Long time asymptotics for fully nonlinear Bellman equations: a Backward SDE approach

Andrea Cosso, Marco Fuhrman, Huyen Pham

Long time asymptotics for fully nonlinear Bellman equations: a Backward SDE approach

Andrea COSSO* Marco FUHRMAN† Huyên PHAM‡

October 4, 2014

Abstract

We study the large time behavior of solutions to fully nonlinear parabolic equations of Hamilton-Jacobi-Bellman type arising typically in stochastic control theory with control both on drift and diffusion coefficients. We prove that, as time horizon goes to infinity, the long run average solution is characterized by a nonlinear ergodic equation. Our results hold under dissipativity conditions, and without any nondegeneracy assumption on the diffusion term. Our approach uses mainly probabilistic arguments relying on new backward SDE representation for nonlinear parabolic, elliptic and ergodic equations.

MSC Classification: 60H30, 60J60, 49J20.

Keywords: Hamilton-Jacobi-Bellman type equation, stochastic control, large time behavior, backward SDE, ergodic type Bellman equation.

*Laboratoire de Probabilités et Modèles Aléatoires, CNRS, UMR 7599, Université Paris Diderot, cosso at math.univ-paris-diderot.fr

†Dipartimento di Matematica, Politecnico di Milano, marco.fuhrman at polimi.it

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

1 Introduction

Let us consider the fully nonlinear parabolic equation of the form

$$\frac{\partial v}{\partial T} - \sup_{a\in A}\big[\mathcal{L}^a v + f\big(x,a,\frac{v}{T+1}\big)\big] \;=\; 0, \quad \text{on } (0,\infty)\times\mathbb{R}^d, \tag{1.1}$$

with $v(0,.) = h$ on $\mathbb{R}^d$, where $\mathcal{L}^a$ is the second-order differential operator:

$$\mathcal{L}^a v \;=\; b(x,a).D_x v + \frac{1}{2}\mathrm{tr}(\sigma\sigma^\intercal(x,a)D_x^2 v).$$

Here A is some Borel subset of $\mathbb{R}^q$, b,σ are continuous functions on $\mathbb{R}^d\times\mathbb{R}^q$, and $f = f(x,a,y)$ is a measurable function on $\mathbb{R}^d\times\mathbb{R}^q\times\mathbb{R}$ satisfying some conditions to be specified later on. When the generator $f = f(x,a)$ does not depend on y, equation (1.1) is the dynamic programming equation, also called Hamilton-Jacobi-Bellman (HJB) equation, associated to the stochastic control problem:

$$v(T,x) \;:=\; \sup_{\alpha\in\mathcal{A}}\mathbb{E}_x\Big[\int_0^T f(X_t^\alpha,\alpha_t)dt + h(X_T^\alpha)\Big],$$

where X^α is the controlled diffusion process

$$dX_t^\alpha \;=\; b(X_t^\alpha,\alpha_t)dt + \sigma(X_t^\alpha,\alpha_t)dW_t, \tag{1.2}$$

driven by a d-dimensional Brownian motion W on a probability space $(\Omega,\mathcal{F},\mathbb{P})$ equipped with the natural filtration of W, and given a control $\alpha\in\mathcal{A}$, i.e., an A-valued adapted process. In the general case $f = f(x,a,y)$, we shall see that under suitable conditions, there exists a unique viscosity solution $v = v(T,x)$ to the generalized parabolic HJB equation (1.1), and our aim is to investigate the large time behavior of $v(T,.)$ as T goes to infinity. It turns out that this asymptotic problem is related to the generalized *ergodic* HJB equation:

$$\lambda - \sup_{a\in A}\big[\mathcal{L}^a\phi + f(x,a,\lambda)\big] \;=\; 0, \quad \text{on } \mathbb{R}^d. \tag{1.3}$$

Asymptotics for stochastic control and related HJB equation have been studied in various settings by many authors since the works [4] and [2]. In the PDE literature, we refer for instance to [3] in a periodic setting, [28] under Dirichlet conditions, or [12] in the whole space. In these cited papers, the HJB equation is semi-linear, i.e., the nonlinearity appears only in the first order derivative. Recently, by combining PDE and stochastic analysis arguments, the papers [15], [16] and [24] studied large time behavior of semi-linear HJB equations with quadratic nonlinearity in gradients. We would like also to point out the recent paper [14], which studied large time behavior of solutions to semi-linear HJB equations by a probabilistic approach relying on ergodic BSDE introduced in [11]. Interestingly, the authors are able to prove in their context a rate of convergence for the solution to the parabolic equation towards the ergodic equation under weak dissipativity conditions. Long time asymptotics of solutions to HJB equations has been also considered in the context of risk-sensitive stochastic control and utility maximization problem, see e.g. [10] and [21].

Our motivation is to develop a systematic study applicable to a large class of fully nonlinear HJB type equation, and to give natural conditions on the dynamics of the control system ensuring ergodicity. The principal novelty of this paper is to consider control on both drift and diffusion coefficients $b(x,a)$, $\sigma(x,a)$, possibly degenerate, and satisfying dissipativity conditions, instead of periodicity condition. In this case, we do not have in general smooth solution to the HJB equation. Another original feature of our framework is the dependence of $f(x,a,y)$ on y, which occurs for example in stochastic control with recursive utility functions. Our first main result is to prove the existence of a viscosity solution pair $(\lambda,\phi) \in \mathbb{R} \times C(\mathbb{R}^d)$, with ϕ Lipschitz, to the ergodic fully nonlinear equation (1.3). We adopt the following approach. We consider the sequence of fully nonlinear elliptic HJB equation for $\beta > 0$:

$$\beta v^\beta - \sup_{a \in A} \big[\mathcal{L}^a v^\beta + f\big(x,a,\beta v^\beta\big) \big] \;=\; 0, \quad \text{on } \mathbb{R}^d, \tag{1.4}$$

and obtain the existence and uniqueness of a solution v^β to (1.4) by combining analytical and probabilistic methods. More precisely, following the randomization approach of [19] for representing parabolic HJB equations, we introduce a class of Backward Stochastic Differential Equations (BSDEs) with nonpositive jumps over an infinite horizon, and supported by a forward regime switching process (X,I) where

$$dX_t \;=\; b(X_t,I_t)dt + \sigma(X_t,I_t)dW_t, \tag{1.5}$$

and I is a pure jump process valued in A. The minimal solution Y^β to this class of elliptic BSDEs is shown to exist and to provide the unique (viscosity) solution v^β to (1.4), and the key point is to derive uniform Lipschitz estimate for the sequence $(v^\beta)_\beta$. This is achieved by ergodicity properties on the forward process X, and suitable estimation on the minimal solution Y^β. Then, by standard analytical approximation procedures (when β goes to zero) as in [15] or [11], we obtain the existence of a pair (λ,ϕ) solution to (1.3). Moreover, the function ϕ admits a probabilistic representation in terms of a new class of BSDEs, namely ergodic BSDEs with nonpositive jumps. Ergodic BSDEs have been introduced in [11] and then in [23], and related to optimal control on the drift of diffusions. We extend this connection to the case of controlled diffusion coefficient by imposing a nonpositive jump constraint on the ergodic BSDE.

Next, our main theorem is to prove that for any solution (λ,ϕ) to (1.3), we have the convergence of the solution to the parabolic generalized HJB equation (1.1):

$$\frac{v(T,.)}{T} \;\longrightarrow\; \lambda, \quad \text{in } C(\mathbb{R}^d), \quad \text{as } T \to \infty. \tag{1.6}$$

Here, convergence "in $C(\mathbb{R}^d)$" stands for locally uniform convergence in $\mathbb{R}^d$. This shows as a byproduct that λ in (1.3) is unique. The main difficulty with respect to the semi-linear HJB case is that we do not have in general a smooth solution and an optimal control for the finite horizon and ergodic stochastic control, and the classical arguments as in [15] or [16] do not work anymore. Moreover, when $f = f(x,a,y)$ depends also on y, we do not even have a stochastic control representation of the function v. Our arguments for proving (1.6) rely on the BSDE representation of solution to (1.1) and (1.3), corresponding comparison theorems, and dual representation of such BSDEs in terms of equivalent probability measures introduced in (2.3). Furthermore, we can strengthen the convergence result (1.6) by a verification theorem: under the condition that the ergodic equation (1.3) admits a classical component solution ϕ, we have

$$v(T,.) - \big(\lambda T + \phi\big) \quad \longrightarrow \quad c, \qquad \text{in } \; C(\mathbb{R}^d), \quad \text{as } \; T \to \infty, \tag{1.7}$$

for some constant c.

The paper is organized as follows. Section 2 introduces some notations, formulates the dissipativity conditions on b, σ, and assumptions on f. We then state ergodicity properties on the regime switching process (X, I) in (1.5) as well as on the controlled diffusion process X^α in (1.2). In Section 3, we prove the existence and uniqueness of a solution to the fully nonlinear elliptic HJB equation (1.4) and its relation to BSDE with nonpositive jumps over an infinite horizon. Section 4 is concerned with the ergodic equation (1.3) and its probabilistic representation in terms of ergodic BSDE with nonpositive jumps. Convergence results (1.6) and (1.7) are studied in Section 5. We collect in the Appendix some proofs and technical estimates needed in the paper.

2 Ergodicity properties

2.1 Notations 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 Poisson random measure μ on $\mathbb{R}_+ \times A$, where A is a compact subset of $\mathbb{R}^q$, endowed with its Borel σ-field $\mathcal{B}(A)$. We assume that the random measure μ has the intensity measure $\vartheta(da)dt$ for some finite measure ϑ on $(A, \mathcal{B}(A))$. We set $\tilde{\mu}(dt, da) = \mu(dt, da) - \vartheta(da)dt$ 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 also denote, for any $T > 0$, $\mathcal{P}_T$ the σ-field of $\mathbb{F}$-predictable subsets of $[0, T] \times \Omega$. Let us introduce some additional notations. We denote by:

- $\mathbf{L^p}(\mathcal{F}_t)$, $p \geq 1$, $t \geq 0$, the set of $\mathcal{F}_t$-measurable random variables X such that $\mathbb{E}[|X|^p] < \infty$.

- $\mathbf{S^2(t, T)}$, $0 \leq t < T < \infty$, the set of real-valued càdlàg adapted processes $Y = (Y_s)_{t\leq s\leq T}$ satisfying
$$\mathbb{E}\Big[\sup_{t\leq s\leq T} |Y_s|^2\Big] \; < \; \infty.$$
We also define $\mathbf{S^2_{\text{loc}}} := \cap_{T>0}\mathbf{S^2(0, T)}$.

- $\mathbf{L^p(W; t, T)}$, $p \geq 1$, $0 \leq t < T < \infty$, the set of $\mathbb{R}^d$-valued predictable processes $Z = (Z_s)_{t\leq s\leq T}$ such that
$$\mathbb{E}\bigg[\Big(\int_t^T |Z_s|^2 ds\Big)^{\frac{p}{2}}\bigg] < \infty.$$
We also define $\mathbf{L^p_{\text{loc}}(W)} := \cap_{T>0}\mathbf{L^p(W; 0, T)}$.

- $\mathbf{L^p}(\tilde{\mu};\mathbf{t},\mathbf{T})$, $p \geq 1$, $0 \leq t < T < \infty$, the set of $\mathcal{P}_T \otimes \mathcal{B}(A)$-measurable maps $U\colon [t,T] \times \Omega \times A \to \mathbb{R}$ such that

$$\mathbb{E}\Big[\int_t^T \Big(\int_A |U_s(a)|^2 \vartheta(da)\Big)^{\frac{p}{2}} ds\Big] \ < \ \infty.$$

  We also define $\mathbf{L^p_{loc}}(\tilde{\mu}) := \cap_{T>0} \mathbf{L^p}(\tilde{\mu};\mathbf{0},\mathbf{T})$.

- $\mathbf{K^2}(\mathbf{t},\mathbf{T})$, $0 \leq t < T < \infty$, the set of nondecreasing càdlàg predictable processes $K = (K_s)_{t \leq s \leq T}$ such that $\mathbb{E}[|K_T|^2] < \infty$ and $K_t = 0$. We also define $\mathbf{K^2_{loc}} := \cap_{T>0}\mathbf{K^2}(\mathbf{0},\mathbf{T})$.

We are given some continuous functions $b\colon \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^d$, $\sigma\colon \mathbb{R}^d \times \mathbb{R}^q \to \mathbb{R}^{d\times d}$, and consider the forward regime switching process (X,I) governed by the stochastic differential equation in $\mathbb{R}^d \times \mathbb{R}^q$:

$$\begin{cases} dX_t = b(X_t,I_t)\,dt + \sigma(X_t,I_t)\,dW_t, \\ dI_t = \int_A (a - I_{t-})\,\mu(dt,da). \end{cases} \tag{2.1}$$

We note that the fact that σ is a square matrix does not involve any loss of generality, since we are not going to assume any nondegeneracy condition. In particular, some rows or columns of σ may be equal to zero. In the following we use the notation $M^\intercal$ for the transpose of any matrix M, and $\|M\|^2 = \mathrm{tr}(MM^\intercal)$ for the Hilbert-Schmidt norm. We shall make the following assumptions on the coefficients b and σ.

(H1)

(i) There exists a positive constant L_1 such that for all $x, x' \in \mathbb{R}^d$, $a, a' \in \mathbb{R}^q$,

$$|b(x,a) - b(x',a')| + \|\sigma(x,a) - \sigma(x',a')\| \quad \leq \quad L_1\,(|x-x'| + |a-a'|).$$

(ii) There exists a constant $\gamma > 0$ such that for all $x, x' \in \mathbb{R}^d$, $a \in A$,

$$(x-x').(b(x,a) - b(x',a)) + \frac{1}{2}\|\sigma(x,a) - \sigma(x',a)\|^2 \quad \leq \quad -\gamma\,|x-x'|^2. \tag{2.2}$$

It is well-known that under **(H1)**(i), there exists a unique solution $(X_t^{x,a}, I_t^a)_{t\geq 0}$ to (2.1) starting from $(x,a) \in \mathbb{R}^d \times \mathbb{R}^q$ at time $t = 0$. Notice that when $a \in A$, then $I_t^a \in A$ for all $t \geq 0$. Condition **(H1)**(ii) is called *dissipativity condition* and will ensure the ergodicity of the process X, as stated in the next paragraph.

Example 2.1 Let $b(x,a) = B(a)x + D(a)$, $\sigma(x,a) = \Sigma(a)$ for some vector valued Lipschitz function D, and matrix valued Lipschitz functions B, Σ on A, such that B is uniformly stable:

$$x.B(a)x \quad \leq \quad -\gamma|x|^2, \quad \forall x \in \mathbb{R}^d,\ a \in A.$$

In this case, **(H1)** is satisfied, and this example corresponds to a controlled Ornstein-Uhlenbeck process with uncertain mean-reversion and volatility. □

We also consider some real-valued continuous function f on $\mathbb{R}^d \times \mathbb{R}^q \times \mathbb{R}$ satisfying the following assumption:

(H2)

(i) There exists a positive constant L_2 such that for all $x, x' \in \mathbb{R}^d$, $a, a' \in \mathbb{R}^q$, $y, y' \in \mathbb{R}$,

$$|f(x,a,y) - f(x',a',y')| \leq L_2\,(|x-x'| + |a-a'| + |y-y'|).$$

(ii) The function $y \in \mathbb{R} \longmapsto f(x,a,y)$ is nonincreasing for all $(x,a) \in \mathbb{R}^d \times \mathbb{R}^q$.

We end this paragraph of notations by introducing the following set of probability measures, which shall play an important role in the sequel for establishing estimates. Let $\mathcal{V}_n$ be the set of $\mathcal{P} \otimes \mathcal{B}(A)$-measurable maps valued in $[1, n+1]$, $\mathcal{V} = \cup_{n \in \mathbb{N}} \mathcal{V}_n$, and consider for $\nu \in \mathcal{V}$, the probability measure $\mathbb{P}^\nu$ equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{F}_T)$, for any $T > 0$, with Radon-Nikodym density:

$$\frac{d\mathbb{P}^\nu}{d\mathbb{P}}\Big|_{\mathcal{F}_t} = \zeta_t^\nu := \mathcal{E}_t\Big(\int_0^{\cdot} \int_A (\nu_s(a) - 1)\tilde{\mu}(ds, da)\Big), \tag{2.3}$$

for $0 \leq t \leq T$, where $\mathcal{E}(.)$ is the Doléans-Dade exponential. Actually, since $\nu \in \mathcal{V}$ is essentially bounded, it is shown in Lemma 2.4 in [19] that $(\zeta^\nu)_{0 \leq t \leq T}$ is a uniformly integrable $\mathbb{P}$-martingale, with $\zeta_T^\nu \in \mathbf{L^2}(\mathcal{F}_T)$, for any $T > 0$, and so it defines a probability measure $\mathbb{P}^\nu$ via (2.3). We shall denote by $\mathbb{E}^\nu$ the expectation under $\mathbb{P}^\nu$. Moreover, by Girsanov theorem, the compensator of μ under $\mathbb{P}^\nu$ is $\nu_t(a)\vartheta(da)dt$, while W remains a Brownian motion independent of μ under $\mathbb{P}^\nu$. We denote by $\tilde{\mu}^\nu(dt, da) = \mu(dt, da) - \nu_t(a)\vartheta(da)dt$ the compensated martingale measure of μ under $\mathbb{P}^\nu$.

2.2 Ergodicity

We now use the dissipativity condition in **(H1)**(ii) to state moment estimates and stability results on the component X of (2.1).

Lemma 2.1 *Let Assumption* **(H1)** *hold.*
(i) *There exists a positive constant $C = C_{b,\sigma}$ depending only on b, σ such that for all $x \in \mathbb{R}^d$ and $a \in A$,*

$$\sup_{t \geq 0} \sup_{\nu \in \mathcal{V}} \mathbb{E}^\nu\big[|X_t^{x,a}|^2\big] \leq C(1 + |x|^2), \tag{2.4}$$

(ii) *For all $t \geq 0$, $x, x' \in \mathbb{R}^d$, $a \in A$,*

$$\sup_{\nu \in \mathcal{V}} \mathbb{E}^\nu\big[|X_t^{x,a} - X_t^{x',a}|^2\big] \leq |x - x'|^2 e^{-2\gamma t}. \tag{2.5}$$

The proof relies on rather standard arguments based on Itô's formula and Gronwall's lemma, and is reported in the Appendix.

Remark 2.1 We shall need the following generalization of estimate (2.4), for all $x \in \mathbb{R}^d$ and $a \in A$,

$$\sup_{s \geq t} \sup_{\nu \in \mathcal{V}} \mathbb{E}^\nu \big[|X_s^{x,a}|^2 \big| \mathcal{F}_t \big] \quad \leq \quad C(1 + |x|^2), \tag{2.6}$$

which is valid with the same constant $C = C_{b,\sigma}$, independent of t, as in (2.4), when Assumption **(H1)** holds. $\Box$

Let $\underline{\alpha}\colon \mathbb{R}^d \to A$ be a feedback control and let $X = X^{\underline{\alpha}}$ be the associated diffusion process governed by

$$dX_t \quad = \quad b(X_t, \underline{\alpha}(X_t))\, dt + \sigma(X_t, \underline{\alpha}(X_t))\, dW_t. \tag{2.7}$$

Suppose that the functions

$$\underline{b}(x) \; := \; b(x, \underline{\alpha}(x)), \qquad \underline{\sigma}(x) \; := \; \sigma(x, \underline{\alpha}(x)) \tag{2.8}$$

are Lipschitz. Then, equation (2.7) defines a time-homogeneous Markov process $\{X_t^{\underline{\alpha}}, t \geq 0\}$, and we denote by $(P_t^{\underline{\alpha}})_{t \geq 0}$ the associated semigroup, which acts on $B(\mathbb{R}^d)$, the set of bounded measurable functions φ, by

$$P_t^{\underline{\alpha}} \varphi(x) \quad = \quad \mathbb{E}_x \big[\varphi(X_t^{\underline{\alpha}}) \big], \qquad t \geq 0, \; x \in \mathbb{R}^d.$$

Notice that $(P_t^{\underline{\alpha}})_{t \geq 0}$ has the Feller property, i.e., for any $f \in C_b(\mathbb{R}^d)$, the space of continuous and bounded functions on $\mathbb{R}^d$, we have that $P_t^{\underline{\alpha}} f \in C_b(\mathbb{R}^d)$. The next result shows the ergodicity of $X^{\underline{\alpha}}$.

Proposition 2.1 *Let $\underline{\alpha}\colon \mathbb{R}^d \to A$ be a feedback control such that $\underline{b}, \underline{\sigma}$ in* (2.8) *satisfy Assumption* **(H1)**. *Then $X^{\underline{\alpha}}$ is ergodic, i.e., the following assertions are valid:*

(i) *There exists a unique invariant probability measure $\rho = \rho^{\underline{\alpha}}$ on $\mathbb{R}^d$:*

$$\int P_t^{\underline{\alpha}} \varphi(x) \rho(dx) \quad = \quad \int \varphi(x) \rho(dx), \quad \forall\, t \geq 0, \; \varphi \in B(\mathbb{R}^d).$$

(ii) *$X_t^{\underline{\alpha}}$ converges weakly to ρ as $t \to \infty$:*

$$P_t^{\underline{\alpha}} \varphi(x) \quad \longrightarrow \quad \int \varphi(x) \rho(dx), \quad \text{as } \; t \to \infty, \quad \forall\, x \in \mathbb{R}^d, \, \varphi \in C_b(\mathbb{R}^d). \tag{2.9}$$

Moreover, $\int |x|^2 \rho(dx) < \infty$, and the convergence (2.9) *holds for all continuous φ satisfying a linear growth condition.*

The proof is based on the "pullback" method (see, e.g., Theorem 6.3.2 in [8]) and is detailed in the Appendix.

3 Elliptic HJB equation

For any $\beta > 0$, let us consider the fully nonlinear elliptic equation of HJB type:

$$\beta\, v^\beta - \sup_{a\in A}\big[\mathcal{L}^a v^\beta + f(x,a,\beta v^\beta)\big] \;=\; 0, \quad \text{on } \mathbb{R}^d, \tag{3.1}$$

where

$$\mathcal{L}^a\varphi \;=\; b(x,a).D_x\varphi + \frac{1}{2}\mathrm{tr}\big(\sigma\sigma^\intercal(x,a)D_x^2\varphi\big).$$

Notice that in the particular case where $f = f(x,a)$ does not depend on y, the equation (3.1) is the dynamic programming equation associated to the stochastic control problem on infinite horizon:

$$v^\beta(x) \;:=\; \sup_{\alpha\in\mathcal{A}} \mathbb{E}_x\Big[\int_0^\infty e^{-\beta t} f(X_t^\alpha,\alpha_t)dt\Big].$$

In this case, it is easy to see from the Lipschitz and growth condition on f in **(H2)**(i), and the estimates (2.4), (2.5) that the sequence of functions $(v^\beta)_\beta$ satisfies the uniform estimates:

$$v^\beta(x) \;\le\; \frac{C}{\beta}(1+|x|), \qquad |v^\beta(x)-v^\beta(x')| \;\le\; C|x-x'|, \quad \forall x,x'\in\mathbb{R}^d, \tag{3.2}$$

for some positive constant C independent of β.

In the general case $f = f(x,a,y)$, this section is devoted to the existence and uniqueness of a viscosity solution v^β to (3.1), and to uniform estimate on $(v^\beta)_\beta$ as in (3.2). To this purpose, we introduce the following class of BSDE with nonpositive jumps over an infinite horizon, for any $\beta > 0$:

$$\begin{aligned} Y_t^\beta \;&=\; Y_T^\beta - \beta\int_t^T Y_s^\beta ds + \int_t^T f(X_s,I_s,\beta Y_s^\beta)ds + K_T^\beta - K_t^\beta \\ &\quad - \int_t^T Z_s^\beta dW_s - \int_t^T\int_A U_s^\beta(a)\tilde{\mu}(ds,da), \qquad 0\le t\le T,\ \forall T\in(0,\infty), \end{aligned} \tag{3.3}$$

and

$$U_t^\beta(a) \;\le\; 0, \qquad dt\otimes d\mathbb{P}\otimes\vartheta(da)\text{-a.e.} \tag{3.4}$$

BSDEs driven by Brownian motion over an infinite horizon have been introduced in [9], [22], studied also in [5] and extended in [25], and related to elliptic semi-linear PDEs. Here, we extend this definition to BSDEs driven by Brownian motion and Poisson random measure, and with the nonpositivity constraint on the jump component. A *minimal solution* to the elliptic BSDE with nonpositive jumps (3.3)-(3.4) is a quadruple $(Y^\beta,Z^\beta,U^\beta,K^\beta) \in \mathbf{S}^2_{\mathrm{loc}}\times\mathbf{L}^2_{\mathrm{loc}}(\mathbf{W})\times\mathbf{L}^2_{\mathrm{loc}}(\tilde{\mu})\times\mathbf{K}^2_{\mathrm{loc}}$ satisfying (3.3)-(3.4), with $|Y_t^\beta| \le C(1+|X_t|)$, for all $t\ge 0$ and for some constant C, such that for any other solution $(\bar{Y}^\beta,\bar{Z}^\beta,\bar{U}^\beta,\bar{K}^\beta) \in \mathbf{S}^2_{\mathrm{loc}}\times\mathbf{L}^2_{\mathrm{loc}}(\mathbf{W})\times\mathbf{L}^2_{\mathrm{loc}}(\tilde{\mu})\times\mathbf{K}^2_{\mathrm{loc}}$ to (3.3)-(3.4), satisfying $|\bar{Y}_t^\beta| \le C'(1+|X_t|)$ for some constant C', we have $Y_t^\beta \le \bar{Y}_t^\beta$, $\mathbb{P}$-a.s., for all $t\ge 0$.

Remark 3.1 There exists at most one minimal solution to (3.3)-(3.4). Indeed, let (Y, Z, U, K) and $(\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K})$ be two minimal solutions to (3.3)-(3.4). The uniqueness of the Y component is clear by definition. Regarding the other components, taking the difference between the two backward equations we obtain

$$\int_0^t \big(Z_s - \tilde{Z}_s\big) dW_s \;=\; K_t - \tilde{K}_t - \int_0^t \int_A \big(U_s(a) - \tilde{U}_s(a)\big) \tilde{\mu}(ds, da), \tag{3.5}$$

for all $0 \leq t \leq T$, $\mathbb{P}$-almost surely. Then, we see that the right-hand side is a finite variation process, while the left-hand side has not finite variation, unless $Z = \tilde{Z}$. Now, from (3.5), we obtain the identity

$$\int_0^t \int_A \big(U_s(a) - \tilde{U}_s(a)\big) \mu(ds, da) \;=\; \int_0^t \int_A \big(U_s(a) - \tilde{U}_s(a)\big) \vartheta(da) ds + K_t - \tilde{K}_t,$$

where the right-hand side is a predictable process, therefore it has no totally inaccessible jumps (see, e.g., Proposition 2.24, Chapter I, in [18]); on the other hand, the left-hand side is a pure-jump process with totally inaccessible jumps, unless $U = \tilde{U}$. As a consequence, we must have $U = \tilde{U}$, from which it follows that $K = \tilde{K}$. $\Box$

In the sequel, we prove by a penalization approach the existence of the minimal solution to (3.3)-(3.4), which shall provide the solution to the elliptic nonlinear HJB equation (3.1). Then, by using this probabilistic representation of v^β, we shall state uniform Lipschitz estimate on $(v^\beta)_\beta$.

3.1 Elliptic penalized BSDE

For any $\beta > 0$ and $n \in \mathbb{N}$, we consider the penalized BSDE on $[0, \infty)$, $\mathbb{P}$-a.s.,

$$\begin{aligned} Y_t^{\beta,n} \;&=\; Y_T^{\beta,n} - \beta \int_t^T Y_s^{\beta,n} \, ds + \int_t^T f(X_s, I_s, \beta Y_s^{\beta,n}) \, ds + n \int_t^T \int_A (U_s^{\beta,n}(a))_+ \vartheta(da) \, ds \\ &\quad - \int_t^T Z_s^{\beta,n} \, dW_s - \int_t^T \int_A U_s^{\beta,n}(a) \, \tilde{\mu}(ds, da), \qquad 0 \leq t \leq T < \infty, \end{aligned} \tag{3.6}$$

where $h_+ = \max(h, 0)$ denotes the positive part of the function h.

We first state an a priori estimate on the above elliptic penalized BSDE.

Lemma 3.1 *Suppose that* **(H1)** *holds. Let $(x, a), (x', a') \in \mathbb{R}^d \times A$ and $(Y^{1,\beta,n}, Z^{1,\beta,n}, U^{1,\beta,n})$ (resp. $(Y^{2,\beta,n}, Z^{2,\beta,n}, U^{2,\beta,n})$) be a solution in $\mathbf{S}^2_{\mathrm{loc}} \times \mathbf{L}^2_{\mathrm{loc}}(\mathbf{W}) \times \mathbf{L}^2_{\mathrm{loc}}(\tilde{\mu})$ to (3.6), with $(X, I) = (X^{x,a}, I^a)$ (resp. $(X, I) = (X^{x',a'}, I^{a'})$) and $f = f_1$ (resp. $f = f_2$) satisfying assumption* **(H2)**. *Set $\Delta_t Y = Y_t^{1,\beta,n} - Y_t^{2,\beta,n}$, $\Delta_t Z = Z_t^{1,\beta,n} - Z_t^{2,\beta,n}$, $\Delta_t U(a'') = U_t^{1,\beta,n}(a'') - U_t^{2,\beta,n}(a'')$, $\Delta'_t f_1 = f_1(X_t^{x,a}, I_t^a, \beta Y_t^{1,\beta,n}) - f_1(X_t^{x',a'}, I_t^{a'}, \beta Y_t^{1,\beta,n})$, and $\Delta_t f = f_1(X_t^{x',a'}, I_t^{a'}, \beta Y_t^{2,\beta,n}) - f_2(X_t^{x',a'}, I_t^{a'}, \beta Y_t^{2,\beta,n})$, $t \geq 0$, $a'' \in A$. Then, there exists $\nu \in \mathcal{V}_n$ such that for all $T \in (0, \infty)$,*

$$|\Delta_t Y|^2 \;\leq\; \mathbb{E}^\nu \bigg[e^{-2\beta(T-t)} |\Delta_T Y|^2 + 2 \int_t^T e^{-2\beta(s-t)} \Delta_s Y (\Delta'_s f_1 + \Delta_s f) ds \bigg| \mathcal{F}_t \bigg], \tag{3.7}$$

for all $0 \leq t \leq T$.

Proof. See Appendix. □

The next result states the existence and uniqueness of a solution to (3.6), and uniform estimate on the solution.

Proposition 3.1 *Let Assumptions* **(H1)** *and* **(H2)** *hold. Then, for any* $(x, a, \beta, n) \in \mathbb{R}^d \times A \times (0, \infty) \times \mathbb{N}$*, there exists a solution* $(Y^{x,a,\beta,n}, Z^{x,a,\beta,n}, U^{x,a,\beta,n}) \in \mathbf{S^2_{\mathrm{loc}}} \times \mathbf{L^2_{\mathrm{loc}}}(\mathbf{W}) \times \mathbf{L^2_{\mathrm{loc}}}(\tilde{\mu})$ *to* (3.6)*, with* $(X, I) = (X^{x,a}, I^a)$*, and satisfying:*

$$|Y_t^{x,a,\beta,n}| \;\leq\; \frac{C_{b,\sigma,f}}{\beta}(1 + |X_t^{x,a}|), \quad \forall t \geq 0, \tag{3.8}$$

for some positive constant $C_{b,\sigma,f}$ *depending only on* b, σ, f*. Moreover, this solution is unique in the class of triplets* $(Y, Z, U) \in \mathbf{S^2_{\mathrm{loc}}} \times \mathbf{L^2_{\mathrm{loc}}}(\mathbf{W}) \times \mathbf{L^2_{\mathrm{loc}}}(\tilde{\mu})$ *satisfying the condition* $|Y_t| \leq C(1 + |X_t^{x,a}|)$*, for all* $t \geq 0$ *and for some positive constant* C *(possibly depending on* x*,* a*,* β *and* n*).*

Proof. See Appendix. □

For any $(x, a, \beta, n) \in \mathbb{R}^d \times A \times (0, \infty) \times \mathbb{N}$, we notice that $Y_0^{x,a,\beta,n}$ is a constant since it is $\mathcal{F}_0$-measurable. Therefore, for each $\beta > 0$, $n \in \mathbb{N}$, we introduce the function $v^{\beta,n} \colon \mathbb{R}^d \times A \to \mathbb{R}$ defined as

$$v^{\beta,n}(x, a) \;:=\; Y_0^{x,a,\beta,n}, \qquad (x, a) \in \mathbb{R}^d \times A. \tag{3.9}$$

Let us now investigate some key properties of the function $v^{\beta,n}$. We first state a uniform Lipschitz estimate on $(v^{\beta,n})$.

Lemma 3.2 *Let Assumptions* **(H1)** *and* **(H2)** *hold. For any* $(\beta, n) \in (0, \infty) \times \mathbb{N}$*, the function* $v^{\beta,n}$ *is such that:* $Y_t^{x,a,\beta,n} = v^{\beta,n}(X_t^{x,a}, I_t^a)$*, for all* $t \geq 0$*, and* $(x, a) \in \mathbb{R}^d \times A$*. Moreover, there exists some positive constant* C *depending only on* b, σ, f*, and independent of* β, n *such that*

$$v^{\beta,n}(x, a) \;\leq\; \frac{C}{\beta}\big(1 + |x|\big), \tag{3.10}$$

$$\big|v^{\beta,n}(x, a) - v^{\beta,n}(x', a)\big| \;\leq\; C|x - x'|, \tag{3.11}$$

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

Proof. See Appendix. □

As expected, for fixed (β, n), the function $v^{\beta,n}$ is related to the elliptic integro-differential equation:

$$\begin{aligned} &\beta\, v^{\beta,n}(x, a) - \mathcal{L}^a v^{\beta,n}(x, a) - \mathcal{M}^a v^{\beta,n}(x, a) - f\big(x, a, \beta v^{\beta,n}(x, a)\big) \\ &\qquad - n \int_A (v^{\beta,n}(x, a') - v^{\beta,n}(x, a))_+\, \vartheta(da') \;=\; 0, \quad \text{on } \mathbb{R}^d \times A, \end{aligned} \tag{3.12}$$

where

$$\mathcal{M}^a \varphi(a) \;=\; \int_A \big(\varphi(a') - \varphi(a)\big)\vartheta(da'),$$

for any $\varphi \in C(A)$. More precisely, we have the following result.

Proposition 3.2 *Let Assumptions* **(H1)** *and* **(H2)** *hold. Then, the function* $v^{\beta,n}$ *in* (3.9) *is a continuous viscosity solution to* (3.12)*, i.e., it is continuous on* $\mathbb{R}^d \times A$ *and it is a viscosity supersolution* (*resp. subsolution*) *to* (3.12)*, namely*

$$\begin{aligned}\beta\,\varphi(x,a) \quad &\geq \ (\textit{resp.} \ \leq) \quad \mathcal{L}^a\varphi(x,a) + \mathcal{M}^a\varphi(x,a) + f\big(x,a,\beta\varphi(x,a)\big) \\ &\qquad\qquad + n\int_A (\varphi(x,a') - \varphi(x,a))_+\,\vartheta(da')\end{aligned}$$

for any $(x,a) \in \mathbb{R}^d \times A$ *and any* $\varphi \in C^2(\mathbb{R}^d \times \mathbb{R}^q)$ *such that*

$$0 \;=\; (v^{\beta,n} - \varphi)(x,a) \;=\; \min_{\mathbb{R}^d\times A}(v^{\beta,n} - \varphi) \quad (\textit{resp.}\ \max_{\mathbb{R}^d\times A}(v^{\beta,n} - \varphi)).$$

Proof. See Appendix. □

3.2 Elliptic BSDE with nonpositive jumps

We can now prove the existence of the minimal solution to the elliptic BSDE with nonpositive jumps (3.3)-(3.4).

Proposition 3.3 *Let Assumptions* **(H1)** *and* **(H2)** *hold. Then, for any* $\beta > 0$ *and* $(x,a) \in \mathbb{R}^d \times A$ *there exists a solution* $(Y^{x,a,\beta}, Z^{x,a,\beta}, U^{x,a,\beta}, K^{x,a,\beta}) \in \mathbf{S^2_{\mathrm{loc}}} \times \mathbf{L^2_{\mathrm{loc}}(W)} \times \mathbf{L^2_{\mathrm{loc}}}(\tilde\mu) \times \mathbf{K^2_{\mathrm{loc}}}$ *to* (3.3)-(3.4)*, with* $(X,I) = (X^{x,a}, I^a)$*. Moreover*

(i) $Y^{x,a,\beta}$ *is the increasing limit of* $(Y^{x,a,\beta,n})_n$ *and satisfies*

$$|Y_t^{x,a,\beta}| \;\leq\; \frac{C}{\beta}(1 + |X_t^{x,a}|), \quad \forall t \geq 0, \tag{3.13}$$

for some positive constant C *independent of* β, x, a, t.

(ii) $(Z^{x,a,\beta}_{|[0,T]}, U^{x,a,\beta}_{|[0,T]})$*, for any* $T > 0$*, is the strong* (*resp. weak*) *limit of* $(Z^{x,a,\beta,n}_{|[0,T]}, U^{x,a,\beta,n}_{|[0,T]})_n$ *in* $\mathbf{L^p(W;0,T)} \times \mathbf{L^p}(\tilde\mu;\mathbf{0,T})$*, with* $p \in [1,2)$*,* (*resp. in* $\mathbf{L^2(W;0,T)} \times \mathbf{L^2}(\tilde\mu;\mathbf{0,T})$).

(iii) $K_t^{x,a,\beta}$ *is the weak limit of* $(K_t^{x,a,\beta,n})_n$ *in* $\mathbf{L^2}(\mathcal{F}_t)$*, for any* $t \geq 0$.

Furthermore, this solution is minimal in the class of quadruplets $(Y,Z,U,K) \in \mathbf{S^2_{\mathrm{loc}}} \times \mathbf{L^2_{\mathrm{loc}}(W)} \times \mathbf{L^2_{\mathrm{loc}}}(\tilde\mu) \times \mathbf{K^2_{\mathrm{loc}}}$ *satisfying the condition* $|Y_t| \leq C(1 + |X_t^{x,a}|)$*, for all* $t \geq 0$ *and for some positive constant* C *(possibly depending on* x*,* a*, and* β*).*

Proof. See Appendix. □

For any $\beta > 0$, let us introduce the deterministic function $v^\beta \colon \mathbb{R}^d \times A \to \mathbb{R}$ defined as follows

$$v^\beta(x,a) \;:=\; Y_0^{x,a,\beta}, \qquad \forall\,(x,a) \in \mathbb{R}^d \times A. \tag{3.14}$$

From point (i) of Proposition 3.3, it follows that $(v^{\beta,n})_n$ converges increasingly to v^β. Then, the identification $Y_t^{x,a,\beta,n} = v^{\beta,n}(X_t^{x,a}, I_t^a)$ implies that $Y_t^{x,a,\beta} = v^\beta(X_t^{x,a}, I_t^a)$. We shall now investigate the relation between v^β and the fully nonlinear elliptic PDE of HJB type (3.1). More precisely, we shall prove that v^β solves in the viscosity sense equation (3.1). The main issue is to prove that v^β does not depend actually on a. However, as we do not know a priori that the function v^β is continuous in both arguments, we shall rely on discontinuous viscosity solutions arguments as in [19], and make the following assumptions on the set A and the intensity measure ϑ:

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

(Hϑ) The measure ϑ supports the whole set $\mathring{A}$, and the boundary of A: $\partial A = A \backslash \mathring{A}$, is negligible with respect to ϑ.

Notice that equation (3.1) does not depend on ϑ, and this intensity measure only appears in order to give a probabilistic representation of v^β. Therefore, we have the choice to fix an intensity measure ϑ satisfying condition **(Hϑ)**, which is anyway a fairly general condition easy to realize. In the sequel, we shall make the standing assumption that **(Hϑ)** holds.

We can now state the main result of this section.

Theorem 3.1 *Let Assumptions* **(H1)**, **(H2)**, **(HA)**, *and* **(Hϑ)** *hold. Then, for any $\beta > 0$, the function v^β in* (3.14) *does not depend on the variable a on $\mathbb{R}^d \times \mathring{A}$:*

$$v^\beta(x, a) \;=\; v^\beta(x, a'), \qquad a, a' \in \mathring{A}, \tag{3.15}$$

for all $x \in \mathbb{R}^d$, and we set by misuse of notation: $v^\beta(x) = v^\beta(x, a)$, $x \in \mathbb{R}^d$, for any $a \in \mathring{A}$. Then, v^β is the unique continuous viscosity solution to equation (3.1), *i.e., it is continuous on $\mathbb{R}^d$ and it is a viscosity supersolution (resp. subsolution) to* (3.1), *namely:*

$$\beta\, \varphi(x) \;\geq\; (\mathit{resp.}\ \leq) \;\sup_{a \in A} \big[\mathcal{L}^a \varphi(x) + f(x, a, \beta\, \varphi(x))\big],$$

for any $x \in \mathbb{R}^d$ and any $\varphi \in C^2(\mathbb{R}^d)$ such that

$$0 \;=\; (v^\beta - \varphi)(x) \;=\; \min_{\mathbb{R}^d}(v^\beta - \varphi) \quad (\mathit{resp.}\ \max_{\mathbb{R}^d}(v^\beta - \varphi)).$$

Moreover, there exists some positive constant C independent of β such that:

$$|\beta v^\beta(x)| \;\leq\; C(1 + |x|), \quad \forall x \in \mathbb{R}^d, \tag{3.16}$$

$$|v^\beta(x) - v^\beta(x')| \;\leq\; C|x - x'|, \qquad \forall\, x, x' \in \mathbb{R}^d. \tag{3.17}$$

Proof. We use the corresponding result for the parabolic case in [19] to prove the non dependence of v^β on a, and then the viscosity property to the elliptic equation. More precisely, we start by observing that, for any $\beta > 0$ and $T > 0$, $v^{\beta,n}$ is a viscosity solution to the parabolic PDE on $[0, T] \times \mathbb{R}^d \times A$:

$$\begin{aligned} \beta\, w(t, x, a) \;&=\; \frac{\partial w}{\partial t}(t, x, a) + \mathcal{L}^a w(t, x, a) + \mathcal{M}^a w(t, x, a) + f(x, a, \beta\, w(t, x, a)) \\ &\quad + n \int_A (w(t, x, a') - w(t, x, a))_+\, \vartheta(da'), \end{aligned}$$

with terminal condition $w(T, x, a) = v^{\beta,n}(x, a)$, for all $(x, a) \in \mathbb{R}^d \times A$, i.e.

$$\beta\, \varphi(t, x, a) \;\geq\; (\text{resp. } \leq) \;\; \frac{\partial \varphi}{\partial t}(t, x, a) + \mathcal{L}^a \varphi(t, x, a) + \mathcal{M}^a \varphi(t, x, a) + f(x, a, \beta\, \varphi(t, x, a))$$

$$+n\int_A (\varphi(t,x,a') - \varphi(t,x,a))_+ \, \vartheta(da'), \tag{3.18}$$

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

$$0 \;=\; (v^{\beta,n} - \varphi)(t,x,a) \;=\; \min_{[0,T]\times\mathbb{R}^d\times A}(v^{\beta,n} - \varphi) \quad (\text{resp.} \max_{[0,T]\times\mathbb{R}^d\times A}(v^{\beta,n} - \varphi)). \tag{3.19}$$

To prove this, for any $(t,x,a) \in [0,T) \times \mathbb{R}^d \times A$ and any $\varphi \in C^{1,2}([0,T] \times (\mathbb{R}^d \times A))$ satisfying (3.18)-(3.19), set $\psi_t(x',a') := \varphi(t,x',a')$, for all $(x',a') \in \mathbb{R}^d \times \mathbb{R}^q$. Then $\psi_t \in C^2(\mathbb{R}^d \times \mathbb{R}^q)$ and

$$0 \;=\; (v^{\beta,n} - \psi_t)(x,a) \;=\; \min_{\mathbb{R}^d\times A}(v^{\beta,n} - \psi_t) \quad (\text{resp.} \max_{\mathbb{R}^d\times A}(v^{\beta,n} - \psi_t)).$$

Observing that $\partial_t\varphi(t,x,a) \leq 0$ (resp. $\partial_t\varphi(t,x,a) \geq 0$) since $\varphi(t,x,a) = \max_{t'\in[0,T)} \varphi(t',x,a)$ (resp. $\varphi(t,x,a) = \min_{t'\in[0,T)} \varphi(t',x,a)$), it follows from Proposition 3.2 that (3.18) is satisfied. Then, from Theorem 3.1 in [19], we deduce by sending n to infinity that the function v^β does not depend on the variable a, and so (3.15) holds. We should point out that in [19] the terminal condition of the parabolic PDE solved by $v^{\beta,n}$ does not depend on n, contrary to our case. However, the part of the proof of Theorem 3.1 in [19] regarding the independence with respect to the variable a does not involve the terminal condition, so the result still holds. Next, we obtain again from Theorem 3.1 in [19] that v^β solves in the viscosity sense the parabolic PDE on $[0,T) \times \mathbb{R}^d \times A$ (as before, we do not call in the terminal condition):

$$\beta\, w(t,x,a) - \frac{\partial w}{\partial t}(t,x,a) - \sup_{a\in A}\big[\mathcal{L}^a w(t,x,a) + f(x,a,\beta\, w(t,x,a))\big] \;=\; 0.$$

Since this equation holds for any T, and v^β does not depend on t, we obtain that v^β is a viscosity solution to the elliptic equation (3.1). The uniqueness follows from Theorem 7.4 in [17]. Finally, the linear growth and Lipschitz properties (3.16)-(3.17) of v^β are direct consequences of the corresponding properties (3.13) for $Y_0^{x,a,\beta}$ and (3.11) for $v^{\beta,n}$, respectively. □

Remark 3.2 Notice that, from the identification $Y^{x,a,\beta} = v^\beta(X^{x,a})$, for any $x \in \mathbb{R}^d$ and for some $a \in \mathring{A}$, and the Lipschitz property (3.17), it follows that $Y^{x,a,\beta}$ is a continuous process, so that $U^{x,a,\beta} \equiv 0$, while $K^{x,a,\beta}$ is also a continuous process. □

4 Ergodic HJB equation and ergodic BSDE with nonpositive jumps

This section is devoted to the existence of a solution pair (λ, ϕ) to the ergodic HJB equation

$$\lambda - \sup_{a\in A}\big[\mathcal{L}^a\phi + f(x,a,\lambda)\big] \;=\; 0, \quad \text{on } \mathbb{R}^d, \tag{4.1}$$

and to its probabilistic representation in terms of ergodic BSDE with nonpositive jumps. We first give the definition of viscosity solution to equation (4.1).

Definition 4.1 (i) *A pair (λ, ϕ), with λ a real number and $\phi\colon \mathbb{R}^d \to \mathbb{R}$ a lower (resp. upper) semicontinuous function, is called a* viscosity supersolution (*resp.* viscosity subsolution) *to equation* (4.1) *if*

$$\lambda \;\geq\; (resp.\;\leq)\; \sup_{a\in A}\{\mathcal{L}^a\varphi(x) + f(x,a,\lambda)\},$$

for any $x \in \mathbb{R}^d$ and any $\varphi \in C^2(\mathbb{R}^d)$ such that

$$(\phi-\varphi)(x) \;=\; \min_{\mathbb{R}^d}(\phi-\varphi) \quad (resp.\; \max_{\mathbb{R}^d}(\phi-\varphi)).$$

(ii) *A pair (λ, ϕ), with λ a real number and $\phi\colon \mathbb{R}^q \to \mathbb{R}$ a continuous function, is called a* viscosity solution *to equation* (4.1) *if it is both a viscosity supersolution and a viscosity subsolution to* (4.1).

Theorem 4.1 *Let Assumptions* **(H1)**, **(H2)**, **(HA)**, *and* **(Hϑ)** *hold. Then, there exists a viscosity solution pair (λ, ϕ) to* (4.1) *with ϕ Lipschitz on $\mathbb{R}^d$.*

Proof. We follow the approximation procedure as in [15] or [11]: for any $\beta > 0$, let v^β be the solution of (3.1) given in Theorem 3.1, and define $\lambda_\beta \in \mathbb{R}$, and the function $\phi^\beta\colon \mathbb{R}^d \to \mathbb{R}$ by:

$$\lambda_\beta \;:=\; \beta v^\beta(0), \qquad \phi^\beta(x) \;:=\; v^\beta(x) - v^\beta(0), \qquad x \in \mathbb{R}^d.$$

By (3.16)-(3.17), we see that there exists some positive constant C independent of β such that

$$\begin{aligned} \sup_{\beta>0} |\lambda_\beta| &\;<\; \infty, \\ \sup_{\beta>0} |\phi^\beta(x)| &\;\leq\; C|x|, \qquad \forall\, x \in \mathbb{R}^d. \end{aligned}$$

Then, the family of functions $(\phi^\beta)_\beta$ is equicontinuous and uniformly bounded on every compact subset of $\mathbb{R}^d$. Thus, by means of Ascoli-Arzelà theorem (for more details, see (4.4) and (4.5) in [11]), we can construct a sequence $(\beta_k)_k$ decreasing monotonically to zero such that, for all $x \in \mathbb{R}^d$,

$$\lambda_{\beta_k} \overset{k\to\infty}{\longrightarrow} \lambda, \qquad \phi^{\beta_k}(x) \overset{k\to\infty}{\longrightarrow} \phi(x), \tag{4.2}$$

for some real constant λ, and a Lipschitz function $\phi\colon \mathbb{R}^d \to \mathbb{R}$. Moreover, the convergence of ϕ^{β_k} towards ϕ is uniform on compact sets.

Now, from the elliptic equation (3.1) satisfied by v^β, and by definition of $(\lambda_\beta, \phi^\beta)$, we see that ϕ^β is a viscosity solution to:

$$\lambda_\beta + \beta\phi^\beta - \sup_{a\in A}\big[\mathcal{L}^a\phi^\beta + f(x,a,\lambda_\beta + \beta\phi^\beta)\big] \;=\; 0, \quad \text{on } \mathbb{R}^d. \tag{4.3}$$

Let us denote by

$$F_k(x,r,p,M) \;:=\; \lambda_{\beta_k} + \beta_k r$$

$$- \sup_{a \in A} \big[b(x,a).p + \frac{1}{2} \mathrm{tr}\big(\sigma\sigma^\intercal(x,a)M\big) + f(x,a,\lambda_{\beta_k} + \beta_k\, r) \big],$$

so that by (4.3), ϕ^{β_k} is a viscosity solution to:

$$F_k(x, \phi^{\beta_k}, D_x\phi^{\beta_k}, D_x^2\phi^{\beta^k}) \;=\; 0,$$

and set

$$G(x,r,p,M) \;:=\; \lambda - \sup_{a \in A} \big[b(x,a).p + \frac{1}{2} \mathrm{tr}\big(\sigma\sigma^\intercal(x,a)M\big) + f(x,a,\lambda) \big],$$

for all $(x,r,p,M) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d\times d}$. Then, it follows from (4.2) that $\lim_{k\to\infty} F_k(x,r,p,M) = G(x,r,p,M)$. As a consequence, from the method of half-relaxed limits of Barles and Perthame, see Remark 6.3 in [6], we deduce by (4.2) that ϕ is a viscosity solution to:

$$G(x, \phi, D_x\phi, D_x^2\phi) \;=\; 0,$$

i.e., (λ, ϕ) is a viscosity solution to equation (4.1). □

We postpone the uniqueness problem of the ergodic equation (4.1) to the next section, and conclude this section by providing a probabilistic representation formula for a solution to the ergodic equation. Let us introduce the ergodic BSDE with nonpositive jumps, $\mathbb{P}$-a.s.,

$$\begin{aligned} Y_t \;=\; & Y_T + \int_t^T \big(f(X_s, I_s, \lambda) - \lambda\big) ds + K_T - K_t \\ & - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \tilde{\mu}(ds,da), \qquad 0 \le t \le T < \infty \end{aligned} \tag{4.4}$$

and

$$U_t(a) \;\le\; 0, \qquad dt \otimes d\mathbb{P} \otimes \vartheta(da)\text{-a.e.} \tag{4.5}$$

Here, in addition to the components (Y,Z,U,K), the real number λ is part of the unknowns of the ergodic BSDE. We recall that ergodic BSDEs driven by Brownian motion have been defined in [11] for the study of optimal control problems on the drift of a diffusion process, which are related to ergodic semilinear HJB equations. In this paper, we extend this definition by considering the jump constraint (4.5), and our first purpose is to introduce a notion of "minimal" solution to (4.4)-(4.5). However, we notice that the "natural" definition of minimal solution as the solution $(\bar{Y}, \bar{Z}, \bar{U}, \bar{K}, \lambda) \in \mathbf{S^2_{\mathrm{loc}}} \times \mathbf{L^2_{\mathrm{loc}}}(\mathbf{W}) \times \mathbf{L^2_{\mathrm{loc}}}(\tilde{\mu}) \times \mathbf{K^2_{\mathrm{loc}}} \times \mathbb{R}$ to (4.4)-(4.5) such that for any other solution $(\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K}, \tilde{\lambda}) \in \mathbf{S^2_{\mathrm{loc}}} \times \mathbf{L^2_{\mathrm{loc}}}(\mathbf{W}) \times \mathbf{L^2_{\mathrm{loc}}}(\tilde{\mu}) \times \mathbf{K^2_{\mathrm{loc}}} \times \mathbb{R}$ to (4.4)-(4.5) we have $\bar{Y} \le \tilde{Y}$, is not relevant in this case, since $(\bar{Y} - c, \bar{Z}, \bar{U}, \bar{K}, \lambda)$, with $c > 0$, would be another solution to (4.4)-(4.5), contradicting the minimality of $(\bar{Y}, \bar{Z}, \bar{U}, \bar{K}, \lambda)$. For this reason, we give the following definition of minimal solution to the ergodic BSDE with nonpositive jumps (4.4)-(4.5).

Definition 4.2 *A quintuple* $(\underline{Y}, \underline{Z}, \underline{U}, \underline{K}, \lambda) \in \mathbf{S^2_{\mathrm{loc}}} \times \mathbf{L^2_{\mathrm{loc}}}(\mathbf{W}) \times \mathbf{L^2_{\mathrm{loc}}}(\tilde{\mu}) \times \mathbf{K^2_{\mathrm{loc}}} \times \mathbb{R}$ *is called a* minimal solution *to* (4.4)-(4.5) *if, for any* $T > 0$, $(\underline{Y}_{|[0,T]}, \underline{Z}_{|[0,T]}, \underline{U}_{|[0,T]}, \underline{K}_{|[0,T]})$ *is a* *minimal solution to the BSDE with nonpositive jumps on* $[0,T]$ *with terminal condition* $\underline{Y}_T$, $\mathbb{P}$*-a.s.,*

$$\begin{aligned} Y_t &= \underline{Y}_T + \int_t^T \big(f(X_s, I_s, \lambda) - \lambda\big) ds + K_T - K_t \\ &\quad - \int_t^T Z_s dW_s - \int_t^T \int_A U_s(a) \tilde{\mu}(ds, da), \qquad 0 \le t \le T, \end{aligned} \tag{4.6}$$

together with the jump constraint

$$U_t(a) \ \le \ 0, \qquad dt \otimes d\mathbb{P} \otimes \vartheta(da)\text{-a.e.} \tag{4.7}$$

In other words, for any other solution $(\tilde{Y}, \tilde{Z}, \tilde{U}, \tilde{K}) \in \mathbf{S^2(0,T)} \times \mathbf{L^2(W;0,T)} \times \mathbf{L^2}(\tilde{\mu};\mathbf{0,T}) \times \mathbf{K^2(0,T)}$ *to* (4.6)-(4.7) *we have* $\underline{Y}_t \le \tilde{Y}_t$, $\mathbb{P}$*-a.s., for all* $0 \le t \le T$.

Remark 4.1 We do not know a priori if there is uniqueness of a minimal solution to (4.4)-(4.5). This will be discussed later in Remark 5.3. □

We shall now prove that any Lipschitz continuous viscosity solution to the ergodic HJB equation (4.1), admits a probabilistic representation in terms of a minimal solution to the ergodic BSDE (4.4).

Theorem 4.2 *Let Assumptions* **(H1)**, **(H2)**, **(HA)**, *and* **(Hϑ)** *hold. Let* (λ, ϕ), *with* ϕ *Lipschitz, be a viscosity solution to the ergodic equation* (4.1). *Then,* (λ, ϕ) *can be represented by means of a minimal solution to the ergodic BSDE* (4.4)-(4.5), *namely:*

(i) *For any* $a \in \mathring{A}$, *there exists a minimal solution* $(\underline{Y}^{x,a}, \underline{Z}^{x,a}, \underline{U}^{x,a}, \underline{K}^{x,a}, \lambda) \in \mathbf{S^2_{\text{loc}}} \times \mathbf{L^2_{\text{loc}}(W)} \times \mathbf{L^2_{\text{loc}}}(\tilde{\mu}) \times \mathbf{K^2_{\text{loc}}} \times \mathbb{R}$ *to the ergodic BSDE with nonpositive jumps,* $\mathbb{P}$*-a.s.,*

$$\begin{aligned} \underline{Y}_t^{x,a} &= \underline{Y}_T^{x,a} + \int_t^T \big(f(X_s^{x,a}, I_s^a, \lambda) - \lambda\big) ds + \underline{K}_T^{x,a} - \underline{K}_t^{x,a} - \int_t^T \underline{Z}_s^{x,a} dW_s \\ &\quad - \int_t^T \int_A \underline{U}_s^{x,a}(a') \tilde{\mu}(ds, da'), \qquad 0 \le t \le T < \infty, \end{aligned} \tag{4.8}$$

and

$$\underline{U}_t^{x,a}(a') \ \le \ 0, \qquad dt \otimes d\mathbb{P} \otimes \vartheta(da')\text{-a.e.} \tag{4.9}$$

(ii) $\underline{Y}_t^{x,a} = \phi(X_t^{x,a})$ *for* $t \ge 0$, *and, in particular,*

$$\phi(x) \ = \ \underline{Y}_0^{x,a}, \qquad \text{for all } x \in \mathbb{R}^d,$$

for some $a \in \mathring{A}$.

Proof. We start by observing that for any $T > 0$, ϕ is a viscosity solution to the following parabolic equation, in the unknown ψ,

$$\begin{cases} -\dfrac{\partial \psi(t,x)}{\partial t} - \ \sup_{a \in A} \big[\mathcal{L}^a \psi(t,x) + f(x,a,\lambda) - \lambda\big] = 0, & (t,x) \in [0,T) \times \mathbb{R}^d, \\ \psi(T,x) = \phi(x), & x \in \mathbb{R}^d. \end{cases} \tag{4.10}$$

It follows from Theorem 7.4 in [17] that ϕ is the unique uniformly continuous viscosity solution to (4.10). From Theorem 3.1 in [19], we see that this viscosity solution admits a representation in terms of the minimal solution to a BSDE with nonpositive jumps on $[0,T]$. More precisely, for any $T>0$, let $(Y^{x,a,T},Z^{x,a,T},U^{x,a,T},K^{x,a,T})\in\mathbf{S^2(0,T)}\times\mathbf{L^2(W;0,T)}\times\mathbf{L^2}(\tilde\mu;\mathbf{0,T})\times\mathbf{K^2(0,T)}$ be the minimal solution to the BSDE with nonpositive jumps on $[0,T]$, $\mathbb{P}$-a.s.,

$$\begin{aligned}Y_t \;&=\; \phi(X_T^{x,a})+\int_t^T \big(f(X_s^{x,a},I_s^a,\lambda)-\lambda\big)ds+K_T-K_t\\ &\quad -\int_t^T Z_s dW_s-\int_t^T\int_A U_s(a')\tilde\mu(ds,da'),\qquad 0\le t\le T,\end{aligned} \tag{4.11}$$

and

$$U_t(a') \;\le\; 0,\qquad dt\otimes d\mathbb{P}\otimes\vartheta(da')\text{-a.e.} \tag{4.12}$$

Then, from Theorem 3.1 in [19] we see that $Y_t^{x,a,T}=\phi(X_t^{x,a})$, for all $0\le t\le T$. The identification $Y_t^{x,a,T}=\phi(X_t^{x,a})$ also implies that $Y^{x,a,T}$ does not depend on T. Moreover, using the fact that the Y component remains the same, reasoning as in Remark 3.1, we can prove that $Z^{x,a,T}$, $U^{x,a,T}$, and $K^{x,a,T}$ do not depend on T. Let $\underline{Y}^{x,a}_{|[0,T]}=Y^{x,a,T}$, $\underline{Z}^{x,a}_{|[0,T]}=Z^{x,a,T}$, $\underline{U}^{x,a}_{|[0,T]}=U^{x,a,T}$, and $\underline{K}^{x,a}_{|[0,T]}=K^{x,a,T}$, for any $T>0$. Then, we see that, for any $T>0$, the quadruple $(\underline{Y}^{x,a}_{|[0,T]},\underline{Z}^{x,a}_{|[0,T]},\underline{U}^{x,a}_{|[0,T]},\underline{K}^{x,a}_{|[0,T]})$ is the minimal solution to (4.11)-(4.12), from which we deduce that $(\underline{Y}^{x,a},\underline{Z}^{x,a},\underline{U}^{x,a},\underline{K}^{x,a},\lambda)$ is a minimal solution to the ergodic BSDE (4.4)-(4.5). □

Remark 4.2 Notice that, a minimal solution to the ergodic BSDE with nonpositive jumps (4.4)-(4.5) provides a viscosity solution to the ergodic HJB equation (4.1). In other words, the converse of the result stated in Theorem 4.2 is also true. More precisely, let $\phi\colon\mathbb{R}^d\times\mathbb{R}^q\to\mathbb{R}$ be a Lipschitz function and $\lambda\in\mathbb{R}$. For any $x\in\mathbb{R}^d$, consider $a\in\mathring{A}$ and a minimal solution $(\underline{Y}^{x,a},\underline{Z}^{x,a},\underline{U}^{x,a},\underline{K}^{x,a},\lambda)\in\mathbf{S^2_{\mathrm{loc}}}\times\mathbf{L^2_{\mathrm{loc}}(W)}\times\mathbf{L^2_{\mathrm{loc}}}(\tilde\mu)\times\mathbf{K^2_{\mathrm{loc}}}\times\mathbb{R}$ to the ergodic BSDE with nonpositive jumps (4.8)-(4.9), such that $\underline{Y}_t^{x,a}=\phi(X_t^{x,a},I_t^a)$, $t\ge0$. Then, ϕ does not depend on the variable a and (λ,ϕ) is a viscosity solution to the ergodic equation (4.1). Indeed, fix $T>0$, then by Definition 4.2 we know that $(\underline{Y}^{x,a},\underline{Z}^{x,a},\underline{U}^{x,a},\underline{K}^{x,a})$ is the unique minimal solution to the BSDE with nonpositive jumps on $[0,T]$ in (4.6)-(4.7). It follows from Theorem 3.1 in [19] that ϕ does not depend on a (this last property follows from Proposition 3.2 in [19], which does not call in the terminal condition; therefore, even if in our case, as opposite to [19], the terminal condition depends on a, the result is still valid) and it is a viscosity solution to equation (4.10). As a consequence, (λ,ϕ) is a viscosity solution to the ergodic equation (4.1). □

5 Convergence of solutions

Let us consider the parabolic fully nonlinear equation of HJB type:

$$\begin{cases}\dfrac{\partial v}{\partial T}-\sup\limits_{a\in A}\big[\mathcal{L}^a v+f\big(x,a,\dfrac{v}{T+1}\big)\big],\;=\;0, & \text{on } (0,\infty)\times\mathbb{R}^d,\\ v(0,.)\;=\;h, & \text{on } \mathbb{R}^d,\end{cases} \tag{5.1}$$

where h is a Lipschitz function on $\mathbb{R}^d$. Existence and uniqueness of a solution $v(T,x)$ to (5.1) is studied in the next paragraph. Now, let us consider a solution pair (λ,ϕ), with ϕ Lipschitz, to the ergodic equation (4.1). The main result of this section is to prove that

$$\frac{v(T,x)}{T} \overset{T\to\infty}{\longrightarrow} \lambda. \tag{5.2}$$

Consequently, this will show the uniqueness of the component λ in (4.1). We shall end this part of the paper by proving a stronger result than (5.2) under additional assumptions. More precisely, suppose that ϕ belongs to $C^2(\mathbb{R}^d)$ so that (λ,ϕ) is a classical solution to the ergodic equation (4.1), and assume that in the ergodic equation, the supremum is attained at $a=\underline{\alpha}(x)$, for every $x\in\mathbb{R}^d$, for some locally Lipschitz function $\underline{\alpha}\colon A\to\mathbb{R}^d$. Then, the following convergence holds

$$v(T,x)-(\lambda T+\phi(x)) \overset{T\to\infty}{\longrightarrow} c, \tag{5.3}$$

for some constant c. In particular, ϕ is uniquely determined up to a constant.

5.1 Wellposedness of the parabolic equation (5.1)

We shall build a solution to equation (5.1) through BSDE methods, as this construction will be useful in the sequel. More precisely, from Theorem 3.1 in [19], under **(H1)**(i), **(H2)**(i), **(HA)**, and **(Hϑ)**, there exists a uniformly continuous viscosity solution v to equation (5.1), which admits the following probabilistic representation formula

$$v(T,x) \;=\; Y_0^{x,a,T}, \qquad (T,x)\in[0,\infty)\times\mathbb{R}^d, \tag{5.4}$$

for any $a\in\mathring{A}$, where $(Y^{x,a,T},Z^{x,a,T},U^{x,a,T},K^{x,a,T})\in\mathbf{S^2(0,T)}\times\mathbf{L^2(W;0,T)}\times\mathbf{L^2}(\tilde{\mu};\mathbf{0,T})\times\mathbf{K^2(0,T)}$, with $v(T-t,X_t^{x,a})=Y_t^{x,a,T}$ for all $0\le t\le T$, is the unique minimal solution to the BSDE with nonpositive jumps on $[0,T]$, $\mathbb{P}$-a.s.,

$$\begin{aligned} Y_t^{x,a,T} \;&=\; h(X_T^{x,a})+\int_t^T f\big(X_s^{x,a},I_s^a,\frac{Y_s^{x,a,T}}{T-s+1}\big)ds+K_T^{x,a,T}-K_t^{x,a,T} \\ &\quad -\int_t^T Z_s^{x,a,T}dW_s-\int_t^T\int_A U_s^{x,a,T}(a')\tilde{\mu}(ds,da'), \qquad 0\le t\le T \end{aligned} \tag{5.5}$$

and

$$U_t^{x,a,T}(a') \;\le\; 0, \qquad dt\otimes d\mathbb{P}\otimes\vartheta(da')\text{-a.e.} \tag{5.6}$$

Moreover, from Theorem 7.4 in [17], we have that v is the unique uniformly continuous viscosity solution to (5.1) (observe that Theorem 7.4 in [17] applies to elliptic equations on unbounded domains; interpreting t as a space variable, (5.1) can be seen as an elliptic equation on the space domain $[0,\infty)\times\mathbb{R}^d$, so that we can now apply Theorem 7.4 in [17]).

Remark 5.1 Notice that Theorem 3.1 in [19] is designed for *backward* parabolic PDEs, while (5.1) is a forward parabolic equation. However, we can exploit Theorem 3.1 in [19] by proceeding as follows. For any $T>0$, we consider the HJB equation on $[0,T]\times\mathbb{R}^d$:

$$\begin{cases} -\dfrac{\partial v^T}{\partial t}-\sup\limits_{a\in A}\big[\mathcal{L}^a v^T(t,x)+f\big(x,a,\dfrac{v^T}{T-t+1}\big)\big]=0, \quad (t,x)\in[0,T)\times\mathbb{R}^d, \\ v^T(T,x) \;=\; h(x), \quad x\in\mathbb{R}^d. \end{cases} \tag{5.7}$$

Under **(H1)**(i), **(H2)**(i), **(H**A**)**, and **(H**ϑ**)**, it follows from Theorem 3.1 in [19] that there exists a uniformly continuous viscosity solution v^T to equation (5.7), which admits a probabilistic representation formula in terms of the unique minimal solution to a certain BSDE with nonpositive jumps. Define the function $v(t,x) := v^T(T-t,x)$, for all $(t,x) \in [0,T] \times \mathbb{R}^d$ and for any $T > 0$. Notice that v is well-defined, thanks to uniqueness results (see, e.g., Theorem 7.4 in [17]) of viscosity solutions to equation (5.7) (in particular, $v^T(T-t,x) = v^{T'}(T'-t,x)$, for any $(t,x) \in [0,T] \times \mathbb{R}^d$ and $0 \leq T \leq T' < \infty$). Moreover, from the viscosity properties of v^T it follows that v is the unique uniformly continuous viscosity solution to equation (5.1). Then, from the probabilistic representation formula for v^T we deduce the representation formula (5.4) for v. $\Box$

In conclusion, we have proved the following result.

Proposition 5.1 *Let Assumptions* **(H1)**, **(H2)**, **(H**A**)**, *and* **(H**ϑ**)** *hold. Then, the function v given by* (5.4) *is the unique uniformly continuous viscosity solution to* (5.1) *on* $[0,\infty) \times \mathbb{R}^d$.

5.2 First convergence result: the proof of (5.2)

Let (λ, ϕ), with ϕ Lipschitz, be a viscosity solution to the ergodic equation (4.1). Let us introduce the function $w\colon [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ given by

$$w(T,x) \;:=\; v(T,x) - \big(\lambda T + \phi(x)\big), \qquad (T,x) \in [0,\infty) \times \mathbb{R}^d. \tag{5.8}$$

The aim is to state an upper and lower estimate for w, uniformly in time T, so that by dividing by T, we obtain the convergence of the long run average $v(T,.)/T$ to λ. Classical PDE arguments (in the case where f does not depend on y) rely on the smoothness of v and ϕ in order to prove that w is a sub and supersolution to some PDE without cost or gain function. Then by comparison principle, and under ergodicity conditions, one would obtain for w a lower and upper bound function which does not depend on time. In our general framework, the major difficulty is due to the non-regularity in general of v and ϕ, especially when there is singularity of the diffusion coefficient. In this case, it is not clear, even with the notion of viscosity solution, how to derive an equation for w involving the difference of v and ϕ. We circumvent this issue by adopting an alternative approach where we use probabilistic representations formulae for v and ϕ. We are also interested in the case where $f(x,a,y)$ depends on y, that we shall actually tackle by using the nondecreasing feature of f in y and imposing the following additional assumption.

(H3) The function f can be written as $f(x,a,y) = f_0(x,a) + f_1(x,a,y)$, where f_1 can be either the zero function or it satisfies, for all $x \in \mathbb{R}^d$, $a \in A$, $y, y' \in \mathbb{R}$,

$$y \;>\; y' \qquad \Longrightarrow \qquad f_1(x,a,y) - f_1(x,a,y') \;\leq\; -\kappa(y - y'),$$

for some constant $\kappa > 0$.

Theorem 5.1 *Let Assumptions* **(H1)**, **(H2)**, **(H3)**, **(HA)**, *and* **(Hϑ)** *hold. Then, there exists a positive constant C such that, the function w defined in* (5.8) *satisfies*

$$-C(1+|x|) \;\leq\; w(T,x) \;\leq\; C(1+|x|), \quad (T,x) \in [0,\infty) \times \mathbb{R}^d. \tag{5.9}$$

In particular, we have

$$\frac{v(T,x)}{T} \overset{T\to\infty}{\longrightarrow} \lambda. \tag{5.10}$$

Remark 5.2 We report here the *proof of Theorem 5.1 when $f = f(x,a)$ does not depend on y,* since it is much easier. Recall from Remark 5.1 that $v(t,x) = v^T(T-t,x)$, where v^T is the unique uniformly continuous viscosity solution to the Hamilton-Jacobi-Bellman equation (5.7). Therefore, v^T admits a stochastic control representation, which in terms of v reads

$$v(T,x) \;=\; \sup_{\alpha\in\mathcal{A}} \mathbb{E}\Big[\int_0^T f\big(X_s^{x,\alpha},\alpha_s\big)ds + h(X_T^{x,\alpha})\Big], \qquad \forall\,(T,x) \in [0,\infty)\times\mathbb{R}^d,$$

where $\mathcal{A}$ is the set of adapted control processes valued in A, and $(X_t^{x,\alpha})_{t\geq0}$ is the unique solution to the controlled equation (1.2) starting from x at time 0. Similarly, we know from the proof of Theorem 4.2 that ϕ is the unique uniformly continuous viscosity solution to equation (4.10), so that ϕ is given by

$$\phi(x) \;=\; \sup_{\alpha\in\mathcal{A}} \mathbb{E}\Big[\int_0^T \big[f\big(X_s^{x,\alpha},\alpha_s\big) - \lambda\big]ds + \phi(X_T^{x,\alpha})\Big], \qquad \forall\,(T,x) \in [0,\infty)\times\mathbb{R}^d.$$

From the definition of w, we have

$$\begin{aligned} w(T,x) \;&=\; \sup_{\alpha\in\mathcal{A}} \mathbb{E}\Big[\int_0^T f\big(X_s^{x,\alpha},\alpha_s\big)ds + h(X_T^{x,\alpha})\Big] - \sup_{\alpha\in\mathcal{A}} \mathbb{E}\Big[\int_0^T f\big(X_s^{x,\alpha},\alpha_s\big)ds + \phi(X_T^{x,\alpha})\Big] \\ &\leq\; \sup_{\alpha\in\mathcal{A}} \mathbb{E}\big[(h-\phi)(X_T^{x,\alpha})\big]. \end{aligned}$$

Proceeding in a similar way, we obtain

$$w(T,x) \;\geq\; \inf_{\alpha\in\mathcal{A}} \mathbb{E}\big[(h-\phi)(X_T^{x,\alpha})\big].$$

Since h and ϕ are Lipschitz, from estimate (2.4) we deduce (5.9). □

Proof of Theorem 5.1. We recall from (5.4) and Theorem 4.2 the nonlinear Feynman-Kac formulae

$$\phi(X_t^{x,a}) \;=\; \underline{Y}_t^{x,a}, \qquad v(T-t,X_t^{x,a}) \;=\; Y_t^{x,a,T}, \qquad 0\leq t\leq T,$$

for all $(T,x) \in [0,\infty)\times\mathbb{R}^d$, and any $a \in \mathring{A}$. Fix then $a \in \mathring{A}$, and define, for $(T,x) \in [0,\infty)\times\mathbb{R}^d$, the process:

$$\tilde{Y}_t^{x,a,T} \;:=\; Y_t^{x,a,T} - \lambda(T-t) - \underline{Y}_t^{x,a}, \qquad 0\leq t\leq T.$$

Then, by definition of w in (5.8), we have

$$w(T,x) \;=\; \tilde{Y}_0^{x,a,T}, \qquad \forall\,(T,x)\in[0,\infty)\times\mathbb{R}^d. \tag{5.11}$$

Moreover, from the BSDE (5.5)-(5.6) for $Y^{x,a,T}$ and (4.8)-(4.9) for $\underline{Y}^{x,a}$, we derive the BSDE for $\tilde{Y}^{x,a,T}$:

$$
\begin{aligned}
\tilde{Y}_t^{x,a,T} &= (h-\phi)(X_T^{x,a}) + \int_t^T \big(f\big(X_s^{x,a}, I_s^a, \frac{Y_s^{x,a,T}}{T-s+1}\big) - f(X_s^{x,a}, I_s^a, \lambda)\big)ds \\
&\quad + K_T^{x,a,T} - K_t^{x,a,T} - \big(\underline{K}_T^{x,a} - \underline{K}_t^{x,a}\big) - \int_t^T \big(Z_s^{x,a,T} - \underline{Z}_s^{x,a}\big)dW_s \\
&\quad - \int_t^T \int_A \big(U_s^{x,a,T}(a') - \underline{U}_s^{x,a}(a')\big)\tilde{\mu}(ds,da'), \qquad 0 \le t \le T,
\end{aligned}
$$

and

$$
\begin{aligned}
U_t^{x,a,T}(a') &\le 0, \qquad dt \otimes d\mathbb{P} \otimes \vartheta(da')\text{-a.e.} \\
\underline{U}_t^{x,a}(a') &\le 0, \qquad dt \otimes d\mathbb{P} \otimes \vartheta(da')\text{-a.e.}
\end{aligned}
$$

We shall now prove suitable upper and lower bounds for $\tilde{Y}^{x,a,T}$, and thus for $w(T,x)$.

• *Step 1. Upper bound:* $w(T,x) \le C(1+|x|)$. Let us consider the BSDE with nonpositive jumps on $[0,T]$:

$$
\begin{aligned}
\hat{Y}_t^{x,a,T} &= (h-\phi)(X_T^{x,a}) + \int_t^T \big(f\big(X_s^{x,a}, I_s^a, \frac{\hat{Y}_s^{x,a,T} + \lambda(T-s) + \phi(X_s^{x,a})}{T-s+1}\big) \\
&\quad - f(X_s^{x,a}, I_s^a, \lambda)\big)ds + \hat{K}_T^{x,a,T} - \hat{K}_t^{x,a,T} \\
&\quad - \int_t^T \hat{Z}_s^{x,a,T} dW_s - \int_t^T \int_A \hat{U}_s^{x,a,T}(a')\tilde{\mu}(ds,da')
\end{aligned} \tag{5.12}
$$

and

$$
\hat{U}_t^{x,a,T}(a') \le 0, \qquad dt \otimes d\mathbb{P} \otimes \vartheta(da')\text{-a.e.} \tag{5.13}
$$

We know from Theorem 2.1 in [19] that there exists a unique minimal solution ($\hat{Y}^{x,a,T}$, $\hat{Z}^{x,a,T}$, $\hat{U}^{x,a,T}$, $\hat{K}^{x,a,T}$) in $\mathbf{S^2(0,T)} \times \mathbf{L^2(W;0,T)} \times \mathbf{L^2}(\tilde{\mu};\mathbf{0,T}) \times \mathbf{K^2(0,T)}$ to equation (5.12)-(5.13). Set $\bar{Y}_t^{x,a,T} = \hat{Y}_t^{x,a,T} + \lambda(T-t) + \underline{Y}_t^{x,a}$, $t \in [0,T]$, and recall that $\underline{Y}_t^{x,a} = \phi(X_t^{x,a})$. Then, from the BSDEs satisfied by $\hat{Y}^{x,a,T}$ and $\underline{Y}^{x,a}$, we easily see that $(\bar{Y}^{x,a,T}, \hat{Z}^{x,a,T} + \underline{Z}^{x,a}, \hat{U}^{x,a,T} + \underline{U}^{x,a}, \hat{K}^{x,a,T} + \underline{K}^{x,a})$ is a solution to (5.5)-(5.6). From the minimality of $(Y^{x,a,T}, Z^{x,a,T}, U^{x,a,T}, K^{x,a,T})$, we get: $Y_t^{x,a,T} \le \bar{Y}_t^{x,a,T}$, and by subtracting to both sides $\lambda(T-t) + \underline{Y}_t^{x,a}$ we end up with

$$
\tilde{Y}_t^{x,a,T} \le \hat{Y}_t^{x,a,T}, \quad t \in [0,T]. \tag{5.14}
$$

Let us now derive an upper bound for $\hat{Y}_0^{x,a,T}$. To this end, we introduce the associated penalized BSDE with jumps on $[0,T]$, for $n \in \mathbb{N}$:

$$
\begin{aligned}
\hat{Y}_t^{x,a,T,n} &= (h-\phi)(X_T^{x,a}) + \int_t^T \big(f\big(X_s^{x,a}, I_s^a, \frac{\hat{Y}_s^{x,a,T,n} + \lambda(T-s) + \phi(X_s^{x,a})}{T-s+1}\big) \\
&\quad - f(X_s^{x,a}, I_s^a, \lambda)\big)ds + n\int_t^T \int_A (\hat{U}_s^{x,a,T,n}(a'))_+ \vartheta(da')ds \\
&\quad - \int_t^T \hat{Z}_s^{x,a,T,n} dW_s - \int_t^T \int_A \hat{U}_s^{x,a,T,n}(a')\tilde{\mu}(ds,da').
\end{aligned} \tag{5.15}
$$

From the uniform Lipschitz condition on $f(x,a,y)$ with respect to y, together with Assumptions **(H2)**(ii) and **(H3)**, we have

$$
\begin{aligned}
& f\big(X_s^{x,a}, I_s^a, \frac{\hat Y_s^{x,a,T,n} + \lambda(T-s) + \phi(X_s^{x,a})}{T-s+1}\big) \ - f(X_s^{x,a}, I_s^a, \lambda) \\
\leq \ & \rho_s^n\big(\hat Y_s^{x,a,T,n} + \phi(X_s^{x,a}) - \lambda\big), \qquad 0 \leq s \leq T,
\end{aligned}
\tag{5.16}
$$

with (we suppose here that f_1 in **(H3)** is not the zero function; otherwise, ρ^n can be taken equal to zero everywhere and the proof becomes easier)

$$
\rho_s^n \ = \ -\frac{\kappa}{T+1} 1_{\{\hat Y_s^{x,a,T,n} + \phi(X_s^{x,a}) - \lambda > 0\}} - L_2 1_{\{\hat Y_s^{x,a,T,n} + \phi(X_s^{x,a}) - \lambda \leq 0\}}, \qquad 0 \leq s \leq T. \tag{5.17}
$$

Then, applying Itô's formula to $e^{\int_0^t \rho_r^n dr} \hat Y_t^{x,a,T,n}$ between 0 and T, we get from (5.15):

$$
\begin{aligned}
\hat Y_0^{x,a,T,n} \ = \ & e^{\int_0^T \rho_r^n dr}(h-\phi)(X_T^{x,a}) - \int_0^T \rho_s^n e^{\int_0^s \rho_r^n dr} \hat Y_s^{x,a,T,n} ds \\
& + \int_0^T e^{\int_0^s \rho_r^n dr}\Big(f\big(X_s^{x,a}, I_s^a, \frac{\hat Y_s^{x,a,T,n} + \lambda(T-s) + \phi(X_s^{x,a})}{T-s+1}\big) \\
& - f(X_s^{x,a}, I_s^a, \lambda)\Big) ds + n \int_0^T \int_A e^{\int_0^s \rho_r^n dr}(\hat U_s^{x,a,T,n}(a'))_+ \vartheta(da') ds \\
& - \int_0^T e^{\int_0^s \rho_r^n dr} \hat Z_s^{x,a,T,n} dW_s - \int_0^T \int_A e^{\int_0^s \rho_r^n dr} \hat U_s^{x,a,T,n}(a') \tilde\mu(ds, da').
\end{aligned}
$$

Using (5.16), we obtain

$$
\begin{aligned}
\hat Y_0^{x,a,T,n} \ \leq \ & e^{\int_0^T \rho_r^n dr}(h-\phi)(X_T^{x,a}) + \int_0^T \rho_s^n e^{\int_0^s \rho_r^n dr}\big(\phi(X_s^{x,a}) - \lambda\big) ds \\
& + n \int_0^T \int_A e^{\int_0^s \rho_r^n dr}(\hat U_s^{x,a,T,n}(a'))_+ \vartheta(da') ds \\
& - \int_0^T e^{\int_0^s \rho_r^n dr} \hat Z_s^{x,a,T,n} dW_s - \int_0^T \int_A e^{\int_0^s \rho_r^n dr} \hat U_s^{x,a,T,n}(a') \tilde\mu(ds, da').
\end{aligned}
\tag{5.18}
$$

Now, from Proposition 2.1 in [19] we have the following dual representation formula for the right-hand side of (5.18):

$$
\begin{aligned}
& e^{\int_0^T \rho_r^n dr}(h-\phi)(X_T^{x,a}) + \int_0^T \rho_s^n e^{\int_0^s \rho_r^n dr}\big(\phi(X_s^{x,a}) - \lambda\big) ds \\
& + n \int_0^T \int_A e^{\int_0^s \rho_r^n dr}(\hat U_s^{x,a,T,n}(a'))_+ \vartheta(da') ds \\
& - \int_0^T e^{\int_0^s \rho_r^n dr} \hat Z_s^{x,a,T,n} dW_s - \int_0^T \int_A e^{\int_0^s \rho_r^n dr} \hat U_s^{x,a,T,n}(a') \tilde\mu(ds, da') \\
& = \ \sup_{\nu \in \mathcal V} \mathbb E^\nu\Big[e^{\int_0^T \rho_r^n dr}(h-\phi)(X_T^{x,a}) + \int_0^T \rho_s^n e^{\int_0^s \rho_r^n dr}\big(\phi(X_s^{x,a}) - \lambda\big) ds\Big].
\end{aligned}
$$

Therefore, we get

$$
\hat Y_0^{x,a,T,n} \ \leq \ \sup_{\nu \in \mathcal V} \mathbb E^\nu\Big[e^{\int_0^T \rho_r^n dr}(h-\phi)(X_T^{x,a}) + \int_0^T \rho_s^n e^{\int_0^s \rho_r^n dr}\big(\phi(X_s^{x,a}) - \lambda\big) ds\Big].
$$

From the definition of ρ^n in (5.17), we find

$$\hat{Y}_0^{x,a,T,n} \leq \sup_{\nu\in\mathcal{V}} \mathbb{E}^\nu\bigg[|h-\phi|(X_T^{x,a}) + \max\Big(\frac{\kappa}{T+1}, L_2\Big)\int_0^T e^{-\min(\frac{\kappa}{T+1},L_2)s}|\phi(X_s^{x,a})-\lambda|ds\bigg].$$

Recalling that h and ϕ are Lipschitz, from estimate (2.4) we obtain

$$\hat{Y}_0^{x,a,T,n} \leq C(1+|x|),$$

for some positive constant C, independent of x, a, T, and n. Since from Theorem 2.1 in [19] we have that $\hat{Y}_0^{x,a,T,n}$ converges to $\hat{Y}_0^{x,a,T}$, as n goes to infinity, we get the same estimate: $\hat{Y}_0^{x,a,T} \leq C(1+|x|)$, and therefore, from (5.11) and (5.14), we deduce that

$$w(T,x) \leq C(1+|x|).$$

• *Step 2. Lower bound:* $w(T,x) \geq -C(1+|x|)$*.* As in step 1, where we built an upper bound for $\tilde{Y}^{x,a,T}$ using the minimality of $Y^{x,a,T}$, here we shall construct a lower bound for $\tilde{Y}^{x,a,T}$ exploiting the minimality of $\underline{Y}^{x,a}$ in the sense of Definition 4.2. In particular, we fix $T>0$ and we recall that $(\underline{Y}^{x,a}_{|[0,T]}, \underline{Z}^{x,a}_{|[0,T]}, \underline{U}^{x,a}_{|[0,T]}, \underline{K}^{x,a}_{|[0,T]})$ is the minimal solution to (4.6)-(4.7) on $[0,T]$ with terminal condition $\phi(X_T^{x,a})$. Now, let us consider the BSDE with nonnegative jumps on $[0,T]$:

$$\begin{aligned} \check{Y}_t^{x,a,T} &= (h-\phi)(X_T^{x,a}) + \int_t^T \Big(f\big(X_s^{x,a}, I_s^a, \frac{v(T-s,X_s^{x,a})}{T-s+1}\big) - f(X_s^{x,a}, I_s^a, \lambda)\Big)ds \qquad (5.19)\\ &\quad - \big(\check{K}_T^{x,a,T} - \check{K}_t^{x,a,T}\big) - \int_t^T \check{Z}_s^{x,a,T}dW_s - \int_t^T\int_A \check{U}_s^{x,a,T}(a')\tilde{\mu}(ds,da') \end{aligned}$$

and

$$\check{U}_t^{x,a,T}(a') \geq 0, \qquad dt\otimes d\mathbb{P}\otimes\vartheta(da')\text{-a.e.} \qquad (5.20)$$

Theorem 2.1 in [19] gives the existence of a unique maximal solution $(\check{Y}^{x,a,T}, \check{Z}^{x,a,T}, \check{U}^{x,a,T}, \check{K}^{x,a,T}) \in \mathbf{S^2(0,T)}\times\mathbf{L^2(W;0,T)}\times\mathbf{L^2(\tilde{\mu};0,T)}\times\mathbf{K^2(0,T)}$ to (5.19)-(5.20). Actually, Theorem 2.1 in [19] is designed for minimal solutions, while here we deal with maximal solutions; however, it is easy to show that $-\check{Y}^{x,a,T}$ is a minimal solution to a certain BSDE with nonpositive jumps, therefore we can apply Theorem 2.1 to $-\check{Y}^{x,a,T}$. Set $\overline{Y}_t^{x,a,T} = -\check{Y}_t^{x,a,T} + Y_t^{x,a,T} - \lambda(T-t)$, $t\in[0,T]$, then

$$\begin{aligned} &(\overline{Y}^{x,a,T}, \overline{Z}^{x,a,T}, \overline{U}^{x,a,T}, \overline{K}^{x,a,T})\\ &:= (\overline{Y}^{x,a,T}, -\check{Z}^{x,a,T} + Z^{x,a,T}, -\check{U}^{x,a,T} + U^{x,a,T}, \check{K}^{x,a,T} + K^{x,a,T}) \end{aligned}$$

is a solution to (4.6)-(4.7) on $[0,T]$ with terminal condition $\phi(X_T^{x,a})$. From the already mentioned minimality of $(\underline{Y}^{x,a}_{|[0,T]}, \underline{Z}^{x,a}_{|[0,T]}, \underline{U}^{x,a}_{|[0,T]}, \underline{K}^{x,a}_{|[0,T]})$ to (4.6)-(4.7), we see that $\underline{Y}_t^{x,a} \leq \overline{Y}_t^{x,a,T}$, and by subtracting to both sides $Y_t^{x,a,T} - \lambda(T-t)$, we end up with

$$\check{Y}_t^{x,a,T} \leq \tilde{Y}_t^{x,a,T}, \quad t\in[0,T]. \qquad (5.21)$$

We now derive a lower bound for $\check{Y}_0^{x,a,T}$ by means of a dual representation formula. In particular, we see from Theorem 2.2 in [19] that $\check{Y}_0^{x,a,T}$ admits the dual representation formula (observe that, as for Theorem 2.1 in [19], Theorem 2.2 in [19] is designed for minimal solutions, while here we deal with maximal solutions; however, since $-\check{Y}^{x,a,T}$ is a minimal solution to a certain BSDE with nonpositive jumps, we can apply Theorem 2.2 to $-\check{Y}^{x,a,T}$)

$$\check{Y}_0^{x,a,T} = \inf_{\nu\in\mathcal{V}} \mathbb{E}^\nu\Big[(h-\phi)(X_T^{x,a}) + \int_0^T \big(f\big(X_s^{x,a}, I_s^a, \frac{v(T-s, X_s^{x,a})}{T-s+1}\big) - f(X_s^{x,a}, I_s^a, \lambda)\big)ds\Big].$$

From the Lipschitz property of h and ϕ, and estimate (2.4), we have

$$\inf_{\nu\in\mathcal{V}} \mathbb{E}^\nu\big[(h-\phi)(X_T^{x,a})\big] \;\geq\; -C(1+|x|). \tag{5.22}$$

Moreover, from the uniform Lipschitz condition on $f(x,a,y)$ with respect to y, and the nondecreasing property of $y \mapsto f(x,a,y)$ in **(H2)**, there exists some adapted, nonpositive, and bounded process ζ such that

$$f\big(X_s^{x,a}, I_s^a, \frac{v(T-s, X_s^{x,a})}{T-s+1}\big) - f(X_s^{x,a}, I_s^a, \lambda) \;=\; \zeta_s\Big(\frac{v(T-s, X_s^{x,a})}{T-s+1} - \lambda\Big),$$

for all $0 \leq s \leq T$. Therefore, we have

$$\begin{aligned} &\inf_{\nu\in\mathcal{V}} \mathbb{E}^\nu\Big[\int_0^T \big(f\big(X_s^{x,a}, I_s^a, \frac{v(T-s, X_s^{x,a})}{T-s+1}\big) - f(X_s^{x,a}, I_s^a, \lambda)\big)ds\Big] \\ &= \inf_{\nu\in\mathcal{V}} \mathbb{E}^\nu\Big[\int_0^T \zeta_s\Big(\frac{v(T-s, X_s^{x,a})}{T-s+1} - \lambda\Big)ds\Big]. \end{aligned} \tag{5.23}$$

From step 1 and the Lipschitz property of ϕ, it follows that $v(T,x) - \lambda T \leq C(1+|x|)$, and consequently $\frac{v(T,x)}{T+1} - \lambda \leq C(1+|x|)$. Hence, since ζ is nonpositive,

$$\zeta_s\Big(\frac{v(T-s, X_s^{x,a})}{T-s+1} - \lambda\Big) \;\geq\; \zeta_s C(1+|X_s^{x,a}|) \;\geq\; -L_2 C(1+|X_s^{x,a}|),$$

where we used the fact that ζ is bounded by L_2, the Lipschitz constant of f. Plugging the above estimate into (5.23) combined with (2.4), and recalling (5.22), we find

$$\check{Y}_0^{x,a,T} \;\geq\; -C(1+|x|).$$

From (5.11) and (5.21), we conclude that $w(T,x) \geq -C(1+|x|)$. □

Remark 5.3 (i) From Theorem 5.1 and Remark 4.2, we deduce a uniqueness result for the component λ of a minimal solution to the ergodic BSDE with nonpositive jumps (4.4)-(4.5). Indeed, consider a family of minimal solutions to (4.4)-(4.5) as in Remark 4.2. Then, from Theorem 5.1 we see that λ is given by (5.10).
(ii) Let $f \in \mathbf{L}^1_{\mathrm{loc}}([0,\infty);\mathbb{R})$ be such that $\int_0^\infty e^{-\beta t} f(t)dt$ exists for $\beta > 0$. A theorem which states that, under certain conditions on f, if $\lim_{\beta\to 0^+} \beta\int_0^\infty e^{-\beta t} f(t)dt = f_\infty \in \mathbb{R}$ then $\lim_{T\to\infty} \frac{1}{T}\int_0^T f(t)dt = f_\infty$, is called a *Tauberian theorem*, see, e.g., [1]. In our paper, Theorem 4.1 together with Theorem 5.1 can be thought as a "robust" Tauberian theorem. Indeed, in Theorem 4.1 we proved that the convergence of $\beta_k v^{\beta_k}(x) = \beta_k \sup_\alpha \mathbb{E}_x[\int_0^\infty e^{-\beta_k t} f(X_t^\alpha, \alpha_t)dt]$ towards λ, and also of $v^{\beta_k}(x) - v^{\beta_k}(0)$ towards ϕ, allows to construct a viscosity solution (λ, ϕ) to the ergodic equation (4.1). Then, Theorem 5.1 implies the convergence of $\frac{v(T,x)}{T} = \frac{1}{T}\sup_\alpha \mathbb{E}_x[\int_0^T f(X_t^\alpha, \alpha_t)dt + h(X_T^\alpha)]$ towards λ.

□

5.3 Further convergence result via verification: the proof of (5.3)

We conclude this section by presenting, in the form of a verification theorem, the following result, which shows the validity of the convergence (5.3).

Theorem 5.2 *Let Assumptions* **(H1)**, **(H2)**, **(H**A**)**, *and* **(H**ϑ**)** *hold. Suppose that:*

(i) (λ, ϕ), *with* $\phi \in C^2(\mathbb{R}^d)$ *and Lipschitz, is a classical solution to the ergodic equation* (4.1).

(ii) *In the ergodic equation* (4.1), *the supremum is attained at* $a = \underline{\alpha}(x)$, *for every* $x \in \mathbb{R}^d$, *for some function* $\underline{\alpha}\colon A \to \mathbb{R}^d$ *such that* $\underline{b}, \underline{\sigma}$ *in* (2.8) *satisfy Assumption* **(H1)**.

Consider the unique (viscosity) solution v *to* (5.4). *Then, there exists a real constant* c *such that*

$$v(T,x) - (\lambda T + \phi(x)) \stackrel{T\to\infty}{\longrightarrow} c,$$

for all $x \in \operatorname{supp}\rho$, *the support of the invariant measure* ρ *given by Proposition 2.1. In particular, when* $\operatorname{supp}\rho = \mathbb{R}^d$ *we deduce that* ϕ *is uniquely determined up to a constant.*

Remark 5.4 The existence of a smooth solution (λ, ϕ) to the ergodic equation (4.1) is ensured under a uniform ellipticity condition, see, e.g., Theorem 1.1 in [27]. More precisely, suppose that assumptions **(H1)**, **(H2)**, **(H**A**)**, and **(H**ϑ**)** hold and let (λ, ϕ) be a viscosity solution to (4.1) with ϕ Lipschitz, whose existence follows from Theorem 4.1. Then, to exploit Theorem 1.1 in [27], we fix $\delta > 0$ and we consider the elliptic equation in the unknown ψ on the bounded domain $B_R \subset \mathbb{R}^d$ (the open ball of radius $R > 0$ centered at the origin)

$$\begin{aligned} \delta\psi(x) - \sup_{a\in A}\big[\mathcal{L}^a\psi + f(x,a,\lambda) - \lambda + \delta\phi(x)\big] &= 0, \quad \text{on } B_R, &(5.24)\\ \psi &= \phi, \quad \text{on } \partial B_R. &(5.25)\end{aligned}$$

Notice that, thanks to the presence of the term "$\delta\psi(x)$" in equation (5.24), we can apply comparison Theorem 3.3 in [17], from which it follows that ϕ is the unique uniformly continuous viscosity solution to equation (5.24)-(5.25). Let us now impose the following uniform ellipticity condition: there exists $\nu \in (0,1]$, possibly depending on R, such that

$$\nu|\xi|^2 \;\le\; \sum_{i,j=1}^d (\sigma\sigma^\intercal)_{ij}(x,a)\xi_i\xi_j \;\le\; \nu^{-1}|\xi|^2, \qquad \forall\, \xi \in \mathbb{R}^d,$$

for all $x \in \overline{B}_R$ and $a \in A$. Then, as explained in Remark 1.1 of [27], under the above assumption, Theorem 1.1 in [27] holds, and there exists a unique solution $\psi \in C^2(B_R) \cap C(\overline{B}_R)$ to equation (5.24)-(5.25). Theorem 3.3 in [17] implies that ψ coincides with our function ϕ, so that $\phi \in C^2(B_R)$. Since R is arbitrary, we conclude that $\phi \in C^2(\mathbb{R}^d)$. □

Proof. *Step 1.* Notice that, for any $T, S > 0$ and for all $x \in \mathbb{R}^d$, we have

$$v(T+S,x) \;=\; \sup_{\alpha\in\mathcal{A}} \mathbb{E}\bigg[\int_0^T f\Big(X_s^{x,\alpha}, \alpha_s, \frac{v(T+S-s, X_s^{x,\alpha})}{T+S-s+1}\Big)ds + v(S, X_T^{x,\alpha})\bigg], \tag{5.26}$$

where $\mathcal{A}$ is the set of adapted control processes valued in A, and $(X_t^{x,\alpha})_{t\geq 0}$ is the unique solution to the controlled equation (1.2) starting from x at time 0. As a matter of fact, to prove (5.26) we recall from Remark 5.1 that $v(T+S,x) = v^{T+S}(0,x)$, for all $x \in \mathbb{R}^d$, where v^{T+S} is the unique uniformly continuous viscosity solution to the following Hamilton-Jacobi-Bellman equation in the unknown $\tilde{v}^{T+S}$:

$$
\begin{aligned}
-\frac{\partial \tilde{v}^{T+S}}{\partial t} - \sup_{a\in A}\big[\mathcal{L}^a \tilde{v}^{T+S}(t,x) + \ f\big(x,a,\frac{v^{T+S}}{T+S-t+1}\big)\big] \;&=\; 0, \qquad (t,x)\in[0,T]\times\mathbb{R}^d,\\
\tilde{v}^{T+S}(T,x) \;&=\; v^{T+S}(T,x), \qquad x\in\mathbb{R}^d.
\end{aligned}
$$

As a consequence, v^{T+S} is given, for all $(t,x)\in[0,T]\times\mathbb{R}^d$, by the stochastic control representation:

$$
v^{T+S}(t,x) \;=\; \sup_{\alpha\in\mathcal{A}} \mathbb{E}\bigg[\int_0^{T-t} f\big(X_s^{x,\alpha},\alpha_s,\frac{v^{T+S}(s+t,X_s^{x,\alpha})}{T+S-s-t+1}\big)ds + v^{T+S}(T,X_{T-t}^{x,\alpha})\bigg],
$$

which implies (5.26). In particular, we have

$$
v(T+S,x) \;\geq\; \mathbb{E}\bigg[\int_0^T f\big(X_s^{x,\underline{\alpha}},\underline{\alpha}(X_s^{x,\underline{\alpha}}),\frac{v(T+S-s,X_s^{x,\underline{\alpha}})}{T+S-s+1}\big)ds + v(S,X_T^{x,\underline{\alpha}})\bigg].
$$

On the other hand, applying Itô's formula to $\phi(X_s^{x,\underline{\alpha}})$ between 0 and T, and using the optimality of $\underline{\alpha}$ in the ergodic equation (4.1), we obtain

$$
\phi(x) \;=\; \mathbb{E}\bigg[\int_0^T f(X_s^{x,\underline{\alpha}},\underline{\alpha}(X_s^{x,\underline{\alpha}}),\lambda)ds - \lambda T + \phi(X_T^{x,\underline{\alpha}})\bigg].
$$

Therefore, w in (5.8) satisfies

$$
\begin{aligned}
w(T+S,x) \;\geq\; \mathbb{E}\bigg[&w(S,X_T^{x,\underline{\alpha}}) \\
&+ \int_0^T f\big(X_s^{x,\underline{\alpha}},\underline{\alpha}(X_s^{x,\underline{\alpha}}),\frac{v(T+S-s,X_s^{x,\underline{\alpha}})}{T+S-s+1}\big) - f(X_s^{x,\underline{\alpha}},\underline{\alpha}(X_s^{x,\underline{\alpha}}),\lambda)ds\bigg].
\end{aligned}
\tag{5.27}
$$

Step 2. Let us prove that there exists a positive constant C such that

$$
|w(T,x)-w(T,x')| \;\leq\; C|x-x'|,
$$

for all $T\geq 0$ and $x,x'\in\mathbb{R}^d$. Recalling (5.8) and since ϕ is Lipschitz, it is therefore enough to prove that the function v satisfies

$$
|v(T,x)-v(T,x')| \;\leq\; C|x-x'|, \qquad T\geq 0,\ x\in\mathbb{R}^d, \tag{5.28}
$$

for some positive constant C. We know that $v(T,x) = Y_0^{x,a,T}$ is represented by the minimal solution to the BSDE with nonpositive jumps (5.5)-(5.6) on $[0,T]$. We recall from Theorem 2.1 in [19] that $Y^{x,a,T,n}\uparrow Y^{x,a,T}$, where $(Y^{x,a,T,n},Z^{x,a,T,n},U^{x,a,T,n})\in\mathbf{S^2(0,T)}\times\mathbf{L^2(W;0,T)}\times\mathbf{L^2(\tilde{\mu};0,T)}$ is the solution to the penalized BSDE on $[0,T]$:

$$
Y_t^{x,a,T,n} \;=\; h(X_T^{x,a}) + \int_t^T f(X_s^{x,a},I_s^a,\frac{Y_s^{x,a,T,n}}{T-s+1})ds + n\int_t^T\int_A \big(U_s^{x,a,T,n}(a')\big)_+\vartheta(da')ds
$$

$$-\int_t^T Z_s^{x,a,T,n}dW_s - \int_t^T \int_A U_s^{x,a,T,n}(a')\tilde{\mu}(ds,da'), \qquad 0 \le t \le T$$

and

$$U_t^{x,a,T,n}(a') \ \le\ 0, \qquad dt \otimes d\mathbb{P} \otimes \vartheta(da')\text{-a.e.}$$

Then, (5.28) follows once we get:

$$|Y_0^{x,a,T,n} - Y_0^{x',a,T,n}| \ \le\ C|x-x'|,$$

for a constant C that does not depend on x, a, T, and n. This can be done using Girsanov theorem and the dissipativity condition (2.2) in the same way as for (3.11).

Step 3. Now, we proceed as in [15] and we introduce the set Γ which contains all the ω-limits of the family $\{w(T,\cdot)\}_{T>1}$ in $C(\mathbb{R}^d)$ (we endow $C(\mathbb{R}^d)$ with the topology for which $f_j \to f$ in $C(\mathbb{R}^d)$ if and only if f_j converges uniformly to f on any compact subset of $\mathbb{R}^d$). In other words, Γ is given by

$$\Gamma \ :=\ \big\{w_\infty \in C(\mathbb{R}^d)\colon\ w(T_j,\cdot) \to w_\infty \text{ in } C(\mathbb{R}^d) \text{ for some } (T_j)_{j\in\mathbb{N}} \text{ with } T_j \to \infty\big\}.$$

It follows from step 2 that the family $\{w(T,\cdot)\}_{T>1}$ is relatively compact in $C(\mathbb{R}^d)$. In particular, $\Gamma \neq \emptyset$, and any w_∞ in Γ is Lipschitz. To conclude, it suffices to prove that every $w_\infty \in \Gamma$ is equal to the same constant $c \in \mathbb{R}$ on $\operatorname{supp}\rho$.

- *Step 3a.* We first show that any element of Γ is constant on $\operatorname{supp}\rho$. Let $w_\infty \in \Gamma$, therefore there exists a sequence $(T_j)_{j\in\mathbb{N}}$, with $T_j \to \infty$, such that $w(T_j,\cdot) \to w_\infty$ in $C(\mathbb{R}^d)$ as $j \to \infty$. From (5.27) with $S = T_j - T$, we have

$$\begin{aligned} w(T_j,x) \ \ge\ \mathbb{E}\Big[& w(T_j - T, X_T^{x,\underline{\alpha}}) \\ & + \int_0^T f\big(X_s^{x,\underline{\alpha}}, \underline{\alpha}(X_s^{x,\underline{\alpha}}), \frac{v(T_j - s, X_s^{x,\underline{\alpha}})}{T_j - s + 1}\big) - f(X_s^{x,\underline{\alpha}}, \underline{\alpha}(X_s^{x,\underline{\alpha}}), \lambda)ds\Big]. \end{aligned} \tag{5.29}$$

From (5.10) we have for all $s \in [0,T]$,

$$\frac{v(T_j - s, X_s^{x,\underline{\alpha}})}{T_j - s + 1} \ \overset{j\to\infty}{\longrightarrow}\ \lambda, \qquad \mathbb{P}\text{-a.s.}$$

Therefore, sending $j \to \infty$ in (5.29), and by the dominated convergence theorem, we obtain

$$w_\infty(x) \ \ge\ \mathbb{E}\big[w_\infty(X_T^{x,\underline{\alpha}})\big].$$

Moreover, choosing $T := T_j$ and letting $j \to \infty$ in the above inequality, we obtain, from Proposition 2.1:

$$w_\infty(x) \ \ge\ \int_{\mathbb{R}^d} w_\infty(x')\rho(dx').$$

Now, taking the infimum with respect to $x \in \mathbb{R}^d$, we end up with

$$0 \ \ge\ \int_{\mathbb{R}^d} (w_\infty(x') - \inf_{\mathbb{R}^d} w_\infty)\rho(dx') \ \ge\ 0.$$

As a consequence $w_\infty = \inf_{\mathbb{R}^d} w_\infty$, ρ-a.s., therefore w_∞ is constant on $\operatorname{supp}\rho$.

- *Step 3b.* We next prove that every $w_\infty \in \Gamma$ is equal to the same constant c on $\operatorname{supp}\rho$. Proceeding as in the derivation of (5.27), we have:

$$
\begin{aligned}
& \mathbb{E}\big[w(R+S, X^{x,\underline{\alpha}}_{T-R}\big] && (5.30)\\
\geq\ & \mathbb{E}\Big[w(S, X^{x,\underline{\alpha}}_T) + \int_R^T f\big(X^{x,\underline{\alpha}}_s, \underline{\alpha}(X^{x,\underline{\alpha}}_s), \frac{v(T+S-s, X^{x,\underline{\alpha}}_s)}{T+S-s+1}\big) - f(X^{x,\underline{\alpha}}_s, \underline{\alpha}(X^{x,\underline{\alpha}}_s), \lambda)ds\Big].
\end{aligned}
$$

for any $T, S, R > 0$ with $R \leq T$. Suppose that there exist two real constants c_1, c_2 and two diverging sequences $(T_j)_{j\in\mathbb{N}}$ and $(S_j)_{j\in\mathbb{N}}$ such that $w(T_j, \cdot) \to c_1$ and $w(S_j, \cdot) \to c_2$ on $\operatorname{supp}\rho$ as $j \to \infty$. Let us take $T := T_j$, $R = T_j - S$, and $S := S_k$ in (5.30). Then, letting $j \to \infty$ we obtain (notice that, by (5.30) and Lebesgue's dominated convergence theorem, the two integral terms in (5.30) simplify one with the other as $j \to \infty$)

$$
c_1 \ \geq \ \int_{\mathbb{R}^d} w(S_k, x')\rho(dx').
$$

Now, sending $k \to \infty$, we find

$$
c_1 \ \geq \ \lim_{k\to\infty} \int_{\mathbb{R}^d} w(S_k, x')\rho(dx') \ = \ \int_{\mathbb{R}^d} C_2\rho(dx') \ = \ c_2.
$$

Therefore $c_1 \geq c_2$. Changing the role of $(T_j)_{j\in\mathbb{N}}$ and $(S_j)_{j\in\mathbb{N}}$, we also find $c_2 \geq c_1$. Hence, $c_1 = c_2$ and the claim follows. □

Remark 5.5 Under the hypotheses of Theorem 5.2 and when $f = f(x, a)$ does not depend on y, λ can be interpreted as value of an *ergodic control problem* with gain functional

$$
J(x, \alpha) \ := \ \limsup_{T\to\infty} \mathbb{E}\Big[\int_0^T f(X^{x,\alpha}_t, \alpha_t)dt + h(X^{x,\alpha}_T)\Big],
$$

where $X^{x,\alpha}$ is the controlled diffusion process satisfying (1.2), starting from $x \in \mathbb{R}^d$ at time 0, and $\alpha \in \mathcal{A}$ is a control process, i.e., an A-valued adapted process. More precisely, it is clear that $J(x, \alpha)$ depends only on the asymptotic behavior of the trajectories of $X^{x,\alpha}$. Therefore, from the ergodicity of $X^{x,\alpha}$, we expect that there exists a real number λ^*, independent of $x \in \mathbb{R}^d$, such that

$$
\lambda^* \ := \ \sup_{\alpha\in\mathcal{A}} J(x, \alpha), \qquad \forall\, x \in \mathbb{R}^d,
$$

namely, λ^* is the value of the ergodic control problem. Let us prove that $\lambda^* = \lambda$. Firtsly, observe that, since f does not depend on y, the function v in (5.4) admits the stochastic control representation

$$
v(T, x) \ = \ \sup_{\alpha\in\mathcal{A}} \mathbb{E}\Big[\int_0^T f(X^{x,\alpha}_t, \alpha_t)dt + h(X^{x,\alpha}_T)\Big].
$$

From (5.10) we know that, for any $x \in \mathbb{R}^d$,

$$
\lambda \ = \ \lim_{T\to\infty} \frac{v(T, x)}{T} \ = \ \lim_{T\to\infty} \sup_{\alpha\in\mathcal{A}} \frac{1}{T}\mathbb{E}\Big[\int_0^T f(X^{x,\alpha}_t, \alpha_t)dt + h(X^{x,\alpha}_T)\Big]
$$

$$= \lim_{T\to\infty} \sup_{\alpha\in\mathcal{A}} \frac{1}{T}\mathbb{E}\Big[\int_0^T f(X_t^{x,\alpha},\alpha_t)dt\Big], \tag{5.31}$$

where the last equality follows from the fact that $\lim_{T\to\infty}\sup_{\alpha\in\mathcal{A}} \frac{1}{T}\mathbb{E}[h(X_T^{x,\alpha})] = 0$, which is a consequence of the Lipschitz property of h and estimate (2.4). From (5.31) we see that $\lambda^* \leq \lambda$. To prove the reverse inequality, fix $x \in \mathbb{R}^d$, then, applying Itô's formula to $\phi(X_t^{x,\underline{\alpha}})$ between 0 and T, and using the optimality of $\underline{\alpha}$ in the ergodic equation (4.1), we obtain

$$\lambda \;=\; \frac{1}{T}\mathbb{E}\Big[\int_0^T f(X_t^{x,\underline{\alpha}},\underline{\alpha}(X_t^{x,\underline{\alpha}}))dt + \phi(X_T^{x,\underline{\alpha}}) - \phi(x)\Big].$$

From the Lipschitz property of ϕ and estimate (2.4), we have $\frac{1}{T}\mathbb{E}[\phi(X_T^{x,\underline{\alpha}}) - \phi(x)] \to 0$ as $T \to \infty$, therefore

$$\begin{aligned}\lambda \;&=\; \lim_{T\to\infty}\frac{1}{T}\mathbb{E}\Big[\int_0^T f(X_t^{x,\underline{\alpha}},\underline{\alpha}(X_t^{x,\underline{\alpha}}))dt\Big]\\ &\leq\; \sup_{\alpha\in\mathcal{A}}\Big\{\limsup_{T\to\infty}\frac{1}{T}\mathbb{E}\Big[\int_0^T f(X_t^{x,\alpha},\alpha_t)dt\Big]\Big\} \;=\; \lambda^*,\end{aligned}$$

which implies that $\lambda^* = \lambda$. $\square$

Appendix

A Ergodicity proofs

A.1 Proof of Lemma 2.1

• Proof of (i)
Let $t \geq 0$, then an application of Itô's formula to $e^{\gamma s}|X_s^{x,a}|^2$ between 0 and t yields

$$\begin{aligned}e^{\gamma t}|X_t^{x,a}|^2 \;&=\; |x|^2 + \gamma\int_0^t e^{\gamma s}|X_s^{x,a}|^2 ds + 2\int_0^t e^{\gamma s}X_s^{x,a}.b(X_s^{x,a},I_s^a)ds\\ &\quad + \int_0^t e^{\gamma s}\|\sigma(X_s^{x,a},I_s^a)\|^2 ds + 2\int_0^t e^{\gamma s}(X_s^{x,a})^\intercal\sigma(X_s^{x,a},I_s^a)dW_s.\end{aligned}$$

Rearranging the terms in a suitable way so to exploit the dissipativity condition **(H1)**(ii), we obtain

$$\begin{aligned}e^{\gamma t}|X_t^{x,a}|^2 \;&=\; |x|^2 + \gamma\int_0^t e^{\gamma s}|X_s^{x,a}|^2 ds + 2\int_0^t e^{\gamma s}X_s^{x,a}.(b(X_s^{x,a},I_s^a) - b(0,I_s^a))ds\\ &+ \int_0^t e^{\gamma s}\mathrm{tr}\big[\big(\sigma(X_s^{x,a},I_s^a) - \sigma(0,I_s^a)\big)\big(\sigma(X_s^{x,a},I_s^a) - \sigma(0,I_s^a)\big)^\intercal\big]ds\\ &+ 2\int_0^t e^{\gamma s}X_s^{x,a}.b(0,I_s^a)ds + 2\int_0^t e^{\gamma s}\mathrm{tr}\big[\sigma(0,I_s^a)\big(\sigma(X_s^{x,a},I_s^a) - \sigma(0,I_s^a)\big)^\intercal\big]ds\\ &+ \int_0^t e^{\gamma s}\mathrm{tr}\big[\sigma(0,I_s^a)\sigma(0,I_s^a)^\intercal\big]ds + 2\int_0^t e^{\gamma s}(X_s^{x,a})^\intercal\sigma(X_s^{x,a},I_s^a)dW_s.\end{aligned}$$

Using **(H1)**, we find

$$e^{\gamma t}|X_t^{x,a}|^2 \;\leq\; |x|^2 + \gamma\int_0^t e^{\gamma s}|X_s^{x,a}|^2 ds - 2\gamma\int_0^t e^{\gamma s}|X_s^{x,a}|^2 ds + 2M_1\int_0^t e^{\gamma s}|X_s^{x,a}|ds$$

$$+ 2M_1L_1\int_0^t e^{\gamma s}|X_s^{x,a}|ds + M_1^2\int_0^t e^{\gamma s}ds + 2\int_0^t e^{\gamma s}(X_s^{x,a})^\intercal\sigma(X_s^{x,a},I_s^a)dW_s,$$

where $M_1 := \sup_{a\in A}(|b(0,a)| + \|\sigma(0,a)\|)$. From the inequality $|X_s^{x,a}| \leq \varepsilon|X_s^{x,a}|^2 + 1/(4\varepsilon)$, for any $\varepsilon > 0$, we obtain

$$\begin{aligned} e^{\gamma t}|X_t^{x,a}|^2 \ &\leq\ |x|^2 - (\gamma - 2M_1\varepsilon - 2M_1L_1\varepsilon)\int_0^t e^{\gamma s}|X_s^{x,a}|^2 ds \\ &\quad + \left(\frac{M_1 + M_1L_1}{2\varepsilon} + M_1^2\right)\int_0^t e^{\gamma s}ds + 2\int_0^t e^{\gamma s}(X_s^{x,a})^\intercal\sigma(X_s^{x,a},I_s^a)dW_s. \qquad \text{(A.1)} \end{aligned}$$

We can find ε such that $\gamma - 2M_1\varepsilon - 2M_1L_1\varepsilon \geq 0$ (more precisely, if $M_1 = 0$ then ε can be any positive real number; otherwise we take $\varepsilon \leq \gamma/(2M_1 + 2M_1L_1)$), therefore (also multiplying both sides in (A.1) by $e^{-\gamma t}$)

$$\begin{aligned} |X_t^{x,a}|^2 \ &\leq\ e^{-\gamma t}|x|^2 + \left(\frac{M_1 + M_1L_1}{2\varepsilon} + M_1^2\right)\frac{1 - e^{-\gamma t}}{\gamma} \\ &\quad + 2e^{-\gamma t}\int_0^t e^{\gamma s}(X_s^{x,a})^\intercal\sigma(X_s^{x,a},I_s^a)dW_s. \qquad \text{(A.2)} \end{aligned}$$

Now, consider $\nu \in \mathcal{V}$ and recall that W remains a Brownian motion under $\mathbb{P}^\nu$. Then, the following well-known estimate holds under **(H1)**(i): for all $T > 0$ and $p \geq 1$, there exists some positive constant $\bar{C}_{T,p}$ such that

$$\sup_{\nu\in\mathcal{V}}\mathbb{E}^\nu\Big[\sup_{0\leq s\leq T}|X_s^{x,a}|^p\Big] \ \leq\ \bar{C}_{T,p}\big(1 + |x|^p\big), \qquad \forall\,(x,a)\in\mathbb{R}^d\times A. \qquad \text{(A.3)}$$

Estimate (A.3) implies that the local martingale

$$(M_T)_{T\geq 0} \ \ :=\ \ \left(\int_0^T e^{\gamma s}(X_s^{x,a})^\intercal\sigma(X_s^{x,a},I_s^a)dW_s\right)_{T\geq 0}$$

is indeed a $\mathbb{P}^\nu$-martingale. Then, we have $\mathbb{E}^\nu[e^{-\gamma t}M_t] = 0$. Therefore, taking the expectation $\mathbb{E}^\nu$ in (A.2), we find

$$\mathbb{E}^\nu\big[|X_t^{x,a}|^2\big] \ \leq\ |x|^2 + \left(\frac{M_1 + M_1L_1}{2\varepsilon} + M_1^2\right)\frac{1}{\gamma},$$

from which we deduce (2.4) with $C := \sqrt{\max\{1, [(M_1 + M_1L_1)/(2\varepsilon) + M_1^2]/\gamma\}}$.

• Proof of (ii)

Applying Itô's formula to $|X_t^{x,a} - X_t^{x',a}|^2$ we find

$$\begin{aligned} |X_t^{x,a} - X_t^{x',a}|^2 \ &=\ |x - x'|^2 + 2\int_0^t (X_s^{x,a} - X_s^{x',a}).(b(X_s^{x,a},I_s^a) - b(X_s^{x',a},I_s^a))ds \\ &\quad + \int_0^t \|\sigma(X_s^{x,a},I_s^a) - \sigma(X_s^{x',a},I_s^a)\|^2 ds \qquad \text{(A.4)} \\ &\quad + 2\int_0^t (X_s^{x,a} - X_s^{x',a})^\intercal(\sigma(X_s^{x,a},I_s^a) - \sigma(X_s^{x',a},I_s^a))dW_s. \end{aligned}$$

Now, consider $\nu \in \mathcal{V}$ and recall that W remains a Brownian motion under $\mathbb{P}^\nu$. Using estimate (A.3), we see that the local martingale

$$\left(\int_0^t (X_s^{x,a} - X_s^{x',a})^\intercal (\sigma(X_s^{x,a}, I_s^a) - \sigma(X_s^{x',a}, I_s^a)) dW_s \right)_{t\geq 0}$$

is indeed a $\mathbb{P}^\nu$-martingale. Therefore, taking the expectation $\mathbb{E}^\nu$ with respect to $\mathbb{P}^\nu$ in (A.4) and using the dissipativity condition (2.2), we obtain

$$\mathbb{E}^\nu\big[|X_t^{x,a} - X_t^{x',a}|^2\big] \;\leq\; |x-x'|^2 - 2\gamma \int_0^t \mathbb{E}^\nu\big[|X_s^{x,a} - X_s^{x',a}|^2\big] ds,$$

which implies

$$\mathbb{E}^\nu\big[|X_t^{x,a} - X_t^{x',a}|^2\big] \;\leq\; |x-x'|^2 e^{-2\gamma t}.$$

□

A.2 Proof of Proposition 2.1

Step 1. Existence and uniqueness of ρ. Let $\tilde{W} = (\tilde{W}_t)_{t\geq 0}$ be a d-dimensional Brownian motion, independent of W and μ. Then, we define

$$\bar{W}_t \;=\; \begin{cases} W_t, & t \geq 0, \\ \tilde{W}_{-t}, & t < 0. \end{cases}$$

For any $T \in \mathbb{R}$ and $x \in \mathbb{R}^d$, we denote $X^{T,x} = (X_t^{T,x})_{t\geq T}$ the unique solution to the equation on $[T,\infty)$:

$$dX_t \;=\; \underline{b}(X_t)\, dt + \underline{\sigma}(X_t)\, d\bar{W}_t, \quad t \geq T, \qquad X_T \;=\; x. \tag{A.5}$$

From the time-homogeneity of equation (A.5), it follows the law invariance property $\mathcal{L}(X_t^{T,x})$ $= \mathcal{L}(X_{t-T}^x)$, for $t \geq T$, where X^x is the solution to (A.5) starting from x at time 0.
Let $S > T > 0$ and $x \in \mathbb{R}^d$, then, applying Itô's formula to the difference $|X_t^{-S,x} - X_t^{-T,x}|^2$ from $-T$ to $t \in [-T, 0]$, we obtain

$$\begin{aligned} |X_t^{-S,x} - X_t^{-T,x}|^2 \;&=\; |X_{-T}^{-S,x} - x|^2 + 2\int_{-T}^t (X_s^{-S,x} - X_s^{-T,x}).(\underline{b}(X_s^{-S,x}) - \underline{b}(X_s^{-T,x}))ds \\ &\quad + \int_{-T}^t \|\underline{\sigma}(X_s^{-S,x}) - \underline{\sigma}(X_s^{-T,x})\|^2 ds \\ &\quad + 2\int_{-T}^t (X_s^{-S,x} - X_s^{-T,x})^\intercal (\underline{\sigma}(X_s^{-S,x}) - \underline{\sigma}(X_s^{-T,x}))d\bar{W}_s. \end{aligned} \tag{A.6}$$

Taking the expectation and using the dissipativity condition (2.2), we find

$$\mathbb{E}\big[|X_t^{-S,x} - X_t^{-T,x}|^2\big] \;\leq\; \mathbb{E}\big[|X_{-T}^{-S,x} - x|^2\big] - 2\gamma\int_{-T}^t \mathbb{E}\big[|X_s^{-S,x} - X_s^{-T,x}|^2\big] ds,$$

which implies

$$\mathbb{E}\big[|X_0^{-S,x} - X_0^{-T,x}|^2\big] \;\leq\; \mathbb{E}\big[|X_{-T}^{-S,x} - x|^2\big] e^{-2\gamma T}. \tag{A.7}$$

Similar to (2.4), we can prove that there exists a positive constant $\bar{C}$, depending only on the L_1, M_1, and γ, such that

$$\mathbb{E}\big[|X_{-T}^{-S,x} - x|^2\big] \;\leq\; \bar{C}(1+|x|^2). \tag{A.8}$$

Plugging (A.8) into (A.7), we obtain

$$\mathbb{E}\big[|X_0^{-S,x} - X_0^{-T,x}|^2\big] \;\leq\; \bar{C}(1+|x|^2)e^{-2\gamma T}. \tag{A.9}$$

It follows from (A.9) that $(X_0^{-T,x})_{T>0}$ converges, as $T \to \infty$, to some square integrable random variable η^x, which a priori depends on x. Let $x' \in \mathbb{R}^d$, then applying Itô's formula to $|X_s^{-T,x} - X_s^{-T,x'}|^2$ between $-T$ and $t \in [-T,0]$, we find

$$\begin{aligned} |X_t^{-T,x} - X_t^{-T,x'}|^2 \;=\; & |x-x'|^2 + 2\int_{-T}^t (X_s^{-T,x} - X_s^{-T,x'}).(\underline{b}(X_s^{-T,x}) - \underline{b}(X_s^{-T,x'}))ds \\ & + \int_{-T}^t \|\underline{\sigma}(X_s^{-T,x}) - \underline{\sigma}(X_s^{-T,x'})\|^2 ds \\ & + 2\int_{-T}^t (X_s^{-T,x} - X_s^{-T,x'})^\intercal(\underline{\sigma}(X_s^{-T,x}) - \underline{\sigma}(X_s^{-T,x'}))d\bar{W}_s. \end{aligned}$$

Taking the expectation and using the dissipativity condition (2.2), we obtain

$$\mathbb{E}\big[|X_t^{-T,x} - X_t^{-T,x'}|^2\big] \;\leq\; |x-x'|^2 - 2\gamma\int_{-T}^t \mathbb{E}\big[|X_s^{-T,x} - X_s^{-T,x'}|^2\big]ds,$$

which implies

$$\mathbb{E}\big[|X_0^{-T,x} - X_0^{-T,x'}|^2\big] \;\leq\; |x-x'|^2 e^{-2\gamma T} \overset{T\to\infty}{\longrightarrow} 0.$$

As a consequence, $\eta^x = \eta^{x'} =: \eta$. We denote $\rho := \mathcal{L}(\eta)$. Finally, using the law invariance property already recalled, and the fact that convergence in $\mathbf{L^2}(\mathbb{P})$ implies convergence in law, we deduce

$$\mathcal{L}(X_T^x) \;=\; \mathcal{L}(X_0^{-T,x}) \;\longrightarrow\; \rho, \tag{A.10}$$

weakly as $T \to \infty$. From the square integrability of η we see that $\int_{\mathbb{R}^d}|x|^2\rho(dx) < \infty$. Let us now prove the invariance property. Let $\varphi \in C_b(\mathbb{R}^d)$, then, from the Markov property we have

$$P_{t+s}^{\underline{\alpha}}\varphi(x) \;=\; P_t^{\underline{\alpha}}(P_s^{\underline{\alpha}}\varphi)(x), \qquad \forall\, t,s \geq 0.$$

Sending $t \to \infty$, using (A.10) and the Feller property, we obtain

$$\int_{\mathbb{R}^d}\varphi(x)\rho(dx) \;=\; \int_{\mathbb{R}^d} P_s^{\underline{\alpha}}\varphi(x)\rho(dx), \qquad \forall\, s \geq 0, \tag{A.11}$$

for all $\varphi \in C_b(\mathbb{R}^d)$. By a monotone class argument, we see that (A.11) remains true for all $\varphi \in B(\mathbb{R}^d)$, which implies the invariant property of ρ. Concerning the uniqueness of ρ, let us consider another invariance probability measure ν and take $\varphi \in C_b(\mathbb{R}^d)$, then

$$\int_{\mathbb{R}^d}\varphi(x)\nu(dx) \;=\; \int_{\mathbb{R}^d}P_s^{\underline{\alpha}}\varphi(x)\nu(dx) \overset{s\to\infty}{\longrightarrow} \int_{\mathbb{R}^d}\Big(\int_{\mathbb{R}^d}\varphi(x)\rho(dx)\Big)\nu(dx) \;=\; \int_{\mathbb{R}^d}\varphi(x)\rho(dx).$$

Since the result holds for any $\varphi \in C_b(\mathbb{R}^d)$, we deduce the uniqueness of ρ.
Step 2. (2.9) *is valid for any continuous* φ *satisfying a linear growth condition.* For any $R > 0$, consider a continuous function $\chi_R \colon \mathbb{R}^d \to [0,1]$ which is equal to 1 on $B_R \subset \mathbb{R}^d$ (the open ball of radius R centered at the origin) and is equal to 0 on $\mathbb{R}^d \backslash B_{2R}$. Then

$$\begin{aligned}\left|\mathbb{E}[\varphi(X_t^x)] - \int_{\mathbb{R}^d} \varphi(x)\rho(dx)\right| \ &\leq \ \left|\mathbb{E}[\varphi(X_t^x)\chi_R(X_t^x)] - \int_{\mathbb{R}^d} \varphi(x)\chi_R(x)\rho(dx)\right| \\ &\quad + \left|\mathbb{E}[\varphi(X_t^x)(1-\chi_R(X_t^x))] - \int_{\mathbb{R}^d} \varphi(x)(1-\chi_R(x))\rho(dx)\right|.\end{aligned} \tag{A.12}$$

Since $\varphi\chi_R \in C_b(\mathbb{R}^d)$, the first term on the right-hand side of (A.12) goes to zero as $t \to \infty$. If

$$\lim_{R\to\infty} \limsup_{t\to\infty} \left|\mathbb{E}[\varphi(X_t^x)(1-\chi_R(X_t^x))] - \int_{\mathbb{R}^d} \varphi(x)(1-\chi_R(x))\rho(dx)\right| \ = \ 0, \tag{A.13}$$

then, taking first $\limsup_{t\to\infty}$ and then $\lim_{R\to\infty}$ in (A.12), we get the thesis. Therefore it remains to prove (A.13). From Cauchy-Schwarz inequality, the linear growth property of φ, and estimate (A.8), we have that there exists a positive constant C such that

$$\begin{aligned}\left|\mathbb{E}[\varphi(X_t^x)(1-\chi_R(X_t^x))]\right| \ &\leq \ \sqrt{\mathbb{E}[|\varphi(X_t^x)|^2]}\sqrt{\mathbb{E}_x[|1-\chi_R(X_t^x)|^2]} \\ &\leq \ C(1+|x|)\sqrt{\mathbb{E}[|1-\chi_R(X_t^x)|^2]}.\end{aligned}$$

Since the function $|1-\chi_R|^2 \in C_b(\mathbb{R}^d)$, we find

$$\limsup_{t\to\infty} \left|\mathbb{E}[\varphi(X_t^x)(1-\chi_R(X_t^x))]\right| \ \leq \ C(1+|x|)\sqrt{\int_{\mathbb{R}^d} |1-\chi_R(x)|^2\rho(dx)}.$$

Similarly, we have

$$\begin{aligned}\left|\int_{\mathbb{R}^d} \varphi(x)(1-\chi_R(x))\rho(dx)\right| \ &\leq \ \sqrt{\int_{\mathbb{R}^d} |\varphi(x)|^2\rho(dx)}\sqrt{\int_{\mathbb{R}^d} |1-\chi_R(x)|^2\rho(dx)} \\ &\leq \ C\left(1+\sqrt{\int_{\mathbb{R}^d} |x|^2\rho(dx)}\right)\sqrt{\int_{\mathbb{R}^d} |1-\chi_R(x)|^2\rho(dx)},\end{aligned}$$

where we recall that $\int_{\mathbb{R}^d} |x|^2\rho(dx) < \infty$. In conclusion, we obtain

$$\begin{aligned}&\limsup_{t\to\infty} \left|\mathbb{E}[\varphi(X_t^x)(1-\chi_R(X_t^x))] - \int_{\mathbb{R}^d} \varphi(x)(1-\chi_R(x))\rho(dx)\right| \\ &\leq \ C\left(1+|x|+\sqrt{\int_{\mathbb{R}^d} |x|^2\rho(dx)}\right)\sqrt{\int_{\mathbb{R}^d} |1-\chi_R(x)|^2\rho(dx)}.\end{aligned}$$

Notice that

$$\int_{\mathbb{R}^d} |1-\chi_R(x)|^2\rho(dx) \ \leq \ \int_{\mathbb{R}^d\backslash B_R} \rho(dx) \ = \ \rho(\mathbb{R}^d\backslash B_R) \ \stackrel{R\to\infty}{\longrightarrow} \ 0,$$

which implies (A.13). □

B Elliptic BSDEs

B.1 Proof of Lemma 3.1

Let $0 \leq t \leq T < \infty$ and apply Itô's formula to $e^{-2\beta s}|\Delta Y_s|^2$ between t and T, then

$$
\begin{aligned}
e^{-2\beta t}|\Delta Y_t|^2 \;=\;& e^{-2\beta T}|\Delta Y_T|^2 + 2n\int_t^T\int_A e^{-2\beta s}\Delta Y_s\big[(U_s^{1,\beta,n}(a''))_+ - (U_s^{2,\beta,n}(a''))_+\big]\vartheta(da'')ds \\
&+ 2\int_t^T e^{-2\beta s}\Delta Y_s\big(f_1(X_s^{x,a},I_s^a,\beta Y_s^{1,\beta,n}) - f_2(X_s^{x',a'},I_s^{a'},\beta Y_s^{2,\beta,n})\big)ds \\
&- 2\int_t^T e^{-2\beta s}\Delta Y_s\Delta Z_s\,dW_s - 2\int_t^T\int_A e^{-2\beta s}\Delta Y_s\Delta U_s(a'')\,\tilde\mu(ds,da'') \\
&- \int_t^T e^{-2\beta s}|\Delta Z_s|^2 ds - \int_t^T\int_A e^{-2\beta s}|\Delta U_s(a'')|^2\mu(ds,da''). \qquad \text{(B.14)}
\end{aligned}
$$

Notice that, using the nonincreasing property of f_1 in y, we have

$$
\begin{aligned}
&\int_t^T e^{-2\beta s}\Delta Y_s\big(f_1(X_s^{x,a},I_s^a,\beta Y_s^{1,\beta,n}) - f_2(X_s^{x',a'},I_s^{a'},\beta Y_s^{2,\beta,n})\big)ds \\
&\qquad\qquad \leq \int_t^T e^{-2\beta s}\Delta Y_s(\Delta_s' f_1 + \Delta_s f)ds.
\end{aligned}
$$

Now, define the $[1,n+1]$-valued map ν as follows

$$
\nu_t(a'') \;=\; 1 + n\frac{(U_t^{1,\beta,n}(a''))_+ - (U_t^{2,\beta,n}(a''))_+}{\Delta U_t(a'')}1_{\{\Delta U_t(a'')\neq 0\}}, \qquad t \geq 0,\ a'' \in A.
$$

Observe that ν is a $\mathcal{P}\otimes\mathcal{B}(A)$-measurable map satisfying $1 \leq \nu_s(a) \leq n+1$, $ds\otimes d\mathbb{P}\otimes\vartheta(da)$-a.e., then $\nu \in \mathcal{V}_n$. Let us consider the probability measure $\mathbb{P}^\nu$ equivalent to $\mathbb{P}$ on $(\Omega,\mathcal{F}_T)$ with Radon-Nikodym density given by (2.3). Recalling that $\tilde\mu^\nu$ denotes the compensated martingale measure associated to μ under $\mathbb{P}^\nu$, equation (B.14) can be rewritten as follows

$$
\begin{aligned}
& e^{-2\beta t}|\Delta Y_t|^2 + \int_t^T e^{-2\beta s}|\Delta Z_s|^2 ds + \int_t^T\int_A e^{-2\beta s}|\Delta U_s(a'')|^2\mu(ds,da'') \\
&\leq\; e^{-2\beta T}|\Delta Y_T|^2 + 2\int_t^T e^{-2\beta s}\Delta Y_s(\Delta_s' f_1 + \Delta_s f)ds - 2\int_t^T e^{-2\beta s}\Delta Y_s\Delta Z_s\,dW_s \\
&\quad - 2\int_t^T\int_A e^{-2\beta s}\Delta Y_s\Delta U_s(a'')\,\tilde\mu^\nu(ds,da''). \qquad \text{(B.15)}
\end{aligned}
$$

From Lemma 2.5 in [19], we see that the two stochastic integrals on the right-hand side of (B.15) are martingales. Hence, taking the expectation $\mathbb{E}^\nu$, conditional on $\mathcal{F}_t$, with respect to $\mathbb{P}^\nu$ in (B.15), we end up with estimate (3.7). $\square$

B.2 Proof of Proposition 3.1

Uniqueness. Fix $(\beta,n) \in (0,\infty)\times\mathbb{N}$ and consider two solutions $(Y^{1,\beta,n},Z^{1,\beta,n},U^{1,\beta,n})$, $(Y^{2,\beta,n},Z^{2,\beta,n},U^{2,\beta,n}) \in \mathbf{S}^2_{\text{loc}}\times\mathbf{L}^2_{\text{loc}}(\mathbf{W})\times\mathbf{L}^2_{\text{loc}}(\tilde\mu)$ to (3.6). Set $\Delta Y_t = Y_t^{1,\beta,n} - Y_t^{2,\beta,n}$, $\Delta Z_t = Z_t^{1,\beta,n} - Z_t^{2,\beta,n}$, and $\Delta U_t(a') = U_t^{1,\beta,n}(a') - U_t^{2,\beta,n}(a')$, $t \geq 0$, $a' \in A$. Let $0 \leq t \leq T < \infty$.

Then, from estimate (3.7) with $f_1 = f_2 = f$ (so that $\Delta' f_1 = \Delta f = 0$), there exists $\nu \in \mathcal{V}_n$ such that

$$e^{-2\beta t}|\Delta Y_t|^2 \ \leq \ \mathbb{E}^\nu\big[e^{-2\beta T}|\Delta Y_T|^2\big|\mathcal{F}_t\big]. \tag{B.16}$$

Moreover, recall from (2.4) that the following estimate holds

$$\mathbb{E}^\nu\big[|X_T^{x,a}|^2\big] \ \leq \ C_{b,\sigma}(1+|x|^2), \qquad \forall\, T \geq 0.$$

Since $|\Delta Y_T| \leq 2C(1+|X_T^{x,a}|)$, we conclude that $\mathbb{E}^\nu[e^{-2\beta T}|\Delta Y_T|^2] \to 0$, as $T \to \infty$. From (B.16) it follows that $\Delta Y = 0$. Finally, plugging $\Delta Y = 0$ into (B.15), we conclude that $\Delta Z = 0$ and $\Delta U = 0$.

Existence. Step 1. Approximating BSDE. Fix $(x,a,\beta,n) \in \mathbb{R}^d \times \mathbb{R}^q \times (0,\infty) \times \mathbb{N}$, $T > 0$, and consider the backward stochastic differential equation on $[0,T]$ given by, $\mathbb{P}$-a.s.,

$$\begin{aligned} Y_t \ &= \ -\beta\int_t^T Y_s\,ds + n\int_t^T\int_A (U_s(a'))_+\vartheta(da')\,ds + \int_t^T f(X_s^{x,a}, I_s^a, \beta Y_s)\,ds \\ &\quad - \int_t^T Z_s\,dW_s - \int_t^T\int_A U_s(a')\,\tilde\mu(ds,da'), \qquad 0 \leq t \leq T. \end{aligned} \tag{B.17}$$

Notice that (B.17) has a zero terminal condition at the final time T. It follows from Lemma 2.4 in [29] that there exists a unique solution $(Y^T, Z^T, U^T) \in \mathbf{S^2(0,T)} \times \mathbf{L^2(W;0,T)} \times \mathbf{L^2}(\tilde\mu;\mathbf{0,T})$ to (B.17).

Step 2. Estimate for Y^T. Let $t \in [0,T]$, then, from estimate (3.7) with $(Y^{1,\beta,n}, Z^{1,\beta,n}, U^{1,\beta,n}) = (Y^T, Z^T, U^T)$, $(Y^{2,\beta,n}, Z^{2,\beta,n}, U^{2,\beta,n}) = (0,0,0)$, $f_1 = f$, and $f_2 = 0$, there exists $\nu \in \mathcal{V}_n$:

$$\begin{aligned} |Y_t^T|^2 \ &\leq \ 2\int_t^T e^{-2\beta(s-t)}\mathbb{E}^\nu\big[Y_s^T f(X_s^{x,a}, I_s^a, 0)\big|\mathcal{F}_t\big]\,ds \\ &\leq \ 2\int_t^T e^{-2\beta(s-t)}\sqrt{\mathbb{E}^\nu\big[|Y_s^T|^2\big|\mathcal{F}_t\big]}\sqrt{\mathbb{E}^\nu\big[|f(X_s^{x,a}, I_s^a, 0)|^2\big|\mathcal{F}_t\big]}\,ds. \end{aligned} \tag{B.18}$$

Set $g(s) = e^{-2\beta(s-t)}\mathbb{E}^\nu[|Y_s^T|^2|\mathcal{F}_t]$ and $h(s) = 2e^{-\beta(s-t)}\sqrt{\mathbb{E}^\nu[|f(X_s^{x,a}, I_s^a, 0)|^2]}$, for any $s \in [t,T]$. Then, recalling that $g(T) = 0$, inequality (B.18) becomes

$$g(t) \ \leq \ g(T) + \int_t^T \sqrt{g(s)}h(s)ds.$$

Our aim is to derive a Gronwall type estimate for g. To this end, define

$$\tilde g(t) \ := \ g(T) + \int_t^T \sqrt{g(s)}h(s)ds, \qquad 0 \leq t \leq T.$$

Notice that $\tilde g \in C^1([0,T])$. Moreover $g(t) \leq \tilde g(t)$, for any $0 \leq t \leq T$, and

$$\tilde g(T') \ = \ \tilde g(T'') + \int_{T'}^{T''} \sqrt{g(s)}h(s)ds \ \leq \ \tilde g(T'') + \int_{T'}^{T''} \sqrt{\tilde g(s)}h(s)ds, \quad t \leq T' < T'' \leq T.$$

Dividing by $T'' - T'$ and letting $T'' - T' \to 0$, we deduce the differential inequality

$$\tilde g'(s) \ \geq \ -h(s)\sqrt{\tilde g(s)}, \qquad t \leq s \leq T.$$

We have

$$\frac{d\sqrt{\tilde{g}(s)}}{ds} \;\geq\; -\frac{1}{2}h(s),$$

which yields

$$\sqrt{\tilde{g}(T)} - \sqrt{\tilde{g}(t)} \;\geq\; -\frac{1}{2}\int_t^T h(s)ds.$$

Therefore, we find

$$|Y_t^T| \;=\; \sqrt{g(t)} \;\leq\; \sqrt{\tilde{g}(t)} \;\leq\; \int_t^T e^{-\beta(s-t)}\sqrt{\mathbb{E}^\nu\big[|f(X_s^{x,a}, I_s^a, 0)|^2\big|\mathcal{F}_t\big]}ds.$$

Recalling that $|f(x,a,0)| \leq L_2|x|+M_2$, with $M_2 := \sup_{a\in A}|f(0,a,0)|$, so that $|f(x,a,0)|^2 \leq 2L_2^2|x|^2 + 2M_2^2$, and using the inequality $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, for any $a, b \in \mathbb{R}_+$, we find

$$|Y_t^T| \;\leq\; \sqrt{2}L_2\int_t^\infty e^{-\beta(s-t)}\sqrt{\mathbb{E}^\nu\big[|X_s^{x,a}|^2\big|\mathcal{F}_t\big]}ds + \sqrt{2}M_2\int_t^\infty e^{-\beta(s-t)}ds.$$

From estimate (2.6), we have

$$\begin{aligned} |Y_t^T| \;&\leq\; \sqrt{2}\big(L_2\sqrt{C_{b,\sigma}}\big(1+|X_t^{x,a}|\big) + M_2\big)\int_t^\infty e^{-\beta(s-t)}ds \\ &=\; \sqrt{2}\frac{L_2\sqrt{C_{b,\sigma}}(1+|X_t^{x,a}|) + M_2}{\beta}. \end{aligned} \tag{B.19}$$

Step 3. Convergence of $(Y^T)_{T>0}$. Let $T, T' > 0$, with $T < T'$, and denote $\Delta Y_t = Y_t^T - Y_t^{T'}$, $0 \leq t \leq T$. Let $t \in [0,T]$, then estimate (3.7) reads

$$|\Delta Y_t|^2 \;\leq\; e^{-2\beta(T-t)}\mathbb{E}^\nu\big[|\Delta Y_T|^2\big|\mathcal{F}_t\big] \;\overset{T\to\infty}{\longrightarrow}\; 0, \tag{B.20}$$

where the convergence result follows from (B.19) and (2.6). Let us now consider the family of real-valued càdlàg adapted processes $(Y^T)_{T>0}$. It follows from (B.20) that, for any $t \geq 0$, the family $(Y_t^T(\omega))_{T>0}$ is Cauchy for almost every ω, so that it converges $\mathbb{P}$-a.s. to some $\mathcal{F}_t$-measurable random variable Y_t, which is bounded from the right-hand side of (B.19). Moreover, using again (B.20), (B.19), and (2.6), we see that, for any $0 \leq S < T \wedge T'$, with $T, T' > 0$, we have

$$\sup_{0\leq t\leq S}|Y_t^{T'} - Y_t^T| \;\leq\; e^{-\beta(T\wedge T'-S)}C_0\sup_{0\leq t\leq S}\big(1+|X_t^{x,a}|\big) \;\overset{T,T'\to\infty}{\longrightarrow}\; 0, \tag{B.21}$$

where C_0 is a positive constant independent of S, T, T'. In other words, the family $(Y^T)_{T>0}$ converges $\mathbb{P}$-a.s. to Y uniformly on compact subsets of $\mathbb{R}_+$. Since each Y^T is a càdlàg process, it follows that Y is càdlàg, as well. Finally, from estimate (B.19) we see that $Y \in \mathbf{S}^2_{\mathrm{loc}}$ and

$$|Y_t| \;\leq\; \frac{C}{\beta}\big(1+|X_t^{x,a}|\big), \qquad \forall\, t \geq 0.$$

Step 4. Convergence of $(Z^T, U^T)_{T>0}$. Let $S, T, T' > 0$, with $S < T < T'$. Then, applying Itô's formula to $e^{-2\beta t}|Y_t^{T'} - Y_t^T|^2$ between 0 and S, and taking the expectation, we find

$$\mathbb{E}\int_0^S e^{-2\beta s}|Z_s^{T'} - Z_s^T|^2ds + \mathbb{E}\int_0^S\int_A e^{-2\beta s}|U_s^{T'}(a') - U_s^T(a')|^2\vartheta(da')ds$$

$$
\begin{aligned}
&= e^{-2\beta S}\mathbb{E}\big[|Y_S^{T'} - Y_S^T|^2\big] - |Y_0^{T'} - Y_0^T|^2 \\
&\quad + 2\mathbb{E}\int_0^S e^{-2\beta s}\big(Y_s^{T'} - Y_s^T\big)\big(f(X_s^{x,a}, I_s^a, \beta Y_s^{T'}) - f(X_s^{x,a}, I_s^a, \beta Y_s^T)\big)ds \\
&\quad + 2n\mathbb{E}\int_0^S\int_A e^{-2\beta s}\big(Y_s^{T'} - Y_s^T\big)\big((U_s^{T'}(a'))_+ - (U_s^T(a'))_+\big)\vartheta(da')ds.
\end{aligned}
$$

Since the map $y \mapsto f(x,a,y)$ is nonincreasing, we get (using also the inequality $ab \leq a^2/2 + b^2/2$, for any $a,b\in\mathbb{R}$)

$$
\begin{aligned}
&\mathbb{E}\int_0^S e^{-2\beta s}|Z_s^{T'} - Z_s^T|^2 ds + \mathbb{E}\int_0^S\int_A e^{-2\beta s}|U_s^{T'}(a') - U_s^T(a')|^2\vartheta(da')ds \\
&\leq e^{-2\beta S}\mathbb{E}\big[|Y_S^{T'} - Y_S^T|^2\big] + 2n\mathbb{E}\int_0^S\int_A e^{-2\beta s}|Y_s^{T'} - Y_s^T||(U_s^{T'}(a'))_+ - (U_s^T(a'))_+|\vartheta(da')ds \\
&\leq e^{-2\beta S}\mathbb{E}\big[|Y_S^{T'} - Y_S^T|^2\big] + 2n^2\vartheta(A)\mathbb{E}\int_0^S e^{-2\beta s}|Y_s^{T'} - Y_s^T|^2 ds \\
&\quad + \frac{1}{2}\mathbb{E}\int_0^S\int_A e^{-2\beta s}|(U_s^{T'}(a'))_+ - (U_s^T(a'))_+|^2\vartheta(da')ds.
\end{aligned}
$$

Multiplying the previous inequality by $e^{2\beta S}$, we obtain

$$
\begin{aligned}
&\mathbb{E}\int_0^S |Z_s^{T'} - Z_s^T|^2 ds + \frac{1}{2}\mathbb{E}\int_0^S\int_A |U_s^{T'}(a') - U_s^T(a')|^2\vartheta(da')ds \\
&\leq \mathbb{E}\int_0^S e^{2\beta(S-s)}|Z_s^{T'} - Z_s^T|^2 ds + \frac{1}{2}\mathbb{E}\int_0^S\int_A e^{2\beta(S-s)}|U_s^{T'}(a') - U_s^T(a')|^2\vartheta(da')ds \\
&\leq \mathbb{E}\big[|Y_S^{T'} - Y_S^T|^2\big] + 2n^2\vartheta(A)\mathbb{E}\int_0^S e^{2\beta(S-s)}|Y_s^{T'} - Y_s^T|^2 ds \overset{T,T'\to\infty}{\longrightarrow} 0,
\end{aligned}
$$

where the convergence to zero follows from estimate (B.21). Then, for any $S > 0$, we see that the family $(Z^T_{|[0,S]}, U^T_{|[0,S]})_{T>S}$ is Cauchy in the Hilbert space $\mathbf{L^2(W;0,S)}\times\mathbf{L^2}(\tilde{\mu};\mathbf{0,S})$. Therefore, we deduce that there exists $(\bar{Z}^S,\bar{U}^S) \in \mathbf{L^2(W;0,S)}\times\mathbf{L^2}(\tilde{\mu};\mathbf{0,S})$ such that $(Z^T_{|[0,S]}, U^T_{|[0,S]})_{T>S}$ converges to $(\bar{Z}^S,\bar{U}^S)$ in $\mathbf{L^2(W;0,S)}\times\mathbf{L^2}(\tilde{\mu};\mathbf{0,S})$, i.e.,

$$
\mathbb{E}\int_0^S |Z_s^T - \bar{Z}_s^S|^2 ds + \mathbb{E}\int_0^S\int_A |U_s^T(a') - \bar{U}_s^S(a')|^2\vartheta(da')ds \overset{T\to\infty}{\longrightarrow} 0.
$$

Notice that $\bar{Z}^{S'}_{|[0,S]} = \bar{Z}^S$ and $\bar{U}^{S'}_{|[0,S]} = \bar{U}^S$, for any $0\leq S\leq S' < \infty$. Indeed, $(\bar{Z}^{S'}_{|[0,S]}, \bar{U}^{S'}_{|[0,S]})$, as $(\bar{Z}^S,\bar{U}^S)$, is the limit in $\mathbf{L^2(W;0,S)}\times\mathbf{L^2}(\tilde{\mu};\mathbf{0,S})$ of $(Z^T_{|[0,S]}, U^T_{|[0,S]})_{T>S}$. Hence, we define $Z_s = \bar{Z}_s^S$ and $U_s = \bar{U}_s^S$, for all $s\in[0,S]$ and for any $S>0$. Observe that $(Z,U) \in \mathbf{L^2_{\mathrm{loc}}(W)}\times\mathbf{L^2_{\mathrm{loc}}}(\tilde{\mu})$. Moreover, for any $S>0$, $(Z^T_{|[0,S]}, U^T_{|[0,S]})_{T>S}$ converges to $(Z_{|[0,S]}, U_{|[0,S]})$ in $\mathbf{L^2(W;0,S)}\times\mathbf{L^2}(\tilde{\mu};\mathbf{0,S})$, i.e.,

$$
\mathbb{E}\int_0^S |Z_s^T - Z_s|^2 ds + \mathbb{E}\int_0^S\int_A |U_s^T(a') - U_s(a')|^2\vartheta(da')ds \overset{T\to\infty}{\longrightarrow} 0. \tag{B.22}
$$

Now, fix $S\in[0,T]$ and consider the BSDE satisfied by (Y^T,Z^T,U^T) on $[0,S]$:

$$
Y_t^T = Y_S^T - \beta\int_t^S Y_s^T\, ds + n\int_t^S\int_A (U_s^T(a'))_+\vartheta(da')\, ds + \int_t^S f(X_s^{x,a}, I_s^a, \beta Y_s^T)\, ds
$$

$$- \int_t^S Z_s^T \, dW_s - \int_t^S \int_A U_s^T(a') \, \tilde{\mu}(ds, da'), \qquad 0 \leq t \leq S.$$

From (B.21) and (B.22), we can pass to the limit in the above BSDE by letting $T \to \infty$, keeping S fixed. Then, we deduce that (Y, Z, U) solves the penalized BSDE (3.6) on $[0, S]$. Since S is arbitrary, it follows that (Y, Z, U) solves equation (3.6) on $[0, \infty)$. $\Box$

B.3 Proof of Lemma 3.2

The linear growth of $v^{\beta,n}$ follows from (3.9) and the estimate on $Y^{x,a,\beta,n}$ of Proposition 3.1. Concerning the identification $Y_t^{x,a,\beta,n} = v^{\beta,n}(X_t^{x,a}, I_t^a)$, it is a consequence, as usual, of the flow property $(X_T^{x,a}, I_T^a) = (X_{T-t}^{X_t^{x,a}, I_t^a}, I_{T-t}^{I_t^a})$ $\mathbb{P}$-a.s., for any $0 \leq t \leq T < \infty$, and from the uniqueness for the penalized BSDE. Finally, regarding the uniform Lipschitz condition (3.11) of $v^{\beta,n}$ with respect to x, consider $x, x' \in \mathbb{R}^d$ and set $\Delta Y_t = Y_t^{x,a,\beta,n} - Y_t^{x',a,\beta,n}$, $\Delta Z_t = Z_t^{x,a,\beta,n} - Z_t^{x',a,\beta,n}$, $\Delta U_t(a') = U_t^{x,a,\beta,n}(a') - U_t^{x',a,\beta,n}(a')$, $t \geq 0$, $a' \in A$. Let $T \in (0, \infty)$, then from estimate (3.7) there exists $\nu \in \mathcal{V}_n$:

$$\begin{aligned}
|\bar{Y}_0|^2 \ &\leq \ \mathbb{E}^\nu\big[e^{-2\beta T}|\bar{Y}_T|^2\big] \\
&\quad + 2\int_0^T e^{-2\beta s}\mathbb{E}^\nu\big[\bar{Y}_s\big(f(X_s^{x,a}, I_s^a, \beta Y_s^{x,a,\beta,n}) - f(X_s^{x',a}, I_s^a, \beta Y_s^{x,a,\beta,n})\big)\big] ds \\
&\leq \ \mathbb{E}^\nu\big[e^{-2\beta T}|\bar{Y}_T|^2\big] \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(B.23)} \\
&\quad + 2\int_0^T e^{-2\beta s}\sqrt{\mathbb{E}^\nu\big[|\bar{Y}_s|^2\big]}\sqrt{\mathbb{E}^\nu\big[\big|f(X_s^{x,a}, I_s^a, \beta Y_s^{x,a,\beta,n}) - f(X_s^{x',a}, I_s^a, \beta Y_s^{x,a,\beta,n})\big|^2\big]} ds.
\end{aligned}$$

Set $g(s) = e^{-2\beta s}\mathbb{E}^\nu[|Y_s|^2]$ and $h(s) = 2e^{-\beta s}\sqrt{\mathbb{E}^\nu[|f(X_s^{x,a}, I_s^a, \beta Y_s^{x,a,\beta,n}) - f(X_s^{x',a}, I_s^a, \beta Y_s^{x,a,\beta,n})|^2]}$, for any $s \in [0, T]$, and proceed as in (B.18). Then, we conclude that

$$\begin{aligned}
|\bar{Y}_0|^2 \ &\leq \ \sqrt{\mathbb{E}^\nu\big[e^{-2\beta T}|\bar{Y}_T|^2\big]} + \int_0^T e^{-\beta s}\sqrt{\mathbb{E}^\nu\big[\big|f(X_s^{x,a}, I_s^a, \beta Y_s^{x,a,\beta,n}) - f(X_s^{x',a}, I_s^a, \beta Y_s^{x,a,\beta,n})\big|^2\big]} ds \\
&\leq \ \sqrt{\mathbb{E}^\nu\big[e^{-2\beta T}|\bar{Y}_T|^2\big]} + L_2 \int_0^T e^{-\beta s}\sqrt{\mathbb{E}^\nu\big[\big|X_s^{x,a} - X_s^{x',a}\big|^2\big]} ds.
\end{aligned}$$

Therefore, recalling that $|\bar{Y}_T| \leq 2C_{b,\sigma,f}(1 + |X_T^{x,a}|)/\beta$ and using estimate (2.4), we obtain

$$\mathbb{E}^\nu\big[e^{-2\beta T}|\bar{Y}_T|^2\big] \ \overset{T\to\infty}{\longrightarrow} \ 0.$$

By Lemma 2.1, we find

$$\mathbb{E}^\nu\big[|X_s^{x,a} - X_s^{x',a}|^2\big] \ \leq \ e^{-2\gamma s}|x - x'|^2, \qquad \forall\, s \geq 0.$$

Then, we deduce that

$$|\bar{Y}_0| \ \leq \ L_2 \int_0^\infty e^{-(\beta+\gamma)s}|x - x'| ds \ = \ \frac{L_2}{\beta + \gamma}|x - x'| \ \leq \ \frac{L_2}{\gamma}|x - x'|,$$

which implies (3.11). $\Box$

B.4 Proof of Proposition 3.2

Continuity. Fix $(\beta, n) \in (0,\infty)\times\mathbb{N}$. Let $x, x' \in \mathbb{R}^d$ and $a, a' \in \mathbb{R}^q$. Set $\Delta Y_t = Y_t^{x,a,\beta,n} - Y_t^{x',a',\beta,n}$, $\Delta Z_t = Z_t^{x,a,\beta,n} - Z_t^{x',a',\beta,n}$, $\Delta U_t = U_t^{x,a,\beta,n} - U_t^{x',a',\beta,n}$. Then, from estimate (3.7) we find, for any $0 \le r \le T$,

$$\begin{aligned}
|\Delta Y_r|^2 \;\le\; & e^{-2\beta T}\mathbb{E}^\nu\big[|\Delta Y_T|^2\big] \\
& + 2\mathbb{E}^\nu\bigg[\int_r^T e^{-2\beta s}\Delta Y_s\big(f(X_s^{x,a}, I_s^a, \beta Y_s^{x,a,\beta,n}) - f(X_s^{x',a'}, I_s^{a'}, \beta Y_s^{x,a,\beta,n})\big)ds\bigg] \\
\le\; & e^{-2\beta T}\mathbb{E}^\nu\big[|\Delta Y_T|^2\big] + \mathbb{E}^\nu\int_r^T e^{-2\beta s}|\Delta Y_s|^2 ds \\
& + \mathbb{E}^\nu\int_r^T e^{-2\beta s}\big|f(X_s^{x,a}, I_s^a, \beta Y_s^{x,a,\beta,n}) - f(X_s^{x',a'}, I_s^{a'}, \beta Y_s^{x,a,\beta,n})\big|^2 ds.
\end{aligned}$$

From Gronwall's lemma applied to the map $s \mapsto \mathbb{E}^\nu[|\Delta Y_s|^2]$ we obtain

$$\begin{aligned}
|\Delta Y_0|^2 \;=\; & e^{\frac{1-e^{-2\beta T}}{2\beta}}\bigg(e^{-2\beta T}\mathbb{E}^\nu\big[|\Delta Y_T|^2\big] \\
& + \mathbb{E}^\nu\int_0^T e^{-2\beta s}\big|f(X_s^{x,a}, I_s^a, \beta Y_s^{x,a,\beta,n}) - f(X_s^{x',a'}, I_s^{a'}, \beta Y_s^{x,a,\beta,n})\big|^2 ds\bigg).
\end{aligned}$$

From the Lipschitz property of f in **(H2)**, we find

$$\begin{aligned}
|\Delta Y_0|^2 \;=\; & e^{\frac{1-e^{-2\beta T}}{2\beta}}\bigg(e^{-2\beta T}\mathbb{E}^\nu\big[|\Delta Y_T|^2\big] \\
& + 2L_2^2\int_0^T e^{-2\beta s}\big\{\mathbb{E}^\nu\big[|X_s^{x,a} - X_s^{x',a'}|^2\big] + \mathbb{E}^\nu\big[|I_s^a - I_s^{a'}|^2\big]\big\}ds\bigg). \qquad \text{(B.24)}
\end{aligned}$$

Now, for any $\varepsilon > 0$, applying Itô's formula to $e^{(2\gamma-\varepsilon-\varepsilon L_1^2)t}|X_t^{x,a} - X_t^{x',a'}|^2$ and proceeding as in the proof of estimate (2.5), we obtain

$$\mathbb{E}^\nu\big[|X_t^{x,a} - X_t^{x',a'}|^2\big] \;\le\; |x-x'|^2 + \Big(1+\frac{2}{\varepsilon}\Big)L_1^2\int_0^t e^{(2\gamma-\varepsilon-\varepsilon L_1^2)(s-t)}\mathbb{E}^\nu\big[|I_s^a - I_s^{a'}|^2\big]ds.$$

Denote by T_1 the first jump time of the marked point process $(T_n, \alpha_n)_{n\ge1}$ associated to the Poisson random measure μ. Notice that the two processes I^a and $I^{a'}$ coincide after T_1, while we have $I_s^a = a$ and $I_s^{a'} = a'$ before T_1. In other words, $|I_s^a - I_s^{a'}| = |a-a'|1_{\{s\le T_1\}} \le |a-a'|$. Therefore

$$\begin{aligned}
\mathbb{E}^\nu\big[|X_t^{x,a} - X_t^{x',a'}|^2\big] \;\le\; & |x-x'|^2 + \Big(1+\frac{2}{\varepsilon}\Big)L_1^2|a-a'|^2\int_0^t e^{(2\gamma-\varepsilon-\varepsilon L_1^2)(s-t)}ds \\
\le\; & |x-x'|^2 + \Big(1+\frac{2}{\varepsilon}\Big)L_1^2|a-a'|^2\int_0^\infty e^{-(2\gamma-\varepsilon-\varepsilon L_1^2)s}ds \\
=\; & |x-x'|^2 + \Big(1+\frac{2}{\varepsilon}\Big)L_1^2|a-a'|^2\frac{1}{2\gamma-\varepsilon-\varepsilon L_1^2}.
\end{aligned}$$

Therefore, (B.24) becomes

$$|\Delta Y_0|^2 \;\le\; e^{\frac{1-e^{-2\beta T}}{2\beta}}\bigg(e^{-2\beta T}\mathbb{E}^\nu\big[|\Delta Y_T|^2\big] + C_0\big(|x-x'|^2 + |a-a'|^2\big)\int_0^T e^{-2\beta s}ds\bigg)$$

$$\leq \; e^{\frac{1-e^{-2\beta T}}{2\beta}} \Big(e^{-2\beta T} \mathbb{E}^\nu \big[|\Delta Y_T|^2 \big] + C_0 \big(|x-x'|^2 + |a-a'|^2 \big) \frac{1}{2\beta} \Big), \tag{B.25}$$

for some positive constant C_0, possibly depending on L_1, L_2, ε, but independent of T. Since $|\Delta Y_T| \leq 2C_{b,\sigma,f}(1+|X_T^{x,a}|)/\beta$, using estimate (2.4) we see that $\mathbb{E}[e^{-2\beta T}|\Delta Y_T|^2] \to 0$ as $T \to \infty$. Therefore, letting $T \to \infty$ in (B.25), it follows that $|\Delta Y_0|^2 \to 0$ as $(x',a') \to (x,a)$. Since $\Delta Y_0 = v^{\beta,n}(x,a) - v^{\beta,n}(x',a')$, then $v^{\beta,n}$ is continuous in both arguments.
Viscosity property. We shall now prove the viscosity supersolution property of $v^{\beta,n}$. A similar argument would show that $v^{\beta,n}$ it is a viscosity subsolution to equation (3.12). Let $(\bar{x},\bar{a}) \in \mathbb{R}^d \times \mathbb{R}^q$ and $\varphi \in C^2(\mathbb{R}^d \times \mathbb{R}^q)$ such that

$$0 \;=\; (v^{\beta,n} - \varphi)(\bar{x},\bar{a}) \;=\; \min_{\mathbb{R}^d \times \mathbb{R}^q} (v^{\beta,n} - \varphi). \tag{B.26}$$

Let us proceed by contradiction, assuming that

$$\begin{aligned} &\beta\,\varphi(\bar{x},\bar{a}) - \mathcal{L}^{\bar{a}}\varphi(\bar{x},\bar{a}) - \mathcal{M}^{\bar{a}}\varphi(\bar{x},\bar{a}) - f(\bar{x},\bar{a},\beta v^{\beta,n}(\bar{x},\bar{a})) \\ &\qquad -n\int_A [\varphi(\bar{x},a') - \varphi(\bar{x},\bar{a})]_+\,\vartheta(da') \;=:\; -2\varepsilon \;<\; 0. \end{aligned}$$

Using the continuity of b, σ, f, and $v^{\beta,n}$, we find $\delta > 0$ such that

$$\begin{aligned} &\beta\,\varphi(x,a) - \mathcal{L}^{a}\varphi(x,a) - \mathcal{M}^{a}\varphi(x,a) - f(x,a,\beta v^{\beta,n}(x,a)) \\ &\qquad -n\int_A [\varphi(x,a') - \varphi(x,a)]_+\,\vartheta(da') \;\leq\; -\varepsilon, \end{aligned} \tag{B.27}$$

for any $(x,a) \in \mathbb{R}^d \times \mathbb{R}^q$, with $|x-\bar{x}|, |a-\bar{a}| < \delta$. Define

$$\tau \;:=\; \inf\big\{t \geq 0\colon |X_t^{\bar{x},\bar{a}} - \bar{x}| > \delta,\, |I_t^{\bar{a}} - \bar{a}| > \delta\big\} \wedge \delta$$

Since $(X^{\bar{x},\bar{a}}, I^{\bar{a}})$ is càdlàg, it is in particular right-continuous at time 0. Therefore, $\tau > 0$, $\mathbb{P}$-almost surely. Then, an application of Itô's formula to $e^{-\beta t}\varphi(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}})$ between 0 and τ, using also (B.27), yields

$$\begin{aligned} &e^{-\beta\tau}\varphi(X_\tau^{\bar{x},\bar{a}}, I_\tau^{\bar{a}}) \\ &\geq\; \varphi(\bar{x},\bar{a}) + \varepsilon\frac{1-e^{-\beta\tau}}{\beta} - n\int_0^\tau\int_A e^{-\beta t}\big(\tilde{U}_t^{\bar{x},\bar{a},\beta,n}(a')\big)_+\vartheta(da')dt \\ &\quad - \int_0^\tau e^{-\beta t} f(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}}, \beta v^{\beta,n}(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}}))dt + \int_0^\tau e^{-\beta t}(D_x\varphi(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}}))^\intercal \sigma(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}})dW_t \\ &\quad + \int_0^\tau\int_A e^{-\beta t}\tilde{U}_t^{\bar{x},\bar{a},\beta,n}(a')\tilde{\mu}(dt,da'), \end{aligned} \tag{B.28}$$

where $\tilde{U}_t^{\bar{x},\bar{a},\beta,n}(a') = \varphi(X_t^{\bar{x},\bar{a}}, a') - \varphi(X_t^{\bar{x},\bar{a}}, I_{t-}^{\bar{a}})$. On the other hand, applying Itô's formula to $e^{-\beta t}Y_t^{\bar{x},\bar{a},\beta,n}$ from 0 to τ, and using the identification $Y_t^{\bar{x},\bar{a},\beta,n} = v^{\beta,n}(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}})$, we find

$$\begin{aligned} v^{\beta,n}(\bar{x},\bar{a}) \;&=\; e^{-\beta\tau}v^{\beta,n}(X_\tau^{\bar{x},\bar{a}}, I_\tau^{\bar{a}}) + n\int_0^\tau\int_A e^{-\beta t}\big(U_t^{\bar{x},\bar{a},\beta,n}(a')\big)_+\vartheta(da')dt \\ &\quad + \int_0^\tau e^{-\beta t} f(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}}, \beta v^{\beta,n}(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}}))dr - \int_0^\tau e^{-\beta t} Z_t^{\bar{x},\bar{a},\beta,n}dW_t \end{aligned}$$

$$-\int_0^\tau \int_A e^{-\beta t} U_t^{\bar{x},\bar{a},\beta,n}(a')\tilde{\mu}(dt,da'). \tag{B.29}$$

Plugging identity (B.29) into inequality (B.28), we obtain

$$\begin{aligned}
& e^{-\beta\tau}\varphi(X_\tau^{\bar{x},\bar{a}}, I_\tau^{\bar{a}}) - e^{-\beta\tau} v^{\beta,n}(X_\tau^{\bar{x},\bar{a}}, I_\tau^{\bar{a}}) \\
\geq\ & \varphi(\bar{x},\bar{a}) - v^{\beta,n}(\bar{x},\bar{a}) + \varepsilon\frac{1-e^{-\beta\tau}}{\beta} \\
& - n\int_0^\tau \int_A e^{-\beta t}\big[\big(\tilde{U}_t^{\bar{x},\bar{a},\beta,n}(a')\big)_+ - \big(U_t^{\bar{x},\bar{a},\beta,n}(a')\big)_+\big]\vartheta(da')dt \\
& + \int_0^\tau e^{-\beta t}\big(\sigma^\intercal(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}}) D_x\varphi(X_t^{\bar{x},\bar{a}}, I_t^{\bar{a}}) - Z_t^{\bar{x},\bar{a},\beta,n}\big)dW_t \\
& + \int_0^\tau \int_A e^{-\beta t}\big(\tilde{U}_t^{\bar{x},\bar{a},\beta,n}(a') - U_t^{\bar{x},\bar{a},\beta,n}(a')\big)\tilde{\mu}(dt,da').
\end{aligned} \tag{B.30}$$

Define the $[1,n+1]$-valued $\mathcal{P}\otimes\mathcal{B}(A)$-measurable map ν as follows

$$\nu_t(a') \;=\; 1 + n\frac{(\tilde{U}_t^{\bar{x},\bar{a},\beta,n}(a'))_+ - (U_t^{\bar{x},\bar{a},\beta,n}(a'))_+}{\tilde{U}_t^{\bar{x},\bar{a},\beta,n}(a') - U_t^{\bar{x},\bar{a},\beta,n}(a')} 1_{\{\tilde{U}_t^{\bar{x},\bar{a},\beta,n}(a') - U_t^{\bar{x},\bar{a},\beta,n}(a') \neq 0\}}.$$

Then, we have $\nu \in \mathcal{V}_n$. Let us introduce the probability measure $\mathbb{P}^\nu$ equivalent to $\mathbb{P}$ on $(\Omega,\mathcal{F}_T)$, with $T \geq \tau$ (e.g., $T=\delta$), with Radon-Nikodym density given by (2.3). Then, taking the expectation $\mathbb{E}^\nu$ with respect to $\mathbb{P}^\nu$ in (B.30), (recalling that $\varphi(\bar{x},\bar{a}) = v^{\beta,n}(\bar{x},\bar{a})$)

$$\mathbb{E}^\nu\big[e^{-\beta\tau}\big(\varphi(X_\tau^{\bar{x},\bar{a}}, I_\tau^{\bar{a}}) - v^{\beta,n}(X_\tau^{\bar{x},\bar{a}}, I_\tau^{\bar{a}})\big)\big] \;\geq\; \varepsilon\mathbb{E}^\nu\bigg[\frac{1-e^{-\beta\tau}}{\beta}\bigg]. \tag{B.31}$$

Since $\tau > 0$, $\mathbb{P}$-a.s., we see that the right-hand side of (B.31) is strictly positive. On the other hand, from (B.26) it follows that the left-hand side of (B.31) is nonpositive, therefore we get a contradiction. □

B.5 Proof of Proposition 3.3

Firstly, we prove point (i). To this end, consider $Y^{x,a,\beta,n}$ and $Y^{x,a,\beta,n+1}$. It is useful to fix $T > 0$ and to look at the penalized BSDE (3.6) on $[0,\infty)$ solved by $Y^{x,a,\beta,n}$ (resp. $Y^{x,a,\beta,n+1}$) as a BSDE on $[0,T]$ with terminal condition $Y_T^{x,a,\beta,n}$ (resp. $Y_T^{x,a,\beta,n+1}$) and generator function f. Then, proceeding as in the proof of the comparison Theorem 2.5 of [26] for BSDEs with jumps on $[0,T]$, we can find a probability measure $\mathbb{P}^\nu$ equivalent to $\mathbb{P}$ on $(\Omega,\mathcal{F}_T)$, such that

$$Y_t^{x,a,\beta,n} - Y_t^{x,a,\beta,n+1} \;\leq\; \mathbb{E}^\nu\big[e^{-\beta(T-t)}\big(Y_T^{x,a,\beta,n} - Y_T^{x,a,\beta,n+1}\big)\big|\mathcal{F}_t\big], \tag{B.32}$$

$\mathbb{P}$-a.s., for all $t \in [0,T]$. Now, from estimate (2.6) and since $|Y_T^{x,a,\beta,n}|$, $|Y_T^{x,a,\beta,n+1}| \leq C_{b,\sigma,f}(1+|X_T^{x,a}|)/\beta$, letting $T\to\infty$ in (B.32), we obtain $Y_t^{x,a,\beta,n} \leq Y_t^{x,a,\beta,n+1}$, $\mathbb{P}$-a.s., for all $t \geq 0$. This shows that the sequence $(Y^{x,a,\beta,n})_n$ is monotone increasing. Since it is bounded by $C_{b,\sigma,f}(1+|X_T^{x,a}|)/\beta$, it converges increasingly to some adapted process $Y^{x,a,\beta}$ satisfying $|Y_t^{x,a,\beta}| \leq C_{b,\sigma,f}(1+|X_t^{x,a}|)/\beta$, for all $t \geq 0$.
Now, fix again $T > 0$ and consider the BSDE with nonpositive jumps (3.3)-(3.4), with

$(X,I) = (X^{x,a}, I^a)$, on $[0,T]$ with terminal condition $Y_T^{x,a,\beta}$ and generator function f. From Theorem 2.1 in [19] we know that there exists a unique minimal solution

$$(\tilde{Y}^{x,a,\beta,T}, \tilde{Z}^{x,a,\beta,T}, \tilde{U}^{x,a,\beta,T}, \tilde{K}^{x,a,\beta,T}) \in \mathbf{S^2(0,T)} \times \mathbf{L^2(W;0,T)} \times \mathbf{L^2}(\tilde{\mu};\mathbf{0,T}) \times \mathbf{K^2(0,T)}$$

to this BSDE. Moreover, $\tilde{Y}^{x,a,\beta,T}$ is the increasing limit[1] of $(Y^{x,a,\beta,n})_n$, so that $\tilde{Y}_t^{x,a,\beta,T} = Y_t^{x,a,\beta}$, $\mathbb{P}$-a.s., for all $t \in [0,T]$. We also know that $(\tilde{Z}^{x,a,\beta,T}, \tilde{U}^{x,a,\beta,T})$ is the strong (resp. weak) limit of $(Z^{x,a,\beta,n}, U^{x,a,\beta,n})_n$ in $\mathbf{L^p(W;0,T)} \times \mathbf{L^p}(\tilde{\mu};\mathbf{0,T})$, with $p \in [1,2)$, (resp. in $\mathbf{L^2(W;0,T)} \times \mathbf{L^2}(\tilde{\mu};\mathbf{0,T})$). This implies that $(\tilde{Z}^{x,a,\beta,T}, \tilde{U}^{x,a,\beta,T}) = (\tilde{Z}^{x,a,\beta,T'}_{|[0,T]}, \tilde{U}^{x,a,\beta,T'}_{|[0,T]})$, for all $T' \geq T$. Then, we define $(Z^{x,a,\beta}, U^{x,a,\beta}) \in \mathbf{L^2_{\text{loc}}(W)} \times \mathbf{L^2_{\text{loc}}}(\tilde{\mu})$ as

$$(Z^{x,a,\beta}_{|[0,T]}, U^{x,a,\beta}_{|[0,T]}) \;=\; (\tilde{Z}^{x,a,\beta,T}, \tilde{U}^{x,a,\beta,T}), \qquad \forall\, T > 0. \tag{B.33}$$

This proves point (ii). Concerning point (iii), from Theorem 2.1 in [19] we have that $\tilde{K}_t^{x,a,\beta,T}$ is the weak limit of $(K_t^{x,a,\beta,n})_n$ in $\mathbf{L^2}(\mathcal{F}_t)$, for any $0 \leq t \leq T$. It follows that $\tilde{K}_t^{x,a,\beta,T} = \tilde{K}_t^{x,a,\beta,T'}$, $\mathbb{P}$-a.s., for all $t \in [0,T]$ and for any $T' \geq T$. Therefore, we define $K^{x,a,\beta} \in \mathbf{K^2_{\text{loc}}}$ as follows: $K_t^{x,a,\beta} = \tilde{K}_t^{x,a,\beta,T}$, for all $t \in [0,T]$ and $T > 0$. We see that the quadruple $(Y^{x,a,\beta}, Z^{x,a,\beta}, U^{x,a,\beta}, K^{x,a,\beta})$ solves the backward equation (3.3) on $[0,\infty)$. Regarding the jump constraint (3.4), from Theorem 2.1 in [19] we know that

$$\tilde{U}_t^{x,a,\beta,T} \;\leq\; 0, \qquad dt \otimes d\mathbb{P} \otimes \vartheta(da)\text{-a.e.}$$

Then, from the definition (B.33) of $U^{x,a,\beta}$ we see that (3.4) holds. It remains to prove the minimality condition. Let $(\bar{Y}^{x,a,\beta}, \bar{Z}^{x,a,\beta}, \bar{U}^{x,a,\beta}, \bar{K}^{x,a,\beta}) \in \mathbf{S^2_{\text{loc}}} \times \mathbf{L^2_{\text{loc}}(W)} \times \mathbf{L^2_{\text{loc}}}(\tilde{\mu}) \times \mathbf{K^2_{\text{loc}}}$ be another solution to (3.3)-(3.4), with $|\bar{Y}_t^{x,a,\beta}| \leq C(1 + |X_t^{x,a}|)$, for all $t \geq 0$ and for some positive constant C (possibly depending on x, a, and β). Then, for any $T > 0$, $(\bar{Y}^{x,a,\beta}_{|[0,T]}, \bar{Z}^{x,a,\beta}_{|[0,T]}, \bar{U}^{x,a,\beta}_{|[0,T]}, \bar{K}^{x,a,\beta}_{|[0,T]})$ solves the BSDE (3.3)-(3.4) on $[0,T]$. As before, proceeding as in the proof of the comparison Theorem 2.5 of [26] for BSDEs with jumps on $[0,T]$, we can find a probability measure $\mathbb{P}^\nu$ equivalent to $\mathbb{P}$ on $(\Omega, \mathcal{F}_T)$, such that

$$Y_t^{x,a,\beta,n} - \bar{Y}_t^{x,a,\beta} \;\leq\; \mathbb{E}^\nu\big[e^{-\beta(T-t)}\big(Y_T^{x,a,\beta,n} - \bar{Y}_T^{x,a,\beta}\big)\big|\mathcal{F}_t\big], \tag{B.34}$$

$\mathbb{P}$-a.s., for all $t \in [0,T]$. From $|Y_T^{x,a,\beta,n}| \leq C_{b,\sigma,f}(1 + |X_T^{x,a}|)/\beta$, $|\bar{Y}_T^{x,a,\beta}| \leq C(1 + |X_T^{x,a}|)$ and estimate (2.6), letting $T \to \infty$ in (B.34) we obtain $Y_t^{x,a,\beta,n} \leq \bar{Y}_t^{x,a,\beta}$, $\mathbb{P}$-a.s., for all $t \geq 0$. Then, sending $n \to \infty$, we find $Y_t^{x,a,\beta} \leq \bar{Y}_t^{x,a,\beta}$, $\mathbb{P}$-a.s., for all $t \geq 0$, which proves the minimality of $(Y^{x,a,\beta}, Z^{x,a,\beta}, U^{x,a,\beta}, K^{x,a,\beta})$ and concludes the proof. □

[1] Notice that in Theorem 2.1 in [19], the terminal condition in the penalized BSDE does not depend on n; while, in our case, the penalized BSDE associated with $Y^{x,a,\beta,n}$ has the terminal condition $Y_T^{x,a,\beta,n}$. However, since $Y_T^{x,a,\beta,n}$ converges increasingly $\mathbb{P}$-a.s. to $Y_T^{x,a,\beta}$ as $n \to \infty$, the results of Theorem 2.1 in [19] are still valid.

References

[1] Arendt W. and C. J. Batty (1995): "A complex Tauberian theorem and mean ergodic semigroups", in *Semigroup Forum*, **50**, 351-366, Springer New York.

[2] Arisawa M. and P.-L. Lions (1998): "On ergodic stochastic control", *Comm. in Partial Differential Equations*, **23**, 2187-2217.

[3] Barles G. and P. Souganidis (2001): "Space-time periodic solutions and long time behavior of solutions to quasi-linear parabolic equations", *SIAM J. Math. Anal.*, **32**, 1311-1323.

[4] Bensoussan A. and J. Frehse (1992): "On Bellman equations of ergodic control in $\mathbb{R}^n$", *J. Reine Angew. Math.*, **429**, 125-160.

[5] Briand P. and Y. Hu (1998): "Stability of BSDEs with random terminal time and homogenization of semilinear elliptic PDEs", *Journal of Functional Analysis*, **155**, 455-494.

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

[7] Da Prato G. and J. Zabczyk (1992): *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications **44**, Cambridge University Press.

[8] Da Prato G. and J. Zabczyk (1996): *Ergodicity for infinite dimensional systems*, volume **229**, Cambridge University Press.

[9] Darling R. W. R. (1995): "Constructing gamma-martingales with prescribed limit, using backwards SDE", *The Annals of Probability*, **23**, 1234-1261.

[10] Fleming W. and S.J. Sheu (1999): "Optimal long term growth rate of expected utility of wealth", *Annals of Applied Probability*, **9**, 871-903.

[11] Fuhrman M., Hu Y., and G. Tessitore (2009): "Ergodic BSDEs and optimal ergodic control in Banach spaces", *SIAM J. Control Optim.*, **48**, 1542-1566.

[12] Fujita Y., Ishii H. and P. Loreti (2006): "Asymptotic solutions of Hamilton-Jacobi equations in euclidian n space", *Indiana Univ. Math. J.*, **55**, 1671-1700.

[13] Hata H., Nagai H. and S.J. Sheu (2010): "Asymptotics of the probability minimizing a down side risk", *Ann. Appl. Probab.*, **20**, 52-89.

[14] Hu Y., Madec P.-Y. and A. Richou (2014): "Large time behavior of mild solutions of Hamilton-Jacobi-Bellman equations in infinite dimension by a probabilistic approach", preprint arXiv:1406.5993.

[15] Ichihara N. (2012): "Large time asymptotic problems for optimal stochastic control with superlinear cost", *Stochastic Processes and their Applications*, **122**, 1248-1275.

[16] Ichihara N. and S.-J. Sheu (2013): "Large time behavior of solutions of Hamilton-Jacobi-Bellman equations with quadratic nonlinearity in gradients", *SIAM J. Math. Anal.*, **45**(1) 279-306.

[17] Ishii H. (1989): "On uniqueness and existence of viscosity solutions of fully nonlinear second order elliptic PDE's", *Communications on Pure and Applied Mathematics*, **42**, 15-45.

[18] Jacod J. and A. N. Shiryaev (2003): *Limit Theorems for Stochastic Processes*, second edition, Springer, Berlin.

[19] Kharroubi I. and H. Pham (2012): "Feynman-Kac representation for Hamilton-Jacobi-Bellman IPDEs", to appear in *The Annals of Probability.*

[20] Protter P. E. (2005): *Stochastic integration and differential equations*, volume 21 of Stochastic Modelling and Applied Probability. Springer-Verlag, Berlin. Second edition. Version 2.1, Corrected third printing.

[21] Nagai H. (2003): "Optimal strategies for risk-sensitive portfolio optimization problems with general factors", *SIAM J. Cont. and Optim.*, **41**, 1779-1800.

[22] Darling R. W. R. and E. Pardoux (1997): "Backwards SDE with random terminal time and applications to semilinear elliptic PDE", *The Annals of Probability*, **25**, 1135-1159.

[23] Richou A. (2009): "Ergodic BSDEs and related PDEs with Neumann boundary conditions", to appear in *Stochastic Processes and their Applications*.

[24] Robertson S. and H. Xing (2013): "Large time behavior of solutions to semi-linear equations with quadratic growth in the gradient", preprint.

[25] Royer M. (2004): "BSDEs with a random terminal time driven by a monotone generator and their links with PDEs", *Stochastics and stochastic reports*, **76**, 281-307.

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

[27] Safonov M. V. (1988): "On the classical solution of nonlinear elliptic equations of second order", *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya*, **52**, 1272-1287.

[28] Souplet P. and Q. Zhang (2006): "Global solutions of inhomogeneous Hamilton-Jacobi equations", *J. Anal. Math.*, **99**, 355-396.

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