text{and} \quad \int_A \lambda_\pi(da) \ < \ \infty ,$$

and we denote by $\tilde{\vartheta}(dt, d\ell) = \vartheta(dt, d\ell) - \lambda_\vartheta(d\ell)dt$ and $\tilde{\pi}(dt, da) = \pi(dt, da) - \lambda_\pi(da)dt$ the compensated martingale measures of ϑ and π respectively. We also assume that

(Hλ_π)

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

(ii) The boundary of A: $\partial A = A \setminus \mathring{A}$, is negligible w.r.t. λ_π, i.e. $\lambda_\pi(\partial A) = 0$.

In this context, by taking a random measure μ on $\mathbb{R}_+ \times E$ in the form, $\mu = \vartheta + \pi$, we notice that it remains a Poisson random measure with intensity measure $\lambda(de)dt$ given by

$$\int_E \varphi(e)\lambda(de) \ = \ \int_L \varphi(\ell)\lambda_\vartheta(d\ell) + \int_A \varphi(a)\lambda_\pi(da) ,$$

for any measurable function φ from E to $\mathbb{R}$, and we have the following identifications

$$\mathbf{L^2}(\tilde{\mu}) \ = \ \mathbf{L^2}(\tilde{\vartheta}) \times \mathbf{L^2}(\tilde{\pi}) , \qquad \mathbf{L^2}(\lambda) \ = \ \mathbf{L^2}(\lambda_\vartheta) \times \mathbf{L^2}(\lambda_\pi) , \tag{3.1}$$

where

- $\mathbf{L^2}(\tilde{\vartheta})$ is the set of $\mathcal{P} \otimes \mathcal{B}(L)$-measurable maps $U : \Omega \times [0, T] \times L \to \mathbb{R}$ such that

$$\|U\|_{\mathbf{L^2}(\tilde{\vartheta})} \ := \ \Big(\mathbb{E}\Big[\int_0^T \int_L |U_t(\ell)|^2 \lambda_\vartheta(d\ell)dt\Big]\Big)^{\frac{1}{2}} < \infty ,$$

• $\mathbf{L^2}(\tilde{\pi})$ is the set of $\mathcal{P} \otimes \mathcal{B}(A)$-measurable maps $R : \Omega \times [0,T] \times A \to \mathbb{R}$ such that

$$\|R\|_{\mathbf{L^2}(\tilde{\pi})} \;:=\; \Big(\mathbb{E}\Big[\int_0^T \int_A |R_t(a)|^2 \lambda_\pi(da)dt\Big]\Big)^{\frac{1}{2}} < \infty \,,$$

• $\mathbf{L^2}(\lambda_\vartheta)$ is the set of $\mathcal{B}(L)$-measurable maps $u : L \to \mathbb{R}$ such that

$$|u|_{\mathbf{L^2}(\lambda_\vartheta)} \;:=\; \Big(\int_L |u(\ell)|^2 \lambda_\vartheta(d\ell)\Big]\Big)^{\frac{1}{2}} < \infty \,,$$

• $\mathbf{L^2}(\lambda_\pi)$ is the set of $\mathcal{B}(A)$-measurable maps $r : A \to \mathbb{R}$ such that

$$|r|_{\mathbf{L^2}(\lambda_\pi)} \;:=\; \Big(\int_A |r(a)|^2 \lambda_\pi(da)\Big)^{\frac{1}{2}} < \infty \,.$$

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:

$$dX_s \;=\; b(X_s, I_s)ds + \sigma(X_s, I_s)dW_s + \int_L \beta(X_{s^-}, I_{s^-}, \ell)\tilde{\vartheta}(ds, d\ell), \tag{3.2}$$

$$dI_s \;=\; \int_A \big(a - I_{s^-}\big)\pi(ds, da). \tag{3.3}$$

In other words, I is the pure jump process valued in A associated to the Poisson random measure π, which changes the coefficients of jump-diffusion process X. We make the usual assumptions on the forward jump-diffusion coefficients:

(HFC)

(i) There exists a constant C such that

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

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

(ii) There exists a constant C such that

$$\begin{aligned} \big|\beta(x, a, \ell)\big| \;&\leq\; C(1 + |x|)\big(1 \wedge |\ell|\big), \\ \big|\beta(x, a, \ell) - \beta(x', a', \ell)\big| \;&\leq\; C\big(|x - x'| + |a - a'|\big)\big(1 \wedge |\ell|\big), \end{aligned}$$

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

Remark 3.1 We do not make any ellipticity assumption on σ. In particular, some lines and columns of σ 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, σ and β 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:

$$\begin{aligned} dX_s \;&=\; b(X_s, \Theta_s, I_s)ds + \sigma(X_s, \Theta_s, I_s)dW_s + \int_L \beta(X_{s^-}, \Theta_{s^-}, I_{s^-}, \ell)\tilde{\vartheta}(ds, d\ell), \\ d\Theta_s \;&=\; ds \\ dI_s \;&=\; \int_A \big(a - I_{s^-}\big)\pi(ds, da). \end{aligned}$$

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

Under these conditions, existence and uniqueness of a solution $(X_s^{t,x,a}, I_s^{t,a})_{t\leq s\leq T}$ to (3.2)-(3.3) 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.

$$\mathbb{E}\Big[\sup_{t\leq s\leq T}|X_s^{t,x,a}|^p + |I_s^{t,a}|^p\Big] \;\leq\; C_p(1+|x|^p+|a|^p)\;, \tag{3.4}$$

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 $f:\mathbb{R}^d\times\mathbb{R}^q\times\mathbb{R}\times\mathbb{R}^d\times\mathbf{L}^2(\lambda_\vartheta)\to\mathbb{R}$. We make the following assumptions on the BSDE coefficients:

(HBC1)

(i) The functions g and $f(.,0,0,0)$ satisfy a polynomial growth condition:

$$\sup_{x\in\mathbb{R}^d,\,a\in\mathbb{R}^q}\frac{|g(x,a)|+|f(x,a,0,0,0)|}{1+|x|^m+|a|^m} \;<\; \infty,$$

for some $m\geq 0$.

(ii) There exists some constant C s.t.

$$\begin{aligned}&|f(x,a,y,z,u)-f(x',a',y',z',u')|\\ \leq\;& C\big(|x-x'|+|a-a'|+|y-y'|+|z-z'|+|u-u'|_{\mathbf{L^2}(\lambda_\vartheta)}\big)\;,\end{aligned}$$

for all $x,x'\in\mathbb{R}^d$, $y,y'\in\mathbb{R}$, $z,z'\in\mathbb{R}^d$, $a,a'\in\mathbb{R}^q$ and $u,u'\in\mathbf{L^2}(\lambda_\vartheta)$.

(HBC2) The generator function f satisfies the monotonicity condition:

$$f(x,a,y,z,u)-f(x,a,y,z,u') \;\leq\; \int_L\gamma(x,a,\ell,y,z,u,u')(u(\ell)-u'(\ell))\lambda_\vartheta(d\ell)\;,$$

for all $x\in\mathbb{R}^d$, $a\in\mathbb{R}^q$, $z\in\mathbb{R}^d$, $y\in\mathbb{R}$ and $u,u'\in\mathbf{L^2}(\lambda_\vartheta)$, where $\gamma:\mathbb{R}^d\times E\times\mathbb{R}\times\mathbb{R}^d\times\mathbf{L^2}(\lambda_\vartheta)\times\mathbf{L^2}(\lambda_\vartheta)\to\mathbb{R}$ is a $\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(E)\otimes\mathcal{B}(\mathbb{R})\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(\mathbf{L^2}(\lambda_\vartheta))\otimes\mathcal{B}(\mathbf{L^2}(\lambda_\vartheta))$-measurable map satisfying: $C_1(1\wedge|\ell|)\leq\gamma(x,a,\ell,y,z,u,u')\leq C_2(1\wedge|\ell|)$, for $\ell\in L$, with two constants $-1<C_1\leq 0\leq C_2$.

Let us also consider an assumption on the dependence of f w.r.t. the jump component used in [2], and stronger than **(HBC2)**.

(HBC2') The generator function f is of the form

$$f(x,a,y,z,u) \;=\; h\big(x,a,y,z,\int_L u(\ell)\delta(x,\ell)\lambda_\vartheta(d\ell)\big)$$

for $(x,a,y,z,u)\in\mathbb{R}^d\times\mathbb{R}^q\times\mathbb{R}\times\mathbb{R}^d\times\mathbf{L^2}(\lambda)$, where

- δ is a measurable function on $\mathbb{R}^d \times L$ satisfying:

$$\begin{aligned} 0 &\leq \delta(x,\ell) \leq C(1 \wedge |\ell|), \\ |\delta(x,\ell) - \delta(x',\ell)| &\leq C|x-x'|(1 \wedge |\ell|^2), \quad x, x' \in \mathbb{R}^d, \ell \in L, \end{aligned}$$

  for some positive constant C.

- h is a continuous function on $\mathbb{R}^d \times \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$ such that $\rho \mapsto h(x,a,y,z,\rho)$ is nondecreasing for all $(x,a,y,z) \in \mathbb{R}^d \times \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^d$, and satisfying for some positive constant C:

$$|h(x,a,y,z,\rho) - h(x,a,y,z,\rho')| \leq C|\rho - \rho'|, \quad \rho, \rho' \in \mathbb{R},$$

  for all $(x,a,y,z) \in \mathbb{R}^d \times \mathbb{R}^q \times \mathbb{R} \times \mathbb{R}^d$.

Now with the identification (3.1), the BSDE problem (2.1)-(2.2) takes the following form: find the minimal solution $(Y,Z,U,R,K) \in \mathbf{S^2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde{\vartheta}) \times \mathbf{L^2}(\tilde{\pi}) \times \mathbf{K^2}$ to

$$\begin{aligned} Y_t &= g(X_T, I_T) + \int_t^T f\big(X_s, I_s, Y_s, Z_s, U_s\big) ds + K_T - K_t \\ &\quad - \int_t^T Z_s.dW_s - \int_t^T \int_L U_s(\ell)\tilde{\vartheta}(ds,d\ell) - \int_t^T \int_A R_s(a)\tilde{\pi}(ds,da), \end{aligned} \tag{3.5}$$

with

$$R_t(a) \leq 0, \quad d\mathbb{P} \otimes dt \otimes \lambda_\pi(da) \ \ a.e. \tag{3.6}$$

The main goal of this paper is to relate the BSDE (3.5) with A-nonpositive jumps (3.6) to the following nonlinear IPDE of HJB type:

$$-\frac{\partial w}{\partial t} - \sup_{a \in A}\Big[\mathcal{L}^a w + f\big(.,a,w,\sigma^\intercal(.,a)D_x w, \mathcal{M}^a w\big)\Big] = 0, \quad \text{on } [0,T) \times \mathbb{R}^d, \tag{3.7}$$

$$w(T,x) = \sup_{a \in A} g(x,a), \quad x \in \mathbb{R}^d, \tag{3.8}$$

where

$$\begin{aligned} \mathcal{L}^a w(t,x) &= b(x,a).D_x w(t,x) + \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x,a)D_x^2 w(t,x)) \\ &\quad + \int_L \big[w(t,x+\beta(x,a,\ell)) - w(t,x) - \beta(x,a,\ell).D_x w(t,x)\big]\lambda_\vartheta(d\ell), \\ \mathcal{M}^a w(t,x) &= \big(w(t,x+\beta(x,a,\ell)) - w(t,x)\big)_{\ell \in L}, \end{aligned}$$

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

Notice that under **(HBC1)**, **(HBC2)** and (3.4) (which follows from **(HFC)**), and with the identification (3.1), the generator $F(t,\omega,y,z,u,r) = f(X_t(\omega), I_t(\omega), y, z, u)$ 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 *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* $\{(\bar Y_s^{t,x,a}, \bar Z_s^{t,x,a}, \bar U_s^{t,x,a}, \bar R_s^{t,x,a}, \bar K_s^{t,x,a}), t \leq s \leq T\}$ *to the BSDE* (3.5)-(3.6) *when* $(X,I) = \{(X_s^{t,x,a}, I_s^{t,a}),\ t \leq s \leq T\}$, *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: for some* $p \geq 2$,

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \frac{|\bar v(t,x)|}{1+|x|^p} \quad < \quad \infty\,. \tag{3.9}$$

Proof. Under (**HBC1**) and since A is compact, we observe that there exists some $m \geq 0$ such that

$$C_{f,g} \;:=\; \sup_{x\in\mathbb{R}^d, a\in A} \frac{|g(x,a)| + |f(x,a,y,z,u)|}{1+|x|^m + |y| + |z| + |u|_{\mathbf{L^2}(\lambda_\vartheta)}} \quad < \quad \infty. \tag{3.10}$$

Let us then consider the smooth function $\bar v(t,x) = \bar C e^{\rho(T-t)}(1+|x|^p)$ for some positive constants $\bar C$ and ρ to be determined later, and with $p = \max(2,m)$. We claim that for $\bar C$ and ρ large enough, the function $\bar v$ is a classical supersolution to (3.7)-(3.8). Indeed, observe first that from the growth condition on g in (3.10), there exists $\bar C > 0$ s.t. $\hat g(x) := \sup_{a\in A} g(x,a) \leq \bar C(1+|x|^p)$ for all $x \in \mathbb{R}^d$. For such $\bar C$, we then have: $\bar v(T,.) \geq \hat g$. 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, σ, β by **(HFC)** (recall that A is compact), such that

$$\begin{aligned} -\frac{\partial \bar v}{\partial t} - \sup_{a\in A}\Big[\mathcal{L}^a \bar v + f\big(.,a,\bar v, \sigma^\intercal(.,a)D_x\bar v, \mathcal{M}^a \bar v\big)\Big] \quad &\geq \quad (\rho - C)\bar v \\ &\geq \quad 0, \end{aligned}$$

by choosing $\rho \geq C$. Let us now define the quintuple $(\bar Y, \bar Z, \bar U, \bar R, \bar K)$ by:

$$\begin{aligned} \bar Y_t \;&=\; \bar v(t,X_t) \;\text{ for }\; t < T, \quad \bar Y_T \;=\; g(X_T, I_T), \\ \bar Z_t \;&=\; \sigma^\intercal(X_{t^-}, I_{t^-}) D_x \bar v(t, X_{t^-}), \quad t \leq T, \\ \bar U_t \;&=\; \mathcal{M}^{I_{t^-}} \bar v(t, X_{t^-}), \quad \bar R_t \;=\; 0, \quad\; t \leq T \\ \bar K_t \;&=\; \int_0^t \Big[-\frac{\partial \bar v}{\partial t}(s,X_s) - \mathcal{L}^{I_s}\bar v(s,X_s) - f(X_s, I_s, \bar Z_s, \bar U_s)\Big] ds, \quad t < T \\ \bar K_T \;&=\; \bar K_{T^-} + \bar v(T, X_T) - g(X_T, I_T). \end{aligned}$$

From the supersolution property of $\bar v$ to (3.7)-(3.8), the process $\bar K$ is nondecreasing. Moreover, from the polynomial growth condition on $\bar v$, linear growth condition on b, σ, growth condition (3.10) on f, g and the estimate (3.4), we see that $(\bar Y, \bar Z, \bar U, \bar R, \bar K)$ lies in $\mathbf{S^2} \times \mathbf{L^2(W)} \times \mathbf{L^2}(\tilde\vartheta) \times \mathbf{L^2}(\tilde\pi) \times \mathbf{K^2}$. Finally, by applying Itô's formula to $\bar v(t,X_t)$, we conclude that $(\bar Y, \bar Z, \bar U, \bar R, \bar K)$ is solution a to (3.5), and the constraint (3.6) is trivially satisfied. □

Under **(HFC)**, **(HBC1)** and **(HBC2)**, we then get from Theorem 2.1 the existence of a unique minimal solution $\{(Y_s^{t,x,a}, Z_s^{t,x,a}, U_s^{t,x,a}, R_s^{t,x,a}, K_s^{t,x,a}),\ t \leq s \leq T\}$ to (3.5)-(3.6) when $(X,I) = \{(X_s^{t,x,a}, I_s^{t,a}),\ t \leq s \leq T\}$. Moreover, as we shall see in the next paragraph, this minimal solution is written in this Markovian context as: $Y_s^{t,x,a} = v(s, X_s^{t,x,a}, I_s^{t,x,a})$ where v is the deterministic function defined on $[0,T] \times \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}$ by:

$$v(t,x,a) \quad := \quad Y_t^{t,x,a}, \quad (t,x,a) \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^q. \tag{3.11}$$

We aim at proving that the function v defined by (3.11) does not depend actually on its argument a, and is a solution in a sense to be precised to the parabolic IPDE (3.7)-(3.8). 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 (3.7)-(3.8). 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

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

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

Definition 3.1 *(Viscosity solutions to* (3.7)-(3.8)*)*
(i) A function w, lsc (resp. usc) on $[0,T] \times \mathbb{R}^d$, is called a viscosity supersolution (resp. subsolution) to (3.7)-(3.8) *if*

$$w(T,x) \quad \geq (\textit{resp.} \;\; \leq) \quad \sup_{a \in A} g(x,a) \;,$$

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

$$\Big(-\frac{\partial \varphi}{\partial t} - \sup_{a \in A} \Big[\mathcal{L}^a \varphi + f(.,a,w,\sigma^{\intercal}(.,a) D_x \varphi, \mathcal{M}^a \varphi) \Big] \Big)(t,x) \quad \geq (\textit{resp.} \;\; \leq) \quad 0,$$

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

$$(w - \varphi)(t,x) \quad = \quad \min_{[0,T] \times \mathbb{R}^d} (w - \varphi) \;\; (\textit{resp.} \;\; \max_{[0,T] \times \mathbb{R}^d} (w - \varphi)) \;.$$

(ii) A locally bounded function w on $[0,T) \times \mathbb{R}^d$ is called a viscosity solution to (3.7)-(3.8) *if w_* is a viscosity supersolution and w^* is a viscosity subsolution to* (3.7)-(3.8).

We can now state the main result of this paper.

Theorem 3.1 *Assume that conditions **(HA)**, **(Hλ_π)**, **(HFC)**, **(HBC1)**, and **(HBC2)** hold. The function v in* (3.11) *does not depend on the variable a on $[0,T) \times \mathbb{R} \times \mathring{A}$ i.e.*

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

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:

$$v(t,x) \quad = \quad v(t,x,a), \quad (t,x) \in [0,T) \times \mathbb{R}^d, \tag{3.12}$$

for any $a \in \mathring{A}$. Then v is a viscosity solution to (3.7) *and a viscosity subsolution to* (3.8)*. Moreover, if **(HBC2')** holds, v is a viscosity supersolution to* (3.8)*.*

Remark 3.2 *1.* Once we have a uniqueness result for the fully nonlinear IPDE (3.7)-(3.8), Theorem 3.1 provides a Feynman-Kac representation of this unique solution by means of the minimal solution to the BSDE (3.5)-(3.6). This suggests consequently an original probabilistic numerical approximation of the nonlinear IPDE (3.7)-(3.8) by discretization and simulation of the minimal solution to the BSDE (3.5)-(3.6). This issue, especially the treatment of the nonpositive jump constraint, has been recently investigated in [14] and [15], 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 $\bar{\lambda}_\pi := \int_A \lambda_\pi(da)$, and so the pure jump process I is perfectly simulatable once we know how to simulate the distribution $\lambda_\pi(da)/\bar{\lambda}_\pi$ 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 (3.7)-(3.8). In the case where the generator function $f(x, a)$ does not depend on (y, z, u) (see Remark 3.3 below), comparison principle is proved in [23], and the result can be extended by same arguments when $f(x, a, y, z)$ depends also on y, z under the Lipschitz condition **(HBC1)**(ii). When f also depends on u, comparison principle is proved by [2] in the semilinear IPDE case, i.e. when A is reduced to a singleton, under condition **(HBC2')**. We also mention recent results on comparison principles for IPDE in [3] and references therein. □

Remark 3.3 *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, u). We then observe that the nonlinear IPDE (3.7) is the Hamilton-Jacobi-Bellman (HJB) equation associated to the following stochastic control problem: let us introduce the controlled jump-diffusion process:

$$dX_s^\alpha = b(X_s^\alpha, \alpha_s)ds + \sigma(X_s^\alpha, \alpha_s)dW_s + \int_L \beta(X_{s^-}^\alpha, \alpha_s, \ell)\tilde{\vartheta}(ds, d\ell), \tag{3.13}$$

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

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

where $\{X_s^{t,x,\alpha}, t \leq s \leq T\}$ denotes the solution to (3.13) starting from x at $s = t$, given a control $\alpha \in \mathcal{A}_{\mathbb{F}^0}$. It is well-known (see e.g. [23] or [18]) that the value function w is characterized as the unique viscosity solution to the dynamic programming HJB equation (3.7)-(3.8), and therefore by Theorem 3.1, $w = v$. 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 (3.5)-(3.6).
2. Combining the BSDE representation of Theorem 3.1 together with the dual representation in Theorem 2.2, we obtain an original representation for the value function of stochastic control problem:

$$\sup_{\alpha\in\mathcal{A}_{\mathbb{F}^0}} \mathbb{E}\Big[\int_0^T f(X_t^\alpha,\alpha_t)dt + g(X_T^\alpha,\alpha_T)\Big] \quad = \quad \sup_{\nu\in\mathcal{V}_A} \mathbb{E}^\nu\Big[\int_0^T f(X_t,I_t)dt + g(X_T,I_T)\Big]$$

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 [8]. 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 ϑ governing the controlled state process X^α, 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, σ and β of the controlled process are path-dependent. In this case, we know from the recent works by Ekren, Touzi, and Zhang [7] 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 [12] 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 σ is non degenerate, in contrast with [7] (see their Assumption 4.7). □

The rest of this paper is devoted to the proof of Theorem 3.1.

3.2 Viscosity property of the penalized BSDE

Let us consider the Markov penalized BSDE associated to (3.5)-(3.6):

$$\begin{aligned} Y_t^n \;&=\; g(X_T,I_T) + \int_t^T f(X_s,I_s,Y_s^n,Z_s^n,U_s^n)ds + n\int_t^T\int_A [R_s^n(a)]^+\lambda_\pi(da)ds \\ &\quad - \int_t^T Z_s^n.dW_s - \int_t^T\int_L U_s^n(\ell)\tilde{\vartheta}(ds,d\ell) - \int_t^T\int_A R_s^n(a)\tilde{\pi}(ds,da)\;, \end{aligned} \tag{3.14}$$

and denote by $\{(Y_s^{n,t,x,a}, Z_s^{n,t,x,a}, U_s^{n,t,x,a}, R_s^{n,t,x,a}),\ t \leq s \leq T\}$ the unique solution to (3.14) when $(X, I) = \{(X_s^{t,x,a}, I_s^{t,a}),\ t \leq s \leq T\}$ 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 [2] that $Y_s^{n,t,x,a} = v_n(s, X_s^{t,x,a}, I_s^{t,a})$, $t \leq s \leq T$, where v_n is the deterministic function defined on $[0, T] \times \mathbb{R}^d \times \mathbb{R}^q$ by:

$$v_n(t, x, a) \quad := \quad Y_t^{n,t,x,a}, \quad (t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q. \tag{3.15}$$

From the convergence result (Theorem 2.1) of the penalized solution, we deduce that the minimal solution of the constrained BSDE is actually in the form $Y_s^{t,x,a} = v(s, X_s^{t,x,a}, I_s^{t,a})$, $t \leq s \leq T$, with a deterministic function v defined in (3.11).

Moreover, from the uniform estimate (2.5) and Lemma 3.1, there exists some positive constant C s.t. for all n,

$$\begin{aligned} |v_n(t, x, a)|^2 \quad \leq \quad & C\Big(\mathbb{E}|g(X_T^{t,x,a}, I_T^{t,a})|^2 + \mathbb{E}\Big[\int_t^T |f(X_s^{t,x,a}, I_s^{t,a}, 0, 0, 0)|^2 ds\Big] \\ & + \mathbb{E}\Big[\sup_{t \leq s \leq T} |\bar{v}(s, X_s^{t,x,a})|^2\Big]\Big), \end{aligned}$$

for all $(t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q$. From the polynomial growth condition in **(HBC1)**(i) for g and f, (3.9) for $\bar{v}$, and the estimate (3.4) 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 and some $p \geq 2$, such that for all n:

$$|v_n(t, x, a)| + |v(t, x, a)| \quad \leq \quad C_v\big(1 + |x|^p + |a|^p\big), \quad \forall (t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q. \tag{3.16}$$

We now consider the parabolic semi-linear penalized IPDE for any n:

$$\begin{aligned} -\frac{\partial v_n}{\partial t}(t, x, a) - \mathcal{L}^a v_n(t, x, a) - f\big(x, a, v_n, \sigma^\intercal(x, a) D_x v_n, \mathcal{M}^a v_n\big) & & \quad (3.17)\\ - \int_A [v_n(t, x, a') - v_n(t, x, a)] \lambda_\pi(da') & & \\ -n \int_A [v_n(t, x, a') - v_n(t, x, a)]^+ \lambda_\pi(da') & = 0, \quad \text{on } [0, T) \times \mathbb{R}^d \times \mathbb{R}^q, \\ v_n(T, ., .) & = g, \quad \text{on } \mathbb{R}^d \times \mathbb{R}^q. & \quad (3.18) \end{aligned}$$

From Theorem 3.4 in Barles et al. [2], we have the well-known property that the penalized BSDE with jumps (2.3) provides a viscosity solution to the penalized IPDE (3.17)-(3.18). Actually, the relation in their paper is obtained under **(HBC2')**, which allows the authors to get comparison theorem for BSDE, but such comparison theorem also holds under the weaker condition **(HBC2)** as shown in [26], and we then get the following result.

Proposition 3.1 *Let Assumptions **(HFC)**, **(HBC1)**, and **(HBC2)** hold. The function v_n in* (3.15) *is a continuous viscosity solution to* (3.17)-(3.18)*, i.e. it is continuous on $[0, T] \times \mathbb{R}^d \times \mathbb{R}^q$, a viscosity supersolution (resp. subsolution) to* (3.18)*:*

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

for any $(x,a) \in \mathbb{R}^d \times \mathbb{R}^q$*, and a viscosity supersolution (resp. subsolution) to* (3.17)*:*

$$
\begin{aligned}
& -\frac{\partial \varphi}{\partial t}(t,x,a) - \mathcal{L}^a \varphi(t,x,a) \qquad (3.19)\\
& -f(x,a,v_n(t,x,a),\sigma^{\intercal}(x,a)D_x\varphi(t,x,a),\mathcal{M}^a\varphi(t,x,a))\\
-\int_A [\varphi(t,x,a') - \varphi(t,x,a)]\lambda_\pi(da') - n\int_A [\varphi(t,x,a') - \varphi(t,x,a)]^+\lambda_\pi(da') \quad & \geq \quad (\textit{resp.} \leq)\ 0,
\end{aligned}
$$

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*

$$
(v_n - \varphi)(t,x,a) \quad = \quad \min_{[0,T]\times\mathbb{R}^d\times\mathbb{R}^q}(v_n - \varphi) \quad (\textit{resp.} \max_{[0,T]\times\mathbb{R}^d\times\mathbb{R}^q}(v_n - \varphi))\ . \qquad (3.20)
$$

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 (3.17) 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 *Assume that* ***(Hλ_π)****,* ***(HFC)****, and* ***(HBC1)*** *hold. In the definition of* v_n *being a viscosity supersolution (resp. subsolution) to* (3.17) *at a point* $(t,x,a) \in [0,T) \times \mathbb{R}^d \times \mathring{A}$*, we can replace condition* (3.20) *by:*

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

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

Proof. We treat only the supersolution case as the subsolution case is proved by 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 (3.19) also holds for any test function φ in $C^{1,2,0}([0,T] \times \mathbb{R}^d \times \mathbb{R}^q)$ s.t.

$$
(v_n - \varphi)(t,x,a) \quad = \quad \min_{[0,T]\times\mathbb{R}^d\times\mathbb{R}^q}(v_n - \varphi)\ . \qquad (3.21)
$$

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

$$
\varphi^r(t',x',a') \quad = \quad \varphi(t',x',a')\Big(1 - \Phi\Big(\frac{|x'|^2 + |a'|^2}{r^2}\Big)\Big) - C_v\Phi\Big(\frac{|x'|^2 + |a'|^2}{r^2}\Big)\big(1 + |x'|^p + |a'|^p\big),
$$

where $C_v > 0$ and $p \geq 2$ are the constant and degree appearing in the polynomial growth condition (3.16) 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 $\varphi^r \in C^{1,2,0}([0,T] \times \mathbb{R}^d \times \mathbb{R}^q)$,

$$
(\varphi^r, D_x\varphi^r, D_x^2\varphi^r) \quad \longrightarrow \quad (\varphi, D_x\varphi, D_x^2\varphi) \quad \text{as} \quad r \to \infty \qquad (3.22)
$$

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

$$|\varphi^r(t',x',a')| \;\leq\; C_r\big(1+|x'|^p+|a'|^p\big) \tag{3.23}$$

for all $(t',x',a')\in[0,T]\times\mathbb{R}^q\times\mathbb{R}^d$. Since Φ is valued in $[0,1]$, we deduce from the polynomial growth condition (3.16) satisfied by v_n and (3.21) that $\varphi^r\leq v_n$ 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

$$(v_n-\varphi^r)(t,x,a) \;=\; \min_{[0,T]\times\mathbb{R}^d\times\mathbb{R}^q}(v_n-\varphi^r)\;, \tag{3.24}$$

for r large enough, and we may assume w.l.o.g. that this minimum is strict. Let $(\varphi^r_k)_k$ be a sequence of function in $C^{1,2}([0,T]\times(\mathbb{R}^d\times\mathbb{R}^q))$ satisfying (3.23) and such that

$$(\varphi^r_k, D_x\varphi^r_k, D^2_x\varphi^r_k) \;\longrightarrow\; (\varphi^r, D_x\varphi^r, D^2_x\varphi^r) \quad \text{as} \quad k\to\infty, \tag{3.25}$$

locally uniformly on $[0,T]\times\mathbb{R}^d\times\mathbb{R}^q$. From the growth conditions (3.16) and (3.23) on the continuous functions v_n and φ^r_k, we can assume w.l.o.g. (up to an usual negative perturbation of the function φ^k_r for large (x',a')), that there exists a bounded sequence $(t_k,x_k,a_k)_k$ in $[0,T]\times\mathbb{R}^d\times\mathbb{R}^q$ such that

$$(v_n-\varphi^r_k)(t_k,x_k,a_k) \;=\; \min_{[0,T]\times\mathbb{R}^d\times\mathbb{R}^q}(v_n-\varphi^r_k)\;. \tag{3.26}$$

The sequence $(t_k,x_k,a_k)_k$ converges up to a subsequence, and thus, by (3.24), (3.25) and (3.26), we have

$$(t_k,x_k,a_k) \;\to\; (t,x,a), \quad \text{as} \quad k\to\infty\;. \tag{3.27}$$

Now, from the viscosity supersolution property of v_n at (t_k,x_k,a_k) with the test function φ^r_k, we have

$$\begin{aligned}
&-\frac{\partial\varphi^r_k}{\partial t}(t_k,x_k,a_k)-\mathcal{L}^{a_k}\varphi^r_k(t_k,x_k,a_k)\\
&-f(x_k,a_k,v_n(t_k,x_k,a_k),\sigma^{\intercal}(x_k,a_k)D_x\varphi^r_k(t_k,x_k,a_k),\mathcal{M}^{a_k}\varphi^r_k(t_k,x_k,a_k))\\
&-\int_A[\varphi^r_k(t_k,x_k,a')-\varphi^r_k(t_k,x_k,a_k)]\lambda_\pi(da')\\
&-n\int_A[\varphi^r_k(t_k,x_k,a')-\varphi^r_k(t_k,x_k,a_k)]^+\lambda_\pi(da') \;\;\geq\;\; 0,
\end{aligned}$$

Sending k and r to infinity, and using (3.22), (3.25) and (3.27), we obtain the viscosity supersolution inequality at (t,x,a) with the test function φ.

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

$$0 \;=\; (v_n-\varphi)(t,x,a) \;=\; \min_{[0,T]\times\mathbb{R}^d\times\mathring{A}}(v_n-\varphi). \tag{3.28}$$

By same arguments as in (3.23), we can assume w.l.o.g. that φ 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 (3.16), and since A is compact, we have

$$(v_n - \varphi)(t', x', a') \quad \geq \quad -C(1 + |x'|^p + |d_A(a')|^p), \tag{3.29}$$

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

$$\varphi_\varepsilon(t', x', a') \quad = \quad \varphi(t', x', a') - \Phi\Big(\frac{d_{A_\varepsilon}(a')}{\varepsilon}\Big)C(1 + |x'|^p + |d_A(a')|^p)$$

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

$$A_\varepsilon \quad = \quad \big\{a' \in A \ : \ d_{\partial A}(a') \ \geq \ \varepsilon\big\}, \tag{3.30}$$

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

$$(\varphi_\varepsilon, D_x\varphi_\varepsilon, D_x^2\varphi_\varepsilon) \quad \longrightarrow \quad (\varphi, D_x\varphi, D_x^2\varphi) \quad \text{ as } \quad \varepsilon \to 0, \tag{3.31}$$

locally uniformly on $[0, T] \times \mathbb{R}^d \times \mathring{A}$. We notice from (3.29) and the definition of φ_ε that $\varphi_\varepsilon \leq v_n$ on $[0, T] \times \mathbb{R}^d \times A_\varepsilon^c$. Moreover, since $\varphi_\varepsilon \leq \varphi$ on $[0, T] \times \mathbb{R}^d \times \mathbb{R}^q$, $\varphi_\varepsilon = \varphi$ on $[0, T] \times \mathbb{R}^d \times \mathring{A}_\varepsilon$ and $a \in \mathring{A}$, we get by (3.28) for ε small enough

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

From Step 1, we then have

$$\begin{aligned} &-\frac{\partial\varphi_\varepsilon}{\partial t}(t, x, a) - \mathcal{L}^a\varphi_\varepsilon(t, x, a) \\ &-f(x, a, v_n(t, x, a), \sigma^\intercal(x, a)D_x\varphi_\varepsilon(t, x, a), \mathcal{M}^a\varphi_\varepsilon(t, x, a)) \\ -\int_A [\varphi_\varepsilon(t, x, a') - \varphi_\varepsilon(t, x, a)]\lambda_\pi(da') &- n\int_A [\varphi_\varepsilon(t, x, a') - \varphi_\varepsilon(t, x, a)]^+\lambda_\pi(da') \quad \geq \quad 0 \ . \end{aligned}$$

By sending ε to zero with (3.31), and using $a \in \mathring{A}$ with $(\mathbf{H}\lambda_\pi)$(ii), we get the required viscosity subsolution inequality at (t, x, a) for the test function φ. □

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 (3.5), and since for the minimal solution $(Y^{t,x,a}, Z^{t,x,a}, U^{t,x,a}, R^{t,x,a}, K^{t,x,a})$ to (3.5)-(3.6), the process $K^{t,x,a}$ is predictable, we observe that the A-jump component $R^{t,x,a}$ is expressed in terms of $Y^{t,x,a} = v(., X^{t,x,a}, I^{t,x,a})$ as:

$$R_s^{t,x,a}(a') \quad = \quad v(s, X_{s^-}^{t,x,a}, a') - v(s, X_{s^-}^{t,x,a}, I_{s^-}^{t,x,a}), \quad t \leq s \leq T, \ a' \in A,$$

for all $(t, x, a) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^q$. From the A-nonpositive constraint (3.6), this yields

$$\mathbb{E}\Big[\int_t^{t+h}\int_A \big[v(s, X_s^{t,x,a}, a') - v(s, X_s^{t,x,a}, I_s^{t,x,a})\big]^+\lambda_\pi(da')ds\Big] \quad = \quad 0,$$

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:

$$\int_A \big[v(t,x,a') - v(t,x,a)\big]^+ \lambda_\pi(da') \;=\; 0.$$

Under condition **(H**λ_π**)**(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 (3.11), and instead, we shall rely on viscosity solutions approach to derive the non dependence of $v(t,x,a)$ in $a \in \mathring{A}$. To this end, let us introduce the following first-order PDE:

$$-\big|D_a v(t,x,a)\big| \;=\; 0\;, \quad (t,x,a) \in [0,T) \times \mathbb{R}^d \times \mathring{A}. \tag{3.32}$$

Lemma 3.3 *Let assumptions **(H**λ_π**)**, **(HFC)**, **(HBC1)** and **(HBC2)** hold. The function v is a viscosity supersolution to* (3.32)*: for any $(t,x,a) \in [0,T) \times \mathbb{R}^d \times \mathring{A}$ 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*

$$-\big|D_a\varphi(t,x,a)\big| \;\geq\; 0, \quad i.e. \quad D_a\varphi(t,x,a) \;=\; 0.$$

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. [1] p. 91):

$$v \;=\; v_* \;=\; \liminf_{n\to\infty}{}_* v_n, \tag{3.33}$$

where

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

Let $(t,x,a) \in [0,T) \times \mathbb{R}^d \times \mathring{A}$, 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:

$$(v-\varphi)(t,x,a) \;=\; \text{strict} \min_{[0,T]\times\mathbb{R}^d\times\mathbb{R}^q}(v-\varphi)\,. \tag{3.34}$$

Up to a suitable negative perturbation of φ 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

$$(v_n-\varphi)(t_n,x_n,a_n) \;=\; \min_{[0,T]\times\mathbb{R}^d\times\mathbb{R}^q}(v_n-\varphi)\,. \tag{3.35}$$

From (3.33), (3.34), and (3.35), we then have, up to a subsequence:

$$(t_n,x_n,a_n,v_n(t_n,x_n,a_n)) \;\longrightarrow\; (t,x,a,v(t,x,a)) \quad \text{as} \;\; n\to\infty\,. \tag{3.36}$$

Now, from the viscosity supersolution property of v_n at (t_n, x_n, a_n) with the test function φ, we have by (3.35):

$$
\begin{aligned}
-\frac{\partial \varphi}{\partial t}(t_n, x_n, a_n) - \mathcal{L}^{a_n}\varphi(t_n, x_n, a_n) \\
-f(x_n, a_n, v_n(t_n, x_n, a_n), \sigma^{\intercal}(x_n, a_n)D_x\varphi(t_n, x_n, a_n), \mathcal{M}^{a_n}\varphi(t_n, x_n, a_n)) \\
-\int_A [\varphi(t_n, x_n, a') - \varphi(t_n, x_n, a_n)]\lambda_\pi(da') \\
-n\int_A [\varphi(t_n, x_n, a') - \varphi(t_n, x_n, a_n)]^+\lambda_\pi(da') \quad \geq \quad 0,
\end{aligned}
$$

which implies

$$
\begin{aligned}
& \int_A [\varphi(t_n, x_n, a') - \varphi(t_n, x_n, a_n)]^+\lambda_\pi(da') \\
\leq \quad & \frac{1}{n}\Big[-\frac{\partial \varphi}{\partial t}(t_n, x_n, a_n) - \mathcal{L}^{a_n}\varphi(t_n, x_n, a_n) \\
& \quad - f(x_n, a_n, v_n(t_n, x_n, a_n), \sigma^{\intercal}(x_n, a_n)D_x\varphi(t_n, x_n, a_n), \mathcal{M}^{a_n}\varphi(t_n, x_n, a_n)) \\
& \quad - \int_A [\varphi(t_n, x_n, a') - \varphi(t_n, x_n, a_n)]\lambda_\pi(da')\Big].
\end{aligned}
$$

Sending n to infinity, we get from (3.36), the continuity of coefficients b, σ, β and f, and the dominated convergence theorem:

$$
\int_A [\varphi(t, x, a') - \varphi(t, x, a)]^+\lambda_\pi(da') \quad = \quad 0\,.
$$

Under **(Hλ_π)**, this means that $\varphi(t, x, a) = \max_{a' \in A}\varphi(t, x, a')$. Since $a \in \mathring{A}$, we deduce that $D_a\varphi(t, x, a) = 0$. □

We notice that the PDE (3.32) 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 (3.32), 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 3.4 *Let assumptions **(Hλ_π)**, **(HFC)**, **(HBC1)** and **(HBC2)** hold. For any $(t, x) \in [0, T) \times \mathbb{R}^d$, the function $v(t, x, .)$ is a viscosity supersolution to*

$$
-\big|D_a v(t, x, a)\big| \quad = \quad 0\,, \qquad a \in \mathring{A},
$$

i.e. for any $a \in \mathring{A}$ 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: $-\big|D_a\varphi(a)\big| \geq 0$ (and so $= 0$).

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

$$
(v(t, x, .) - \varphi)(a) \quad = \quad \min_{\mathbb{R}^q}(v(t, x, .) - \varphi)\,. \tag{3.37}
$$

As usual, we may assume w.l.o.g. that this minimum is strict and that φ satisfies the growth condition $\sup_{a' \in \mathbb{R}^q}\frac{|\varphi(a')|}{1 + |a'|^p} < \infty$. 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

$$
\varphi^n(t', x', a') \quad = \quad \varphi(a') - n\big(|t' - t|^2 + |x' - x|^{2p}\big) - |a' - a|^{2p}
$$

for all $(t', x', a') \in [0,T] \times \mathbb{R}^d \times \mathbb{R}^q$. From the growth condition (3.16) on the lsc function v, and the growth condition on the continuous function φ, 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

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

In particular, we have

$$\begin{aligned} v(t,x,a) - \varphi(a) &= (v - \varphi^n)(t,x,a) \ \geq \ (v - \varphi^n)(t_n, x_n, a_n) \\ &= v(t_n, x_n, a_n) - \varphi(a_n) + n(|t_n - t|^2 + |x_n - x|^{2p}) + |a_n - a|^{2p} \\ &\geq v(t_n, x_n, a_n) - v(t, x, a_n) + v(t,x,a) - \varphi(a) \\ &\qquad + n(|t_n - t|^2 + |x_n - x|^{2p}) + |a_n - a|^{2p} \end{aligned} \tag{3.38}$$

by (3.37), which implies from the growth condition (3.16) on v:

$$n(|t_n - t|^2 + |x_n - x|^{2p}) + |a_n - a|^{2p} \quad \leq \quad C(1 + |x_n - x|^p + |a_n - a|^p).$$

Therefore, the sequences $\{n(|t_n - t|^2 + |x_n - x|^{2p})\}_n$ and $(|a - a_n|^{2p})_n$ are bounded and (up to a subsequence) we have: $(t_n, x_n, a_n) \longrightarrow (t, x, a_\infty)$ 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 (3.38), we obtain by sending n to infinity and since the minimum in (3.37) is strict, that $a_\infty = a$, and so:

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

On the other hand, from Lemma 3.3 applied to (t_n, x_n, a_n) with the test function φ^n, we have

$$0 \ = \ D_a \varphi^n(t_n, x_n, a_n) \quad = \quad D_a \varphi(a_n) - 2p(a_n - a)|a_n - a|^{2p-1},$$

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

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

Proposition 3.2 *Let assumptions* **(HA)**, ***(Hλ_π)***, **(HFC)**, **(HBC1)** *and* **(HBC2)** *hold. The function v does not depend on the variable a on $[0,T) \times \mathbb{R}^d \times \mathring{A}$:*

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

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

$$u_n(t,x,a) \quad = \quad \inf_{a' \in A} \left[v(t,x,a') + n|a - a'|^{2p} \right], \quad (t,x,a) \in [0,T] \times \mathbb{R}^d \times A.$$

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 $[0,T] \times \mathbb{R}^d \times A$. Indeed, fix some $(t,x,a) \in [0,T] \times \mathbb{R}^d \times A$. Since v is lsc, there exists a sequence $(a_n)_n$ valued in A such that

$$u_n(t,x,a) \;=\; v(t,x,a_n) + n|a-a_n|^{2p} \;,$$

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

$$u_n(t,x,a) \;\longrightarrow\; v(t,x,a) \quad \text{as} \;\; n \to \infty \,. \tag{3.39}$$

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

$$N(r,\delta) \;=\; \min\Big\{ n \in \mathbb{N} \;:\; n \;\geq\; \frac{2C_v(1+2^{2p-5}+r^p+2^{p-1}\max_{a\in A}|a|^p)}{\left(\frac{\delta}{2}\right)^{2p}} + C_v \Big\}$$

where C_v is the constant in the growth condition (3.16), and define the set $\mathring{A}_\delta$ by

$$\mathring{A}_\delta \;=\; \big\{ a \in A \;:\; d(a,\partial A) := \min_{a'\in\partial A}|a-a'| > \delta \big\} \,.$$

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

$$0 \;=\; (u_n(t,x,.) - \varphi)(a) \;\;=\;\; \min_{\mathbb{R}^q}(u_n(t,x,.) - \varphi) \;, \tag{3.40}$$

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

$$0 \;=\; (v(t,x,.) - \psi)(a_n) \;\;=\;\; \min_{\mathbb{R}^q}(v(t,x,.) - \psi) \;, \tag{3.41}$$

and

$$D_a\psi(a_n) \;\;=\;\; D_a\varphi(a). \tag{3.42}$$

To this end we proceed in two substeps.

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

$$\operatorname*{argmin}_{a'\in A}\big\{ v(t,x,a') + n|a'-a|^{2p} \big\} \;\;\subset\;\; \mathring{A} \,.$$

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

$$v(t,x,a_n) + n|a_n - a|^{2p} \;\;=\;\; \min_{a'\in A}\big[v(t,x,a') + n|a'-a|^{2p} \big] \,.$$

Then we have

$$v(t,x,a_n) + n|a_n - a|^{2p} \;\;\leq\;\; v(t,x,a),$$

and by (3.16), this gives

$$-C_v(1+|x|^p+2^{p-1}\max_{a\in A}|a|^p+2^{p-1}|a_n-a|^p)+n|a_n-a|^{2p} \quad\leq\quad C_v(1+|x|^p+|a|^p)\;.$$

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:

$$(n-C_v)|a_n-a|^{2p} \quad\leq\quad 2C_v(1+2^{2p-5}+|x|^p+2^{p-1}\max_{a\in A}|a|^p)\;.$$

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

$$|a_n-a| \quad\leq\quad \frac{\delta}{2}\;,$$

which shows that $a_n \in \mathring{A}$ since $a \in \mathring{A}_\delta$.

Substep 2.2. Fix $\delta > 0$, $(t,x,a) \in [0,T)\times\mathbb{R}^d\times\mathring{A}_\delta$, and $\varphi \in C^2(\mathbb{R}^q)$ satisfying (3.40). Let us then choose $a_n \in \text{argmin}\,\{v(t,x,a')+n|a'-a|^{2p} : a' \in A\}$, and define $\psi \in C^2(\mathbb{R}^q)$ by:

$$\psi(a') \quad=\quad \varphi(a+a'-a_n)-n|a_n-a|^{2p}, \quad a'\in\mathbb{R}^q.$$

It is clear that ψ satisfies (3.42). Moreover, we have by (3.40) and the inf-convolution definition of u_n:

$$\psi(a') \quad\leq\quad u_n(t,x,a+a'-a_n)-n|a_n-a|^{2p} \;\leq\; v(t,x,a')\;, \quad a'\in\mathbb{R}^q.$$

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

$$\psi(a_n) \quad=\quad \varphi(a)-n|a_n-a|^{2p} \;=\; u_n(t,x,a)-n|a_n-a|^{2p} \;=\; v(t,x,a_n)\;,$$

which shows (3.41).

Step 3. The function u_n does not depend locally on the variable a. From Step 2 and Lemma 3.4, 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

$$-\left|D_a u_n(t,x,a)\right| \quad=\quad 0, \quad a\in\mathring{A}_\delta, \tag{3.43}$$

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 $a \in \mathring{A}_\delta$:

$$u_n(t,x,a) \quad=\quad u_n(t,x,a'), \quad \forall a'\in B(a,\eta), \tag{3.44}$$

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 $\mathring{A}_\delta$. From Theorem 2.1.7 in [5], we deduce that $u_n(t,x,.)$ is locally Lipschitz continuous 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 2.1 (ii) in [1], and the viscosity supersolution property (3.43), we get that this relation (3.43) 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 (3.39) of u_n to v, and the relation (3.44), we get by sending n to infinity that for any $\delta > 0$ the function v satisfies: for any $(t, x, a) \in [0, T) \times \mathbb{R}^d \times \mathring{A}_\delta$

$$v(t, x, a) \quad = \quad v(t, x, a')$$

for all $\eta > 0$ such that $B(a, \eta) \subset \mathring{A}_\delta$ and all $a' \in B(a, \eta)$. Then by sending δ 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 $[0, T) \times \mathbb{R}^d \times \mathring{A}$.

□

3.4 Viscosity properties of the minimal solution to the BSDE with A-nonpositive jumps

From Proposition 3.2, we can define by misuse of notation the function v on $[0, T) \times \mathbb{R}^d$ by

$$v(t, x) \quad = \quad v(t, x, a), \quad (t, x) \in [0, T) \times \mathbb{R}^d, \tag{3.45}$$

for any $a \in \mathring{A}$. Moreover, by the growth condition (3.16), we have for some $p \geq 2$:

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} \frac{|v(t, x)|}{1 + |x|^p} \quad < \quad \infty. \tag{3.46}$$

The aim of this section is to prove that the function v is a viscosity solution to (3.7)-(3.8).

Proof of the viscosity supersolution property to (3.7)**.** We first notice from (3.33) and (3.45) that v is lsc and

$$v(t, x) \; = \; v_*(t, x) \; = \; \liminf_{n \to \infty} {}_* v_n(t, x, a) \tag{3.47}$$

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

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

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

$$\varphi^\varepsilon(t', x', a') \quad = \quad \varphi(t', x') - \varepsilon\big(|t' - t|^2 + |x' - x|^{2p} + |a' - a|^{2p}\big),$$

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

$$(v - \varphi^\varepsilon)(t, x, a) \quad = \quad \text{strict} \min_{[0,T] \times \mathbb{R}^d \times \mathbb{R}^q} (v - \varphi^\varepsilon). \tag{3.48}$$

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

$$(v_n - \varphi^\varepsilon)(t_n, x_n, a_n) \;\; = \;\; \min_{[0,T]\times\mathbb{R}^d\times\mathbb{R}^q} (v_n - \varphi^\varepsilon) \; . \tag{3.49}$$

From (3.47) and (3.48), we obtain by standard arguments that up to a subsequence:

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

Now from the viscosity supersolution property of v_n at (t_n, x_n, a_n) with the test function φ^ε, we have

$$\begin{aligned}
-\frac{\partial \varphi^\varepsilon}{\partial t}(t_n, x_n, a_n) - \mathcal{L}^{a_n} \varphi^\varepsilon(t_n, x_n, a_n) \\
-f\big(x_n, a_n, v_n(t_n, x_n, a_n), \sigma^\intercal(x_n, a_n) D_x \varphi^\varepsilon(t_n, x_n, a_n), \mathcal{M}^{a_n} \varphi^\varepsilon(t_n, x_n, a_n)\big) \\
- \int_A [\varphi^\varepsilon(t_n, x_n, a') - \varphi^\varepsilon(t_n, x_n, a_n)] \lambda_\pi(da') \\
- n \int_A [\varphi^\varepsilon(t_n, x_n, a') - \varphi^\varepsilon(t_n, x_n, a_n)]^+ \lambda_\pi(da') \;\; & \geq \;\; 0 \; .
\end{aligned}$$

Sending n to infinity in the above inequality, we get from the definition of φ^ε and the dominated convergence Theorem:

$$\begin{aligned}
- \frac{\partial \varphi^\varepsilon}{\partial t}(t, x, a) - \mathcal{L}^a \varphi^\varepsilon(t, x, a) \\
-f\big(x, a, v(t, x), \sigma^\intercal(x, a) D_x \varphi^\varepsilon(t, x, a), \mathcal{M}^a \varphi^\varepsilon(t, x, a)\big) \\
+\varepsilon \int_A |a' - a|^{2p} \lambda_\pi(da') \;\; & \geq \;\; 0 \; .
\end{aligned} \tag{3.50}$$

Sending ε to zero, and since $\varphi^\varepsilon(t, x, a) = \varphi(t, x)$, we get

$$-\frac{\partial \varphi}{\partial t}(t, x) - \mathcal{L}^a \varphi(t, x) - f\big(x, a, v(t, x), \sigma^\intercal(x, a) D_x \varphi(t, x), \mathcal{M}^a \varphi(t, x)\big) \;\; \geq \;\; 0 \; .$$

Since a is arbitrarily chosen in $\mathring{A}$, we get from **(HA)** and the continuity of the coefficients b, σ, γ and f in the variable a

$$-\frac{\partial \varphi}{\partial t}(t, x) - \sup_{a \in A} \Big[\mathcal{L}^a \varphi(t, x) + f\big(x, a, v(t, x), \sigma^\intercal(x, a) D_x \varphi(t, x), \mathcal{M}^a \varphi(t, x)\big) \Big] \;\; \geq \;\; 0 \; ,$$

which is the viscosity supersolution property. □

Proof of the viscosity subsolution property to (3.7)**.** Since v is the pointwise limit of the nondecreasing sequence of continuous functions (v_n), and recalling (3.45), we have by [1] p. 91:

$$v^*(t, x) \;\; = \;\; \limsup_{n \to \infty}{}^* v_n(t, x, a) \tag{3.51}$$

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

$$\limsup_{n \to \infty}{}^* v_n(t, x, a) \;\; := \;\; \limsup_{\substack{n \to \infty \\ (t', x', a') \to (t, x, a) \\ t' < T, \, a' \in \mathring{A}}} v_n(t', x', a').$$

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

$$(v^* - \varphi)(t,x) \quad = \quad \max_{[0,T]\times\mathbb{R}^d} (v^* - \varphi) \; . \tag{3.52}$$

We may assume w.l.o.g. that this maximum is strict and that φ satisfies

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \frac{|\varphi(t,x)|}{1+|x|^p} \quad < \quad \infty \; . \tag{3.53}$$

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

$$(t_n, x_n, a_n, v_n(t_n, x_n, a_n)) \quad \longrightarrow \quad (t, x, a, v^*(t,x)) \qquad \text{as} \qquad n \to \infty. \tag{3.54}$$

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

$$\varphi_n(t', x', a') \quad = \quad \varphi(t', x') + n\Big(\frac{d_{A_{\eta_n}}(a')}{\eta_n} \wedge 1 + |t' - t_n|^2 + |x' - x_n|^{2p}\Big)$$

where A_{η_n} is defined by (3.30) for $\varepsilon = \eta_n$ and $(\eta_n)_n$ is a positive sequence converging to 0 s.t. (such sequence exists by $(\mathbf{H}\lambda_\pi)$(ii)):

$$n^2 \lambda_\pi(A \setminus A_{\eta_n}) \quad \longrightarrow \quad 0 \qquad \text{as} \qquad n \to \infty \; . \tag{3.55}$$

From the growth conditions (3.46) and (3.53) on v and φ, we can find a sequence $(\bar{t}_n, \bar{x}_n, \bar{a}_n)$ in $[0,T] \times \mathbb{R}^d \times A$ such that

$$(v_n - \varphi_n)(\bar{t}_n, \bar{x}_n, \bar{a}_n) \quad = \quad \max_{[0,T]\times\mathbb{R}^d\times A} (v_n - \varphi_n) \; , \qquad n \geq 1 \; . \tag{3.56}$$

Using (3.51) and (3.52), we obtain by standard arguments that up to a subsequence

$$n\Big(\frac{1}{\eta_n} d_{A_{\eta_n}}(\bar{a}_n) + |\bar{t}_n - t_n|^p + |\bar{x}_n - x_n|^{2p}\Big) \quad \longrightarrow \quad 0 \qquad \text{as} \;\; n \to \infty \; , \tag{3.57}$$

and

$$v_n(\bar{t}_n, \bar{x}_n, \bar{a}_n) \quad \longrightarrow \quad v^*(t,x) \qquad \text{as} \;\; n \to \infty \; .$$

We deduce from (3.57) and (3.54) that, up to a subsequence:

$$(\bar{t}_n, \bar{x}_n, \bar{a}_n) \quad \longrightarrow \quad (t, x, \bar{a}), \qquad \text{as} \;\; n \to \infty \; . \tag{3.58}$$

for some $\bar{a} \in A$. Moreover, for n large enough, we have $\bar{a}_n \in \mathring{A}$. Indeed, suppose that, up to a subsequence, $\bar{a}_n \in \partial A$ for $n \geq 1$. Then we have $\frac{1}{\eta_n} d_{A_{\eta_n}}(\bar{a}_n) \geq 1$, which contradicts (3.57). Now, from the viscosity subsolution property of v_n at $(\bar{t}_n, \bar{x}_n, \bar{a}_n)$ with the test function φ_n satisfying (3.56), Lemma 3.2, and since $\bar{a}_n \in \mathring{A}$, we have:

$$\begin{aligned}
&- \frac{\partial \varphi_n}{\partial t}(\bar{t}_n, \bar{x}_n, \bar{a}_n) - \mathcal{L}^{\bar{a}_n} \varphi_n(\bar{t}_n, \bar{x}_n, \bar{a}_n) \\
-f(\bar{x}_n, \bar{a}_n, v_n(\bar{t}_n, \bar{x}_n, \bar{a}_n), &\sigma^\intercal(\bar{x}_n, \bar{a}_n) D_x \varphi(\bar{t}_n, \bar{x}_n), \mathcal{M}^{\bar{a}_n} \varphi(\bar{t}_n, \bar{x}_n, \bar{a}_n)) \\
&-(n+1)n \int_A \Big(\frac{d_{A_{\eta_n}}(a')}{\eta_n} \wedge 1\Big) \lambda_\pi(da') \quad \leq \quad 0 \; ,
\end{aligned} \tag{3.59}$$

for all $n \geq 1$. From (3.55) we get

$$(n+1)n \int_A \Big(\frac{d_{A_{\eta_n}}(a')}{\eta_n} \wedge 1 \Big) \lambda_\pi(da') \quad \longrightarrow \quad 0 \quad \text{ as } n \to \infty \tag{3.60}$$

Sending n to infinity into (3.59), and using (3.51), (3.58) and (3.60), we get

$$-\frac{\partial \varphi}{\partial t}(t,x) - \mathcal{L}^{\bar{a}}\varphi(t,x) - f(x, \bar{a}, v^*(t,x), \sigma^\intercal(x,\bar{a}) D_x\varphi(t,x), \mathcal{M}^{\bar{a}}\varphi(t,x)) \quad \leq \quad 0 \ .$$

Since $\bar{a} \in A$, this gives

$$-\frac{\partial \varphi}{\partial t}(t,x) - \sup_{a \in A} \Big[\mathcal{L}^{a}\varphi(t,x) + f\big(x, a, v^*(t,x), \sigma^\intercal(x,a) D_x\varphi(t,x), \mathcal{M}^{a}\varphi(t,x)\big) \Big] \quad \leq \quad 0 \ ,$$

which is the viscosity subsolution property. □

Proof of the viscosity supersolution property to (3.8)**.** Let $(x,a) \in \mathbb{R}^d \times \mathring{A}$. From (3.47), 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

$$(t_n, x_n, a_n, v_n(t_n, x_n, a_n)) \quad \longrightarrow \quad (T, x, a, v_*(T,x)) \qquad \text{as} \quad n \to \infty \ .$$

The sequence of continuous functions $(v_n)_n$ being nondecreasing and $v_n(T,.) = g$ we have

$$v_*(T,x) \quad \geq \quad \lim_{n \to \infty} v_1(t_n, x_n, a_n) \ = \ g(x,a) \ .$$

Since a is arbitrarily chosen in $\mathring{A}$, we deduce that $v_*(T,x) \geq \sup_{a \in \mathring{A}} g(x,a) = \sup_{a \in A} g(x,a)$ by **(H**A**)** and continuity of g in a. □

Proof of the viscosity subsolution property to (3.8)**.** Let $x \in \mathbb{R}^d$. Then we can find by (3.51) a sequence $(t_n, x_n, a_n)_n$ in $[0,T) \times \mathbb{R}^d \times \mathring{A}$ such that

$$(t_n, x_n, v_n(t_n, x_n, a_n)) \quad \to \quad (T, x, v^*(T,x)) \ , \quad \text{ as } \quad n \to \infty \ . \tag{3.61}$$

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

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

for all $(t,x) \in [0,T) \times \mathbb{R}^d$. From **(HFC)**, **(HBC1)** and **(HBC2')**, we see that h is a continuous viscosity supersolution to (3.17)-(3.18), on $[T-\eta, T] \times \bar{B}(x,\eta)$ for η small enough. We can then apply Theorem 3.5 in [2] which gives that

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

for all $n \geq 0$. By applying the above inequality at (t_n, x_n, a_n), and sending n to infinity, together with (3.61), we get the required result. □

4 Conclusion

We introduced a class of BSDEs with partially nonpositive jumps and showed how the minimal solution is related to a fully nonlinear IPDE of HJB type, when considering a Markovian framework with forward regime switching jump-diffusion process. Such BSDE representation can be extended to cover also the case of Hamilton-Jacobi-Bellman-Isaacs equation arising in controller/stopper differential games, see [6]. It is also extended to the non Markovian context in [12]. From a numerical application viewpoint, our BSDE representation leads to original probabilistic approximation scheme for the resolution in high dimension of fully nonlinear HJB equations, as recently investigated in [14] and [15]. We believe that this opens new perspectives for dealing with more general non linear equations and control problems, like for instance mean field games, or control of McKean-Vlasov equations.

References

[1] Barles G. (1994) : *Solutions de viscosité des équations d'Hamilton-Jacobi*, Mathématiques et Applications, Springer Verlag.

[2] Barles G., Buckdahn R. and E. Pardoux (1997) : "Backward stochastic differential equations and integral-partial differential equations", *Stochastics and Stochastics Reports*, **60**, 57-83.

[3] Barles G. and C. Imbert (2008): " Second-Order Elliptic Integro-Differential Equations: Viscosity Solutions Theory Revisited", *Ann. Inst. H. Poincaré, Anal. Non Linéaire*, **25**, 567-585.

[4] Bouchard B. and N. Touzi (2004): "Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations", *Stochastic processes and Their Applications*, **111**, 174-206.

[5] Cannarsa P. and Sinestrari C. (2004): *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control*, Progress in Nonlinear Differential Equations and Their Applications, **58**, Birkhäuser.

[6] Choukroun S., Cosso A. and H. Pham (2013): "Reflected BSDEs with nonpositive jumps, and controller-and-stopper games", preprint arXiv:1308.5511

[7] Ekren I., Touzi N. and J. Zhang (2013): "Viscosity Solutions of Fully Nonlinear Parabolic Path Dependent PDEs: Part I", Preprint.

[8] El Karoui N. (1981): *Les aspects probabilistes du contrôle stochastique*, Lect. Notes in Mathematics 876, Ecole d'Eté de Saint Flour 1979.

[9] El Karoui N, Peng S. and M. C. Quenez (1997): "Backward stochastic differential equations in finance", *Mathematical Finance*, **7**, 1-71.

[10] Essaky E. H. (2008) : "Reflected backward stochastic differential equation with jumps and RCLL obstacle", *Bull. Sci. math.*, **132**, 690-710.

[11] Friedman, A (1975): *Stochastic Differential Equations and Applications, Vol. 1*, Probability and Mathematical Statistics, Vol. 28. Academic Press, New York-London.

[12] Fuhrman M. and H. Pham (2013): "Dual and backward SDE representation for optimal control of non-Markovian SDEs", preprint arXiv:1310.6943.

[13] Henry-Labordère P. (2012): "Counterparty Risk Valuation: A Marked Branching Diffusion Approach", preprint arxiv 1203.2369v1.

[14] Kharroubi I., Langrené N. and H. Pham (2013a): "Discrete time approximation of fully nonlinear HJB equations via BSDEs with nonpositive jumps", preprint arXiv:1311.4505

[15] Kharroubi I., Langrené N. and H. Pham (2013b): "Numerical algorithm for fully nonlinear HJB equations: an approach by control randomization", preprint arXiv:1311.4503

[16] Kharroubi I., Ma J., Pham H. and J. Zhang (2010): " Backward SDEs with constrained jumps and quasi-variational inequalities", *Annals of Probability*, **38**, 794-840.

[17] Ma, J. and J. Zhang (2005): "Representations and regularities for solutions to BSDEs with reflections", *Stochastic processes and their applications*, **115**, 539- 569.

[18] Oksendal B. and A. Sulem (2007): Applied Stochastic Control of Jump Diffusions, 2-nd edition, Universitext, Springer.

[19] Pardoux E. and S. Peng (1990) : "Adapted solution of a backward stochastic differential equation", *Systems and Control Letters*, **14**, 55-61.

[20] Pardoux E. and S. Peng (1992) : "Backward stochastic differential equation and quasilinear parabolic partial differential equations", in *Stochastic partial differential equations and their applications*, B. Rozovskiin R. Sowers (eds), Lect. Notes in Cont. Inf. Sci., **176**, 200-217.

[21] Peng S. (2000): "Monotonic limit theorem for BSDEs and non-linear Doob-Meyer decomposition", *Probab. Theory and Rel. Fields*, 16 (3), 225-234.

[22] Peng S. (2006): "G-expectation, G-Brownian motion and related stochastic calculus of Itô type", Proceedings of 2005, Abel symposium, Springer.

[23] Pham H. (1998): "Optimal stopping of controlled jump diffusion processes: a viscosity solutions approach", *Journal of Mathematical Systems Estimation and Control*, **8**, 27 pp (electronic).

[24] Pham H. (2009): *Continuous-time stochastic control and optimization with financial applications*, Springer, Series SMAP, Vol, 61.

[25] Protter P. and K. Shimbo (2008): "No arbitrage and general semimartingale", in *Festschrift for Thomas Kurtz.*

[26] Royer M. (2006): "Backward stochastic differential equations with jumps and related nonlinear expectations", *Stochastic Processes and their Applications*, **116**, 1358-1376.

[27] Soner M., Touzi N. and J. Zhang (2012): "The wellposedness of second order backward SDEs", *Probability Theory and Related Fields*, **153**, 149-190.

[28] Tang S. and X. Li (1994): "Necessary conditions for optimal control of stochastic systems with jumps", *SIAM J. Control and Optimization*, **32**, 1447-1475.