



**HAL**  
open science

# Path-dependent Hamilton-Jacobi-Bellman equation: Uniqueness of Crandall-Lions viscosity solutions

Andrea Cosso, Fausto Gozzi, Mauro Rosostolato, Francesco Russo

► **To cite this version:**

Andrea Cosso, Fausto Gozzi, Mauro Rosostolato, Francesco Russo. Path-dependent Hamilton-Jacobi-Bellman equation: Uniqueness of Crandall-Lions viscosity solutions. 2021. hal-03285204v1

**HAL Id: hal-03285204**

**<https://hal.science/hal-03285204v1>**

Preprint submitted on 13 Jul 2021 (v1), last revised 3 Aug 2023 (v3)

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

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

# Path-dependent Hamilton-Jacobi-Bellman equation: Uniqueness of Crandall-Lions viscosity solutions

Andrea COSSO\* Fausto GOZZI† Mauro ROSESTOLATO‡ Francesco RUSSO§

July 13, 2021

## Abstract

We prove existence and uniqueness of Crandall-Lions viscosity solutions of Hamilton-Jacobi-Bellman equations in the space of continuous paths, associated to the optimal control of path-dependent SDEs. This seems the first uniqueness result in such a context. More precisely, similarly to the seminal paper [37], the proof of our core result, that is the comparison theorem, is based on the fact that the value function is bigger than any viscosity subsolution and smaller than any viscosity supersolution. Such a result, coupled with the proof that the value function is a viscosity solution (based on the dynamic programming principle, which we prove), implies that the value function is the unique viscosity solution to the Hamilton-Jacobi-Bellman equation. The proof of the comparison theorem in [37] relies on regularity results which are missing in the present infinite-dimensional context, as well as on the local compactness of the finite-dimensional underlying space. We overcome such non-trivial technical difficulties introducing a suitable approximating procedure and a smooth gauge-type function, which allows to generate maxima and minima through an appropriate version of the Borwein-Preiss generalization of Ekeland's variational principle on the space of continuous paths.

**Mathematics Subject Classification (2020):** 93E20, 49L25, 35R15.

**Keywords:** path-dependent SDEs, dynamic programming principle, pathwise derivatives, functional Itô calculus, path-dependent HJB equations, viscosity solutions.

---

\*University of Bologna, Department of Mathematics, Piazza di Porta San Donato 5, 40126 Bologna, Italy; andrea.cosso@unibo.it.

†Luiss University, Department of Economics and Finance, Rome, Italy; fgozzi@luiss.it.

‡Università del Salento, Dipartimento di Matematica e Fisica “Ennio De Giorgi”, 73100 Lecce, Italy; mauro.rosestolato@unisalento.it.

§ENSTA Paris, Institut Polytechnique de Paris, Unité de Mathématiques Appliquées, 828, bd. des Maréchaux, F-91120 Palaiseau, France; francesco.russo@ensta-paris.fr.

# 1 Introduction

The optimal control of path-dependent SDEs arises frequently in applications (for instance in Economics and Finance) where the dynamics are non-Markovian. Such non-Markovianity makes difficult to apply the dynamic programming approach to those problems. Indeed, the standard dynamic programming approach is designed when the state equation is Markovian hence it cannot be applied to such problems as it is.

More precisely, consider the following SDE on a complete probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  where a  $m$ -dimensional Brownian motion  $B = (B_t)_{t \geq 0}$  is defined. Let  $T > 0$ ,  $t \in [0, T]$ ,  $x \in C([0, T]; \mathbb{R}^d)$ , and consider a progressively measurable process  $\alpha: [0, T] \times \Omega \rightarrow A$  (with  $A$  being a Polish space), where  $x$  is the initial path and  $\alpha$  the control process. Let the state process  $X: [0, T] \times \Omega \rightarrow \mathbb{R}^d$  satisfy the following controlled path-dependent SDE:

$$\begin{cases} dX_s = b(s, X, \alpha_s) ds + \sigma(s, X, \alpha_s) dB_s, & s \in (t, T], \\ X_s = x(s), & s \in [0, t]. \end{cases}$$

Here  $X$  denotes the whole path, which, under mild assumptions, belongs to  $C([0, T]; \mathbb{R}^d)$ . We assume  $b: [0, T] \times C([0, T]; \mathbb{R}^d) \times A \rightarrow \mathbb{R}^d$  (as well as  $\sigma$ ) to be non-anticipative, namely, for all  $s \in [0, T]$ ,  $a \in A$ ,  $b(s, x, a)$  and  $\sigma(s, x, a)$  depend on the path  $x \in C([0, T]; \mathbb{R}^d)$  only up to time  $s$ .

The stochastic optimal control problem consists in maximizing the reward functional

$$J(t, x, \alpha) = \mathbb{E} \left[ \int_t^T f(s, X^{t,x,\alpha}, \alpha_s) ds + g(X^{t,x,\alpha}) \right],$$

with  $f$  being non-anticipative, as  $b$  and  $\sigma$  above. The value function is then defined as

$$v(t, x) = \sup_{\alpha} J(t, x, \alpha), \quad \forall (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d),$$

where the supremum is taken over all progressively measurable control processes  $\alpha$ .

We see that the value function is defined on the infinite-dimensional space of continuous paths  $C([0, T]; \mathbb{R}^d)$ , hence it is related to some Hamilton-Jacobi-Bellman equation (HJB equation for short) in infinite dimension.

The “standard” approach to study such problems consists in changing state space transforming the path-dependent SDE into a Markovian SDE, formulated on an infinite-dimensional space  $\mathcal{H}$ , typically  $C([0, T]; \mathbb{R}^d)$  or  $\mathbb{R}^d \times L^2(0, T; \mathbb{R}^d)$ . In this case the associated Hamilton-Jacobi-Bellman equation is a PDE in infinite dimension (see for instance [17, 29]) which contains “standard” Fréchet derivatives in the space  $\mathcal{H}$ . Some results on the viscosity solution approach are given for instance in [30, 31, 45]; however, uniqueness results seems not available up to now, see the discussion in [29, Section 3.14, pages 363-364]).

More recently, another approach has been developed after the seminal work of Dupire [21], which is based on the introduction of a different notion of “finite-dimensional” derivatives (known as horizontal/vertical derivatives) which allows to write the associated Hamilton-Jacobi-Bellman equation without using the derivatives in the space  $\mathcal{H}$ . We call such an

equation a path-dependent Hamilton-Jacobi-Bellman equation (see equation (3.5) below), which belongs to the more general class of path-dependent partial differential equations, that is PDEs where the unknown depends on the paths and the involved derivatives are the Dupire horizontal and vertical derivatives. The definitions of these derivatives will be recalled in Appendix A. There are also other approaches, similar to that introduced by Dupire, but based on slightly different notions of derivatives, see in particular [1, 38, 33].

The theory of path-dependent PDEs is very recent, yet there are already many papers on this subject, see for instance [19, 20, 46, 39, 8, 40, 23, 24, 25, 32, 41, 12, 13, 44, 43, 42, 2, 9, 3, 11]. One stream in the literature (starting with [23] and further developed in [24, 25, 41, 44, 43, 9]) looks at such equations using a modified definition of viscosity solution where maxima and minima are taken in expectation. In this way, roughly speaking, the amount of test functions increases and, hence, uniqueness is easier to prove.

Another stream in the literature looks at path-dependent PDEs using the “standard” definition of viscosity solution adapted to the new derivatives. We call such a definition the “Crandall-Lions” one, recalling for instance their papers [15, 16]. In such a context there are only two papers, namely [11], which only address the path-dependent heat equation, and [47], even if, unfortunately, the present version of this last paper (v1) seems to contain some relevant gaps.

In the present paper we look at this last stream proving existence and uniqueness of Crandall-Lions viscosity solutions of HJB equations associated to the optimal control of path-dependent SDEs. This seems the first uniqueness result in such a context. The proof of uniqueness (or, more precisely, of the comparison theorem, from which uniqueness is derived) is difficult due to the fact that the usual approach adopted in the theory of viscosity solutions relies on fine properties of functions on  $\mathbb{R}^d$ , as for instance Aleksandrov’s theorem and Jensen’s lemma (see on this [14, Appendix, pages 56-58]). The extension of those results to functions defined on the space of continuous paths seems however impracticable (this is probably the reason why, to circumvent such a problem, other notions of solutions have been introduced in the literature). The proposed methodology is instead built on refinements of the original approach developed in [37] and is based on the existence of the candidate solution  $v$ , which is shown to be bigger than any subsolution and smaller than any supersolution. The latter is traditionally based on regularity results which are missing in the present context as well as on the local compactness of the underlying space in order to generate maxima or minima. We overcome those non-trivial technical difficulties firstly relying on suitable approximating procedures, see Lemmas B.3-B.4-B.5 and Theorem B.6 of Appendix B. Moreover, concerning the existence of maxima or minima, instead of the missing local compactness of the underlying space, we exploit its completeness relying on a novel variational principle on  $[0, T] \times C([0, T]; \mathbb{R}^d)$  proved in [11] (see Theorem 4.4), based on the existence of a suitable smooth gauge-type function (see Lemma 4.2) and on a variant of the Borwein-Preiss generalization of Ekeland’s variational principle (see [5, Theorem 2.5.2]).

Once the comparison theorem is proved, we deduce from our existence result (Theorem 3.3) that the value function  $v$  is the unique Crandall-Lions viscosity solution of the path-dependent HJB equation. The existence result is based, as usual, on the dynamic

programming principle, which is proved rigorously in the present paper, see Theorem 2.6.

The rest of the paper is organized as follows. In Section 2 we formulate the stochastic optimal control problem of path-dependent SDEs and prove the dynamic programming principle. In Section 3 we introduce the notion of Crandall-Lions viscosity solution and prove that the value function  $v$  solves in the viscosity sense the path-dependent Hamilton-Jacobi-Bellman equation. In Section 4 we state the smooth variational principle on  $[0, T] \times C([0, T]; \mathbb{R}^d)$  and prove the comparison theorem, from which the uniqueness result follows. In Appendix A we recall the definitions of horizontal and vertical derivatives together with the functional Itô formula. Finally, in Appendix B we report all the results concerning the approximation of the value function needed in the proof of the comparison theorem.

## 2 Path dependent stochastic optimal control problems

### 2.1 Notations and basic setting

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a complete probability space on which a  $m$ -dimensional Brownian motion  $B = (B_t)_{t \geq 0}$  is defined. Let  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  denote the  $\mathbb{P}$ -completion of the filtration generated by  $B$ . Notice that  $\mathbb{F}$  is right-continuous, so that it satisfies the usual conditions. Furthermore, let  $T > 0$  and let  $A$  be a Polish space, with  $\mathcal{B}(A)$  being its Borel  $\sigma$ -algebra. We denote by  $\mathcal{A}$  the family of all  $\mathbb{F}$ -progressively measurable processes  $\alpha: [0, T] \times \Omega \rightarrow A$ . Finally, for every  $p \geq 1$ , we denote by  $\mathbf{S}_p(\mathbb{F})$  the set of  $d$ -dimensional continuous  $\mathbb{F}$ -progressively measurable processes  $X: [0, T] \times \Omega \rightarrow \mathbb{R}^d$  such that

$$\|X\|_{\mathbf{S}_p} := \mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t|^p \right]^{1/p} < \infty. \quad (2.1)$$

The *state space* of the stochastic optimal control problem is the set  $C([0, T]; \mathbb{R}^d)$  of continuous  $d$ -dimensional paths on  $[0, T]$ . For every  $x \in C([0, T]; \mathbb{R}^d)$  and  $t \in [0, T]$ , we denote by  $x(t)$  or  $x_t$  the value of  $x$  at time  $t$  and we set  $x(\cdot \wedge t) := (x(s \wedge t))_{s \in [0, T]}$  or  $x_{\cdot \wedge t} := (x(s \wedge t))_{s \in [0, T]}$ . Observe that  $x(t)$  (or  $x_t$ ) is an element of  $\mathbb{R}^d$ , while  $x(\cdot \wedge t)$  (or  $x_{\cdot \wedge t}$ ) belongs to  $C([0, T]; \mathbb{R}^d)$ . We endow  $C([0, T]; \mathbb{R}^d)$  with the supremum norm  $\|\cdot\|_T$  defined as

$$\|x\|_T = \sup_{s \in [0, T]} |x(s)|, \quad x \in C([0, T]; \mathbb{R}^d),$$

where  $|x(s)|$  denotes the Euclidean norm of  $x(s)$  in  $\mathbb{R}^d$ . We remark that  $(C([0, T]; \mathbb{R}^d), \|\cdot\|_T)$  is a Banach space and we denote by  $\mathcal{B}$  its Borel  $\sigma$ -algebra. We also define, for every  $t \in [0, T]$ , the seminorm  $\|\cdot\|_t$  as

$$\|x\|_t = \|x_{\cdot \wedge t}\|_T, \quad x \in C([0, T]; \mathbb{R}^d).$$

Finally, on  $[0, T] \times C([0, T]; \mathbb{R}^d)$  we define the pseudometric  $d_\infty: ([0, T] \times C([0, T]; \mathbb{R}^d))^2 \rightarrow [0, \infty)$  as

$$d_\infty((t, x), (t', x')) := |t - t'| + \|x(\cdot \wedge t) - x'(\cdot \wedge t')\|_T.$$

We refer to [11, Section 2.1] for more details on such a pseudometric.

## 2.2 Assumptions and state equation

We consider the coefficients

$$b, \sigma, f: [0, T] \times C([0, T]; \mathbb{R}^d) \times A \longrightarrow \mathbb{R}^d, \mathbb{R}^{d \times m}, \mathbb{R}, \quad g: C([0, T]; \mathbb{R}^d) \longrightarrow \mathbb{R},$$

on which we impose the following assumptions.

**Assumption (A).**

- (i) *The maps  $b, \sigma, f, g$  are continuous.*
- (ii) *There exist a constant  $K \geq 0$  such that*

$$\begin{aligned} |b(t, x, a) - b(t, x', a)| + |\sigma(t, x, a) - \sigma(t, x', a)| + |f(t, x, a) - f(t, x', a)| &\leq K \|x - x'\|_t, \\ |g(x) - g(x')| &\leq K \|x - x'\|_T, \\ |b(t, 0, a)| + |\sigma(t, 0, a)| + |f(t, x, a)| + |g(x)| &\leq K, \end{aligned}$$

for all  $a \in A, (t, x), (t', x') \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , with  $|\sigma(t, x, a)| := (\text{tr}(\sigma\sigma^\top)(t, x, a))^{1/2} = (\sum_{i,j} |\sigma_{i,j}(t, x, a)|^2)^{1/2}$  denoting the Frobenius norm of  $\sigma(t, x, a)$ .

**Assumption (B).** *The maps  $b, \sigma, f$  are uniformly continuous in  $t$ , uniformly with respect to the other variables. In particular, there exists a modulus of continuity  $w: [0, \infty) \rightarrow [0, \infty)$  such that*

$$|b(t, x, a) - b(s, x, a)| + |\sigma(t, x, a) - \sigma(s, x, a)| + |f(t, x, a) - f(s, x, a)| \leq w(|t - s|),$$

for all  $t, s \in [0, T], x \in C([0, T]; \mathbb{R}^d), a \in A$ . The map  $w$  is continuous, increasing, subadditive, and  $w(0) = 0$ .

**Remark 2.1.** *By the Lipschitz continuity of  $b, \sigma, f$ , we deduce that they satisfy the following non-anticipativity condition:*

$$b(t, x, a) = b(t, x_{\cdot \wedge t}, a), \quad \sigma(t, x, a) = \sigma(t, x_{\cdot \wedge t}, a), \quad f(t, x, a) = f(t, x_{\cdot \wedge t}, a),$$

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

For every  $t \in [0, T], \xi \in \mathbf{S}_2(\mathbb{F}), \alpha \in \mathcal{A}$ , the state process satisfies the following system of controlled stochastic differential equations:

$$\begin{cases} dX_s = b(s, X, \alpha_s) ds + \sigma(s, X, \alpha_s) dB_s, & s \in (t, T], \\ X_s = \xi_s, & s \in [0, t]. \end{cases} \quad (2.2)$$

**Proposition 2.2.** *Suppose that Assumption (A) holds. Then, for every  $t \in [0, T], \xi \in \mathbf{S}_2(\mathcal{F}_t), \alpha \in \mathcal{A}$ , there exists a unique solution  $X^{t, \xi, \alpha} \in \mathbf{S}_2(\mathbb{F})$  to equation (2.2). Moreover, for any  $\xi' \in C([0, T]; \mathbb{R}^d)$  it holds that*

$$\|X^{t, \xi, \alpha} - X^{t, \xi', \alpha}\|_{\mathbf{S}_2} \leq c \|\xi_{\cdot \wedge t} - \xi'_{\cdot \wedge t}\|_{\mathbf{S}_2}, \quad (2.3)$$

for some constant  $c$ , independent of  $t, \xi, \xi', \alpha$ . Finally, it holds that

$$\lim_{r \rightarrow t^+} \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \sup_{0 \leq s \leq T} |X_{s \wedge r}^{t, \xi, \alpha} - \xi_{s \wedge t}|^2 \right] = 0. \quad (2.4)$$

*Proof.* See [10, Proposition 2.8] for the existence and uniqueness result, together with estimate (2.3). Finally, concerning (2.4) we refer to [10, Remark 2.9].  $\square$

## 2.3 Value function

Given  $t \in [0, T]$  and  $x \in C([0, T]; \mathbb{R}^d)$ , the stochastic optimal control problem consists in finding  $\alpha \in \mathcal{A}$  maximizing the following functional:

$$J(t, x, \alpha) = \mathbb{E} \left[ \int_t^T f(s, X^{t,x,\alpha}, \alpha_s) ds + g(X^{t,x,\alpha}) \right].$$

Finally, the value function is defined as

$$v(t, x) = \sup_{\alpha \in \mathcal{A}} J(t, x, \alpha), \quad \forall (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d). \quad (2.5)$$

**Proposition 2.3.** *Suppose that Assumption (A) holds. Then, the value function  $v$  is bounded, jointly continuous on  $([0, T] \times C([0, T]; \mathbb{R}^d), d_\infty)$ , and there exists a constant  $L \geq 0$  (depending only on  $T, K, c$  in (2.3)) such that*

$$|v(t, x) - v(t, x')| \leq L \|x - x'\|_t, \quad (2.6)$$

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

*Proof.* The boundedness of  $v$  follows directly from the boundedness of  $f$  and  $g$ . On the other hand, the joint continuity of  $v$  can be deduced from [10, Proposition 3.3]. Finally, (2.6) is a consequence of the Lipschitz continuity of  $f, g$  and of estimate (2.3).  $\square$

## 2.4 Dynamic programming principle

In Section 3, Theorem 3.3, we prove that the value function  $v$  is a viscosity solution of a suitable path-dependent Hamilton-Jacobi-Bellman equation. The proof of this property is standard and it is based, as usual, on the dynamic programming principle which is stated below. We prove it relying on [10, Theorem 3.4] and on the two next technical Lemmata 2.4 and 2.5. For other rigorous proofs of the dynamic programming principle in the path-dependent case we refer to [26, 27].

We begin introducing some notations. For every  $t \in [0, T]$ , let  $\mathbb{F}^t = (\mathcal{F}_s^t)_{s \in [0, T]}$  be the  $\mathbb{P}$ -completion of the filtration generated by  $(B_{s \vee t} - B_t)_{s \in [0, T]}$ . Let also  $Prog(\mathbb{F}^t)$  denote the  $\sigma$ -algebra of  $[t, T] \times \Omega$  of all  $(\mathcal{F}_s^t)_{s \in [t, T]}$ -progressive sets. Finally, let  $\mathcal{A}_t$  be the subset of  $\mathcal{A}$  of all  $\mathbb{F}^t$ -progressively measurable processes.

**Lemma 2.4.** *Suppose that Assumption (A) holds. Then, the value function defined by (2.5) satisfies*

$$v(t, x) = \sup_{\alpha \in \mathcal{A}_t} J(t, x, \alpha), \quad \forall (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d). \quad (2.7)$$

**Proof.** Fix  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ . Since  $\mathcal{A}_t \subset \mathcal{A}$ , we see that  $v(t, x) \geq \sup_{\alpha \in \mathcal{A}_t} J(t, x, \alpha)$ . It remains to prove the reverse inequality

$$v(t, x) \leq \sup_{\alpha \in \mathcal{A}_t} J(t, x, \alpha). \quad (2.8)$$

We split the proof of (2.8) into four steps.

STEP I. *Additional notations.* We firstly fix some notations. Let  $(\mathbb{R}^m)^{[0, t]}$  be the set of functions from  $[0, t]$  to  $\mathbb{R}^m$ , endowed with the product  $\sigma$ -algebra  $\mathcal{B}(\mathbb{R}^m)^{[0, t]}$  generated by the finite-dimensional cylindrical sets of the form:  $C_{t_1, \dots, t_n}(H) = \{y \in (\mathbb{R}^m)^{[0, t]} : (y(t_1), \dots, y(t_n)) \in H\}$ , for some  $t_i \in [0, t]$ ,  $H = H_{t_1} \times \dots \times H_{t_n}$ ,  $H_{t_i} \in \mathcal{B}(\mathbb{R}^m)$ . Now, consider the map  $\mathbf{B}^t : \Omega \rightarrow (\mathbb{R}^m)^{[0, t]}$  defined as follows:

$$\mathbf{B}^t : \omega \longmapsto (B_s(\omega))_{0 \leq s \leq t}.$$

Such a map is measurable with respect to  $\mathcal{F}_t$ , as a matter of fact the counterimage through  $\mathbf{B}^t$  of a finite-dimensional cylindrical set  $C_{t_1, \dots, t_n}(H)$  clearly belongs to  $\mathcal{F}_t$ . In addition, the  $\sigma$ -algebra generated by  $\mathbf{B}^t$  coincides with  $\mathcal{G}_t := \sigma(B_s, 0 \leq s \leq t)$ . Notice that

$$\mathcal{F}_t = \mathcal{G}_t \vee \mathcal{N},$$

where  $\mathcal{N}$  is the family of  $\mathbb{P}$ -null sets.

Finally, let  $(E^t, \mathcal{E}^t)$  be the measurable space given by  $E^t = [t, T] \times \Omega$  and  $\mathcal{E}^t = \text{Prog}(\mathbb{F}^t)$ . Then, we denote by  $\mathcal{I}^t : E^t \rightarrow E^t$  the identity map.

STEP II. *Representation of  $\alpha$ .* Given  $\alpha \in \mathcal{A}$ , let us prove that there exists a map  $\mathbf{a}^t : [t, T] \times \Omega \times (\mathbb{R}^m)^{[0, t]} \rightarrow A$  such that:

- 1)  $\mathbf{a}^t$  is measurable with respect to the product  $\sigma$ -algebra  $\text{Prog}(\mathbb{F}^t) \otimes \mathcal{B}(\mathbb{R}^m)^{[0, t]}$ ;
- 2) the processes  $\alpha_{|[t, T]}$  (denoting the restriction of  $\alpha$  to  $[t, T]$ ) and  $(\mathbf{a}^t(s, \cdot, \mathbf{B}^t))_{s \in [t, T]}$  are indistinguishable.

In order to prove the existence of such a map  $\mathbf{a}^t$ , we begin noticing that the following holds:

$$\mathcal{F}_s = \mathcal{F}_t \vee \mathcal{F}_s^t = \mathcal{G}_t \vee \mathcal{F}_s^t, \quad \forall s \in [t, T],$$

where the second equality follows from the fact that  $\mathcal{N}$ , the family of  $\mathbb{P}$ -null sets, is contained in both  $\mathcal{F}_t$  and  $\mathcal{F}_s^t$ . Recalling that  $\alpha$  is  $\mathbb{F}$ -progressively measurable, we have that  $\alpha_{|[t, T]}$  is progressively measurable with respect to the filtration

$$\sigma(\mathcal{G}_t \vee \mathcal{F}_{t \vee s}^t)_{s \in [t, T]}.$$

In other words, the map  $\alpha_{|[t, T]} : [t, T] \times \Omega \rightarrow A$  is  $\text{Prog}(\mathbb{F}^t) \vee (\{\emptyset, [t, T]\} \otimes \mathcal{G}_t)$ -measurable, with  $\{\emptyset, [t, T]\}$  denoting the trivial  $\sigma$ -algebra on  $[t, T]$ .

Now, recall the definitions of  $\mathcal{I}^t$  and  $\mathbf{B}^t$  from STEP I, and let denote still by the same symbol  $\mathbf{B}^t$  the canonical extension of  $\mathbf{B}^t$  to  $[t, T] \times \Omega$  (or, equivalently, to  $E^t$ ), defined as



$\mathbf{B}^t: [t, T] \times \Omega \rightarrow (\mathbb{R}^m)^{[0,t]}$  with  $(t, \omega) \mapsto \mathbf{B}^t(\omega)$ . Then, the  $\sigma$ -algebra generated by the pair  $(\mathcal{I}^t, \mathbf{B}^t): [t, T] \times \Omega \rightarrow E^t \times (\mathbb{R}^m)^{[0,t]}$  coincides with  $Prog(\mathbb{F}^t) \vee (\{\emptyset, [t, T]\} \otimes \mathcal{G}_t)$ . Therefore, by Doob's measurability theorem (see for instance [34, Lemma 1.13]) it follows that the restriction of  $\alpha$  to  $[t, T]$  can be represented as follows:  $\alpha|_{[t,T]} = \mathbf{a}^t(\mathcal{I}^t, \mathbf{B}^t)$ , for some map  $\mathbf{a}^t: [t, T] \times \Omega \times (\mathbb{R}^m)^{[0,t]} \rightarrow A$  satisfying items 1)-2) above.

STEP III. *The stochastic process  $X^{t,x,\mathbf{B}^t}$ .* Given  $\alpha \in \mathcal{A}$ , let  $\mathbf{a}^t$  be as in STEP II. For every  $y \in (\mathbb{R}^m)^{[0,t]}$ , let  $X^{t,x,y}$  be the unique solution in  $\mathbf{S}_2(\mathbb{F})$  to the following equation:

$$\begin{cases} dX_s = b(s, X, \mathbf{a}^t(s, \cdot, y)) ds + \sigma(s, X, \mathbf{a}^t(s, \cdot, y)) dB_s, & s \in (t, T], \\ X_s = x(s), & s \in [0, t]. \end{cases} \quad (2.9)$$

From the proof (see [10, Proposition 2.8]) of the existence of a solution to equation (2.9), based on a fixed point argument, we can also deduce that the random field  $X: [0, T] \times \Omega \times (\mathbb{R}^m)^{[0,t]} \rightarrow \mathbb{R}^d$  is measurable with respect to the product  $\sigma$ -algebra  $Prog(\mathbb{F}^t) \otimes \mathcal{B}(\mathbb{R}^m)^{[0,t]}$ . As a consequence, we can consider the composition of  $X^{t,x,y}$  and  $\mathbf{B}^t$ , denoted  $X^{t,x,\mathbf{B}^t}$ . Using the independence of  $\mathcal{G}_t = \sigma(\mathbf{B}^t)$  and  $\mathcal{F}_T^t$ , we deduce that the process  $X^{t,x,\mathbf{B}^t}$  satisfies the following equation:

$$\begin{cases} dX_s = b(s, X, \mathbf{a}^t(s, \cdot, \mathbf{B}^t)) ds + \sigma(s, X, \mathbf{a}^t(s, \cdot, \mathbf{B}^t)) dB_s, & s \in (t, T], \\ X_s = x(s), & s \in [0, t]. \end{cases}$$

Recalling from STEP II that  $\alpha|_{[t,T]}$  and  $(\mathbf{a}^t(s, \cdot, \mathbf{B}^t))_{s \in [t,T]}$  are indistinguishable, and noticing that the solution to equation (2.10) below depends on  $\alpha$  only through its values on  $[t, T]$  (namely, it depends only on  $\alpha|_{[t,T]}$ ), we conclude that  $X^{t,x,\mathbf{B}^t}$  solves the same equation of  $X^{t,x,\alpha}$ , namely

$$\begin{cases} dX_s = b(s, X, \alpha_s) ds + \sigma(s, X, \alpha_s) dB_s, & s \in (t, T], \\ X_s = x(s), & s \in [0, t]. \end{cases} \quad (2.10)$$

From pathwise uniqueness for equation (2.10), we get that  $X^{t,x,\mathbf{B}^t}$  and  $X^{t,x,\alpha}$  are also indistinguishable.

STEP IV. *The stochastic process  $X^{t,x,\mathbf{B}^t}$ .* Given  $\alpha \in \mathcal{A}$ , let  $\mathbf{a}^t$  be as in STEP II and  $X^{t,x,\mathbf{B}^t}$  as in STEP III. Then, we have

$$\begin{aligned} J(t, x, \alpha) &= \mathbb{E} \left[ \int_t^T f(s, X^{t,x,\alpha}, \alpha_s) ds + g(X^{t,x,\alpha}) \right] \\ &= \mathbb{E} \left[ \int_t^T f(s, X^{t,x,\mathbf{B}^t}, \mathbf{a}^t(s, \cdot, \mathbf{B}^t)) ds + g(X^{t,x,\mathbf{B}^t}) \right]. \end{aligned}$$

Denoting by  $\boldsymbol{\mu}^t$  the probability distribution of  $\mathbf{B}^t$  on  $((\mathbb{R}^m)^{[0,t]}, \mathcal{B}(\mathbb{R}^m)^{[0,t]})$ , and recalling the

independence of  $\mathcal{G}_t = \sigma(\mathbf{B}^t)$  and  $\mathcal{F}_T^t$ , by Fubini's theorem we obtain

$$\begin{aligned} & \mathbb{E} \left[ \int_t^T f(s, X^{t,x,\mathbf{B}^t}, \mathbf{a}^t(s, \cdot, \mathbf{B}^t)) ds + g(X^{t,x,\mathbf{B}^t}) \right] \\ &= \int_{(\mathbb{R}^m)^{[0,t]}} \mathbb{E} \left[ \int_t^T f(s, X^{t,x,y}, \mathbf{a}^t(s, \cdot, y)) ds + g(X^{t,x,y}) \right] \boldsymbol{\mu}^t(dy). \end{aligned}$$

Now, fix some  $a_0 \in A$  and, for every  $y \in (\mathbb{R}^m)^{[0,t]}$ , denote

$$\beta_s^y := a_0 1_{[0,t)}(s) + \mathbf{a}^t(s, \cdot, y) 1_{[t,T]}, \quad \forall s \in [0, T].$$

Notice that  $\beta^y \in \mathcal{A}_t$ . Moreover, recalling that  $X^{t,x,y}$  solves equation (2.9), we see that it solves the same equation of  $X^{t,x,\beta^y}$ . Then, by pathwise uniqueness,  $X^{t,x,y}$  and  $X^{t,x,\beta^y}$  are indistinguishable. In conclusion, we obtain

$$\begin{aligned} & \int_{(\mathbb{R}^m)^{[0,t]}} \mathbb{E} \left[ \int_t^T f(s, X^{t,x,y}, \mathbf{a}^t(s, \cdot, y)) ds + g(X^{t,x,y}) \right] \boldsymbol{\mu}^t(dy) \\ &= \int_{(\mathbb{R}^m)^{[0,t]}} \mathbb{E} \left[ \int_t^T f(s, X^{t,x,\beta^y}, \beta_s^y) ds + g(X^{t,x,\beta^y}) \right] \boldsymbol{\mu}^t(dy) \\ &= \int_{(\mathbb{R}^m)^{[0,t]}} J(t, x, \beta^y) \boldsymbol{\mu}^t(dy) \leq \int_{(\mathbb{R}^m)^{[0,t]}} \sup_{\gamma \in \mathcal{A}_t} J(t, x, \gamma) \boldsymbol{\mu}^t(dy) = \sup_{\gamma \in \mathcal{A}_t} J(t, x, \gamma). \end{aligned}$$

This proves that  $J(t, x, \alpha) \leq \sup_{\gamma \in \mathcal{A}_t} J(t, x, \gamma)$ , for every  $\alpha \in \mathcal{A}$ . Then, inequality (2.8) follows from the arbitrariness of  $\alpha$ .  $\square$

Next lemma expresses in terms of  $v$  the value of the optimal control problem formulated at time  $t$ , with *random* initial condition  $\xi \in \mathbf{S}_2(\mathbb{F})$ . In order to state such a lemma, we introduce the function  $V: [0, T] \times \mathbf{S}_2(\mathbb{F}) \rightarrow \mathbb{R}$  defined as follows:

$$V(t, \xi) = \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \int_t^T f(s, X^{t,\xi,\alpha}, \alpha_s) dr + g(X^{t,\xi,\alpha}) \right], \quad (2.11)$$

for every  $t \in [0, T]$ ,  $\xi \in \mathbf{S}_2(\mathbb{F})$ . Clearly, when  $\xi \equiv x \in C([0, T]; \mathbb{R}^d)$  we have  $V(t, x) = v(t, x)$ .

**Lemma 2.5.** *Suppose that Assumption (A) holds. Let  $t \in [0, T]$  and  $\xi \in \mathbf{S}_2(\mathbb{F})$ , then*

$$V(t, \xi) = \mathbb{E}[v(t, \xi)]. \quad (2.12)$$

**Proof.** We begin noting that for  $t = 0$  it is clear that equality (2.12) holds true, as a matter of fact  $\mathcal{F}_0$  is the family of  $\mathbb{P}$ -null sets, therefore  $\xi$  is a.s. equal to a constant and (2.12) follows from the fact that  $V(t, x) = v(t, x)$ , for every  $x \in C([0, T]; \mathbb{R}^d)$ . For this reason, in the sequel we suppose that  $t > 0$ . We split the rest of the proof into four steps.

**STEP I. Additional notations.** Firstly, we fix some notations. For a fixed  $t \in (0, T]$ , let  $\mathbb{G}^t = (\mathcal{G}_s^t)_{s \geq 0}$  be given by

$$\mathcal{G}_s^t := \mathcal{F}_t \vee \mathcal{F}_s^t, \quad \forall s \geq 0.$$

Moreover, let  $\mathbf{S}_2(\mathbb{G}^t)$  (resp.  $\mathbf{S}_2(\mathcal{F}_t)$ ) be the set of  $d$ -dimensional continuous  $\mathbb{G}^t$ -progressively measurable (resp.  $\mathcal{B}([0, T]) \otimes \mathcal{F}_t$ -measurable) processes  $X: [0, T] \times \Omega \rightarrow \mathbb{R}^d$  satisfying the integrability condition (2.1). Notice that Proposition 2.2 extends to the case with initial condition  $\xi \in \mathbf{S}_2(\mathcal{F}_t)$  rather than  $\xi \in \mathbf{S}_2(\mathbb{F})$ . In particular, given  $\xi \in \mathbf{S}_2(\mathcal{F}_t)$  and  $\alpha \in \mathcal{A}$ , equation (2.2) admits a unique solution  $X^{t, \xi, \alpha} \in \mathbf{S}_2(\mathbb{G}^t)$ . Then, for  $\xi \in \mathbf{S}_2(\mathcal{F}_t)$  we define  $V(t, \xi)$  as in (2.11).

STEP II. *Preliminary remarks.* We begin noting that  $X^{t, \xi, \alpha} = X^{t, \xi \wedge t, \alpha}$ , so that it is enough to prove equality (2.12) with  $\xi \wedge t$  in place of  $\xi$ . More generally, we shall prove the validity of (2.12) in the case when  $\xi \in \mathbf{S}_2(\mathcal{F}_t)$ .

Now, recall that  $v$  is Lipschitz in the variable  $x$  (see Proposition 2.3) and observe that, by the same arguments,  $V$  is also Lipschitz in its second argument. Furthermore, both  $v$  and  $V$  are bounded. Notice also that given  $\xi \in \mathbf{S}_2(\mathcal{F}_t)$  there exists a sequence  $\{\xi_k\}_k \subset \mathbf{S}_2(\mathcal{F}_t)$  converging to  $\xi$ , with  $\xi_k$  taking only a finite number of values. As a consequence, from the continuity of  $v$  and  $V$ , it is enough to prove (2.12) with  $\xi \in \mathbf{S}_2(\mathcal{F}_t)$  taking only a finite number of values. Then, from now on, let us suppose that

$$\xi = \sum_{i=1}^n x_i 1_{E_i}, \quad (2.13)$$

for some  $n \in \mathbb{N}$ ,  $x_i \in C([0, T]; \mathbb{R}^d)$ ,  $E_i \in \mathcal{F}_t$ , with  $\{E_i\}_{i=1, \dots, n}$  being a partition of  $\Omega$ .

STEP III. *Proof of the inequality  $V(t, \xi) \leq \mathbb{E}[v(t, \xi)]$ .* Since  $\xi \in \mathbf{S}_2(\mathbb{F})$  takes only a finite number of values, by [10, Lemma B.3] (here we use that  $t > 0$ , so in particular  $\mathcal{F}_t$  has the property required by [10, Lemma B.3], namely there exists a  $\mathcal{F}_t$ -measurable random variable having uniform distribution on  $[0, 1]$ ) there exists a  $\mathcal{F}_t$ -measurable random variable  $U: \Omega \rightarrow \mathbb{R}$ , having uniform distribution on  $[0, 1]$  and being independent of  $\xi$ . As a consequence, from [10, Lemma B.2] it follows that, for every  $\alpha \in \mathcal{A}$ , there exists a measurable function

$$a: ([0, T] \times \Omega \times C([0, T]; \mathbb{R}^d) \times [0, 1], \text{Prog}(\mathbb{F}^t) \otimes \mathcal{B}(C([0, T]; \mathbb{R}^d)) \otimes \mathcal{B}([0, 1])) \longrightarrow (A, \mathcal{B}(A))$$

such that

$$\beta_s := \alpha_s 1_{[0, t)}(s) + a_s(\xi, U) 1_{[t, T]}(s), \quad \forall s \in [0, T]$$

belongs to  $\mathcal{A}$  and

$$\left( \xi, (a_s(\xi, U))_{s \in [t, T]}, (B_s - B_t)_{s \in [t, T]} \right) \stackrel{\mathcal{L}}{=} \left( \xi, (\alpha_s)_{s \in [t, T]}, (B_s - B_t)_{s \in [t, T]} \right),$$

where  $\stackrel{\mathcal{L}}{=}$  means equality in law. Then, by the same arguments as in [29, Proposition 1.137], we get

$$(X_s^{t, \xi, \alpha}, \alpha_s)_{s \in [0, T]} \stackrel{\mathcal{L}}{=} (X_s^{t, \xi, \beta}, \beta_s)_{s \in [0, T]}.$$

Moreover, recalling (2.13), define

$$\beta_{i, s} := \alpha_s 1_{[0, t)}(s) + a_s(x_i, U) 1_{[t, T]}(s), \quad \forall s \in [0, T], i = 1, \dots, n.$$

Since  $X^{t,\xi,\beta}$  and  $X^{t,x_1,\beta_1} 1_{E_1} + \dots + X^{t,x_n,\beta_n} 1_{E_n}$  solve the same equation, they are  $\mathbb{P}$ -indistinguishable. Hence

$$\begin{aligned} \mathbb{E} \left[ \int_t^T f(s, X^{t,\xi,\alpha}, \alpha_s) ds + g(X^{t,\xi,\alpha}) \right] &= \mathbb{E} \left[ \int_t^T f(s, X^{t,\xi,\beta}, \beta_s) ds + g(X^{t,\xi,\beta}) \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^n \left( \int_t^T f(s, X^{t,x_i,\beta_i}, \beta_{i,s}) ds + g(X^{t,x_i,\beta_i}) \right) 1_{E_i} \right]. \end{aligned}$$

Recalling that both  $\{X^{t,x_i,\beta_i}\}_i$  and  $\{\beta_i\}_i$  are independent of  $\{E_i\}_i$ , we have

$$\begin{aligned} &\mathbb{E} \left[ \sum_{i=1}^n \left( \int_t^T f(s, X^{t,x_i,\beta_i}, \beta_{i,s}) ds + g(X^{t,x_i,\beta_i}) \right) 1_{E_i} \right] \\ &= \mathbb{E} \left[ \sum_{i=1}^n \mathbb{E} \left[ \int_t^T f(s, X^{t,x_i,\beta_i}, \beta_{i,s}) ds + g(X^{t,x_i,\beta_i}) \right] 1_{E_i} \right] \\ &= \sum_{i=1}^n \mathbb{E} \left[ \mathbb{E} \left[ \int_t^T f(s, X^{t,x_i,\beta_i}, \beta_{i,s}) ds + g(X^{t,x_i,\beta_i}) \right] 1_{E_i} \right] \leq \sum_{i=1}^n \mathbb{E} [v(t, x_i) 1_{E_i}] = \mathbb{E}[v(t, \xi)]. \end{aligned}$$

Then, the inequality  $V(t, \xi) \leq \mathbb{E}[v(t, \xi)]$  follows from the arbitrariness of  $\alpha$ .

STEP IV. *Proof of the inequality  $V(t, \xi) \geq \mathbb{E}[v(t, \xi)]$ .* Take  $\xi \in \mathbf{S}_2(\mathcal{F}_t)$  as in (2.13). Then, from equality (2.7) of Lemma 2.4, for every  $\varepsilon > 0$  and  $i = 1, \dots, n$ , there exists  $\beta_i^\varepsilon \in \mathcal{A}_t$  such that

$$v(t, x_i) \leq \mathbb{E} \left[ \int_t^T f(s, X^{t,x_i,\beta_i^\varepsilon}, \beta_{i,s}^\varepsilon) ds + g(X^{t,x_i,\beta_i^\varepsilon}) \right] + \varepsilon.$$

Let

$$\beta^\varepsilon := \sum_{i=1}^n \beta_i 1_{E_i}.$$

We have  $\beta^\varepsilon \in \mathcal{A}$ , moreover  $X^{t,\xi,\beta^\varepsilon}$  and  $X^{t,x_1,\beta_1^\varepsilon} 1_{E_1} + \dots + X^{t,x_n,\beta_n^\varepsilon} 1_{E_n}$  solve the same equation, therefore they are  $\mathbb{P}$ -indistinguishable. Therefore (exploiting the independence of both  $\{X^{t,x_i,\beta_i^\varepsilon}\}_i$  and  $\{\beta_i^\varepsilon\}_i$  from  $\{E_i\}_i$ )

$$\begin{aligned} \mathbb{E}[v(t, \xi)] &= \sum_{i=1}^n \mathbb{E} [v(t, x_i) 1_{E_i}] \\ &\leq \sum_{i=1}^n \mathbb{E} \left[ \mathbb{E} \left[ \int_t^T f(s, X^{t,x_i,\beta_i^\varepsilon}, \beta_{i,s}^\varepsilon) ds + g(X^{t,x_i,\beta_i^\varepsilon}) \right] 1_{E_i} \right] + \varepsilon \\ &= \mathbb{E} \left[ \sum_{i=1}^n \mathbb{E} \left[ \int_t^T f(s, X^{t,x_i,\beta_i^\varepsilon}, \beta_{i,s}^\varepsilon) ds + g(X^{t,x_i,\beta_i^\varepsilon}) \right] 1_{E_i} \right] + \varepsilon \\ &= \mathbb{E} \left[ \sum_{i=1}^n \left( \int_t^T f(s, X^{t,x_i,\beta_i^\varepsilon}, \beta_{i,s}^\varepsilon) ds + g(X^{t,x_i,\beta_i^\varepsilon}) \right) 1_{E_i} \right] + \varepsilon \\ &= \mathbb{E} \left[ \int_t^T f(s, X^{t,\xi,\beta^\varepsilon}, \beta_s^\varepsilon) ds + g(X^{t,\xi,\beta^\varepsilon}) \right] + \varepsilon \leq V(t, \xi) + \varepsilon. \end{aligned}$$

From the arbitrariness of  $\varepsilon$ , the inequality  $\mathbb{E}[v(t, \xi)] \leq V(t, \xi)$  follows.  $\square$

**Theorem 2.6.** *Suppose that Assumption (A) holds. Then the value function  $v$  satisfies the **dynamic programming principle**: for every  $t, s \in [0, T]$ , with  $t \leq s$ , and every  $x \in C([0, T]; \mathbb{R}^d)$  it holds that*

$$v(t, x) = \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \int_t^s f(r, X^{t,x,\alpha}, \alpha_r) dr + v(s, X^{t,x,\alpha}) \right].$$

**Proof.** This follows directly from [10, Theorem 3.4] and Lemma 2.5. As a matter of fact, let  $V$  be the function given by (2.11). From [10, Theorem 3.4] we get the dynamic programming principle for  $V$ :

$$V(t, x) = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[ \int_t^s f(r, X^{t,x,\alpha}, \alpha_r) dr \right] + V(s, X^{t,x,\alpha}) \right\}.$$

Moreover, by Lemma 2.5 we know that

$$V(s, X^{t,x,\alpha}) = \mathbb{E}[v(s, X^{t,x,\alpha})],$$

from which the claim follows.  $\square$

## 3 Path Dependent HJB equations and viscosity solutions

### 3.1 Definition of path-dependent viscosity solutions

In the present paper we adopt the standard definitions of pathwise (or functional) derivatives of a map  $u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$ , as they were introduced in the seminal paper [21], and further developed by [6, 7] and [11, Section 2]. We report in Appendix A a concise presentation of these tools. Just to fix notations, we recall here that the pathwise derivatives of a map  $u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  are given by the horizontal derivative  $\partial_t^H u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  and the vertical derivatives of first and second-order  $\partial_x^V u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}^d$  and  $\partial_{xx}^V u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}^{d \times d}$ . We also refer to Definition A.4 (resp. Definition A.6) for the definition of the class  $C^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  (resp.  $C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$ ). Finally, we recall that for a map  $u \in C^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  the so-called functional Itô's formula holds, see Theorem A.7.

Now, consider the following second-order path-dependent partial differential equation:

$$\begin{cases} \partial_t^H u(t, x) = F(t, x, u(t, x), \partial_x^V u(t, x), \partial_{xx}^V u(t, x)), & (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d), \\ u(T, x) = g(x), & x \in C([0, T]; \mathbb{R}^d), \end{cases} \quad (3.1)$$

with  $F: [0, T] \times C([0, T]; \mathbb{R}^d) \times \mathbb{R} \times \mathbb{R}^d \times \mathcal{S}(d) \rightarrow \mathbb{R}$ , where  $\mathcal{S}(d)$  is the set of symmetric  $d \times d$  matrices.

**Definition 3.1.** We say that a function  $u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  is a **classical solution** of equation (3.1) if it belongs to  $C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  and satisfies (3.1).

**Definition 3.2.** We say that an upper semicontinuous function  $u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  is a **(path-dependent) viscosity subsolution** of equation (3.1) if:

- $u(T, x) \leq g(x)$ , for all  $x \in C([0, T]; \mathbb{R}^d)$ ;
- for any  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$  and  $\varphi \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$ , satisfying

$$(u - \varphi)(t, x) = \sup_{(t', x') \in [0, T] \times C([0, T]; \mathbb{R}^d)} (u - \varphi)(t', x'),$$

with  $(u - \varphi)(t, x) = 0$ , we have

$$- \partial_t^H \varphi(t, x) + F(t, x, u(t, x), \partial_x^V \varphi(t, x), \partial_{xx}^V \varphi(t, x)) \leq 0. \quad (3.2)$$

We say that a lower semicontinuous function  $u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  is a **(path-dependent) viscosity supersolution** of equation (3.1) if:

- $u(T, x) \geq g(x)$ , for all  $x \in C([0, T]; \mathbb{R}^d)$ ;
- for any  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$  and  $\varphi \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$ , satisfying

$$(u - \varphi)(t, x) = \inf_{(t', x') \in [0, T] \times C([0, T]; \mathbb{R}^d)} (u - \varphi)(t', x'),$$

with  $(u - \varphi)(t, x) = 0$ , we have

$$- \partial_t^H \varphi(t, x) + F(t, x, u(t, x), \partial_x^V \varphi(t, x), \partial_{xx}^V \varphi(t, x)) \geq 0. \quad (3.3)$$

We say that a continuous map  $u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  is a **(path-dependent) viscosity solution** of equation (3.1) if  $u$  is both a (path-dependent) viscosity subsolution and a (path-dependent) viscosity supersolution of (3.1).

## 3.2 The value function solves the path-dependent HJB equation

Now, we focus on the path-dependent Hamilton-Jacobi-Bellman equation, namely on equation (3.1) with

$$F(t, x, r, p, M) = - \sup_{a \in A} \left\{ \langle b(t, x, a), p \rangle + \frac{1}{2} \text{tr}[(\sigma \sigma^\top)(t, x, a)M] + f(t, x, a) \right\}. \quad (3.4)$$

Therefore, equation (3.1) becomes

$$\begin{cases} \partial_t^H u(t, x) + \sup_{a \in A} \left\{ \langle b(t, x, a), \partial_x^V u(t, x) \rangle \right. \\ \left. + \frac{1}{2} \text{tr}[(\sigma \sigma^\top)(t, x, a) \partial_{xx}^V u(t, x)] + f(t, x, a) \right\} = 0, & (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d), \\ u(T, x) = g(x), & x \in C([0, T]; \mathbb{R}^d). \end{cases} \quad (3.5)$$

We now prove that the value function  $v$  is a viscosity solution to equation (3.5).

**Theorem 3.3.** *Suppose that Assumptions (A) and (B) hold. The value function  $v$ , defined by (2.5), is a viscosity solution to equation (3.5).*

**Proof.** Recall from Proposition 2.3 that  $v$  is continuous, moreover  $v(T, \cdot) \equiv g(\cdot)$ . Then it remains to prove both the subsolution and the supersolution property on  $[0, T] \times C([0, T]; \mathbb{R}^d)$ .

*Subsolution property.* Let  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$  and  $\varphi \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  be such that

$$(v - \varphi)(t, x) = \sup_{(t', x') \in [0, T] \times C([0, T]; \mathbb{R}^d)} (v - \varphi)(t', x') = 0.$$

From Theorem 2.6 we know that, for every  $h > 0$  sufficiently small,

$$0 = \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[ \frac{1}{h} \int_t^{t+h} f(r, X^{t,x,\alpha}, \alpha_r) dr + \frac{1}{h} (v(t+h, X^{t,x,\alpha}) - v(t, x)) \right] \right\}.$$

Then, there exists  $\alpha^h \in \mathcal{A}$  such that

$$\begin{aligned} -h &\leq \mathbb{E} \left[ \frac{1}{h} \int_t^{t+h} f(r, X^{t,x,\alpha^h}, \alpha_r^h) dr + \frac{1}{h} (v(t+h, X^{t,x,\alpha^h}) - v(t, x)) \right] \\ &\leq \mathbb{E} \left[ \frac{1}{h} \int_t^{t+h} f(r, X^{t,x,\alpha^h}, \alpha_r^h) dr + \frac{1}{h} (\varphi(t+h, X^{t,x,\alpha^h}) - \varphi(t, x)) \right], \end{aligned}$$

where the above inequality follows from the fact that  $v(t, x) = \varphi(t, x)$  and  $v \leq \varphi$ . By the functional Itô formula (A.1), we obtain

$$\begin{aligned} 0 &\leq h + \frac{1}{h} \int_t^{t+h} \mathbb{E} [\partial_t^H \varphi(r, X^{t,x,\alpha^h})] dr + \frac{1}{h} \int_t^{t+h} \mathbb{E} [\langle b(r, X^{t,x,\alpha^h}, \alpha_r^h), \partial_x^V \varphi(r, X^{t,x,\alpha^h}) \rangle] dr \\ &\quad + \frac{1}{h} \int_t^{t+h} \frac{1}{2} \mathbb{E} [\text{tr}[(\sigma \sigma^\top)(r, X^{t,x,\alpha^h}, \alpha_r^h) \partial_{xx}^V \varphi(r, X^{t,x,\alpha^h})]] dr + \frac{1}{h} \int_t^{t+h} \mathbb{E} [f(r, X^{t,x,\alpha^h})] dr. \end{aligned}$$

Recalling that  $b, \sigma, f$  are uniformly continuous in  $(t, x)$ , uniformly with respect to  $a$ , using (2.4), we get

$$\begin{aligned} &\frac{1}{h} \int_t^{t+h} \mathbb{E} [\partial_t^H \varphi(r, X^{t,x,\alpha^h})] dr + \frac{1}{h} \int_t^{t+h} \mathbb{E} [\langle b(r, X^{t,x,\alpha^h}, \alpha_r^h), \partial_x^V \varphi(r, X^{t,x,\alpha^h}) \rangle] \\ &\quad + \frac{1}{2} \text{tr}[(\sigma \sigma^\top)(r, X^{t,x,\alpha^h}, \alpha_r^h) \partial_{xx}^V \varphi(r, X^{t,x,\alpha^h})] + f(r, X^{t,x,\alpha^h}, \alpha_r^h) \Big] dr \\ &= \partial_t^H \varphi(t, x) + \frac{1}{h} \int_t^{t+h} \mathbb{E} [\langle b(t, x, \alpha_r^h), \partial_x^V \varphi(t, x) \rangle \\ &\quad + \frac{1}{2} \text{tr}[(\sigma \sigma^\top)(t, x, \alpha_r^h) \partial_{xx}^V \varphi(t, x)] + f(t, x, \alpha_r^h) \Big] dr + \rho(h), \end{aligned}$$

where  $\rho(h) \rightarrow 0$  as  $h \rightarrow 0^+$ . Then, we obtain

$$\begin{aligned} 0 &\leq h + \rho(h) + \partial_t^H \varphi(t, x) + \frac{1}{h} \int_t^{t+h} \sup_{a \in \mathcal{A}} \left\{ \langle b(t, x, a), \partial_x^V \varphi(t, x) \rangle \right. \\ &\quad \left. + \frac{1}{2} \text{tr}[(\sigma \sigma^\top)(t, x, a) \partial_{xx}^V \varphi(t, x)] + f(t, x, a) \right\} dr. \end{aligned}$$

Sending  $h \rightarrow 0^+$ , we conclude that (3.2) holds (with  $F$  given by (3.4)).

*Supersolution property.* Let  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$  and  $\varphi \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  be such that

$$(v - \varphi)(t, x) = \inf_{(t', x') \in [0, T] \times C([0, T]; \mathbb{R}^d)} (v - \varphi)(t', x') = 0.$$

From Theorem 2.6 we have, for every  $h > 0$  sufficiently small, and for every constant control strategy  $\alpha \equiv a \in A$ ,

$$\begin{aligned} 0 &\geq \mathbb{E} \left[ \frac{1}{h} \int_t^{t+h} f(r, X^{t,x,a}, a) dr + \frac{1}{h} (v(t+h, X^{t,x,a}) - v(t, x)) \right] \\ &\geq \mathbb{E} \left[ \frac{1}{h} \int_t^{t+h} f(r, X^{t,x,a}, a) dr + \frac{1}{h} (\varphi(t+h, X^{t,x,a}) - \varphi(t, x)) \right], \end{aligned}$$

where the above inequality follows from the fact that  $v(t, x) = \varphi(t, x)$  and  $v \geq \varphi$ . Now, by the functional Itô formula (A.1), we obtain

$$\begin{aligned} 0 &\geq \frac{1}{h} \int_t^{t+h} \mathbb{E} [\partial_t^H \varphi(r, X^{t,x,a})] dr + \frac{1}{h} \int_t^{t+h} \mathbb{E} [\langle b(r, X^{t,x,a}, a), \partial_x^V \varphi(r, X^{t,x,a}) \rangle] dr \\ &\quad + \frac{1}{h} \int_t^{t+h} \frac{1}{2} \mathbb{E} [\text{tr} [(\sigma \sigma^\top)(r, X^{t,x,a}, a) \partial_{xx}^V \varphi(r, X^{t,x,a})]] dr + \frac{1}{h} \int_t^{t+h} \mathbb{E} [f(r, X^{t,x,a}, a)] dr. \end{aligned}$$

Letting  $h \rightarrow 0^+$ , exploiting the regularity of  $\varphi$  and the continuity of  $b, \sigma, f$ , we find

$$0 \geq \partial_t^H \varphi(t, x) + \langle b(t, x, a), \partial_x^V \varphi(t, x) \rangle + \frac{1}{2} \text{tr} [(\sigma \sigma^\top)(t, x, a) \partial_{xx}^V \varphi(t, x)] + f(t, x, a).$$

From the arbitrariness of  $a$ , we conclude that (3.3) holds (with  $F$  given by (3.4)).  $\square$

## 4 Uniqueness

### 4.1 Smooth variational principle

This section is devoted to state a smooth variational principle on  $[0, T] \times C([0, T]; \mathbb{R}^d)$ , which will be an essential tool in the proof of the comparison theorem (Theorem 4.5). All the results stated in the present section are proved in [11, Section 3], to which we refer for more details. Here we notice that such a smooth variational principle is obtained from [5, Theorem 2.5.2], which is a generalization of the Borwein-Preiss variant ([4]) of Ekeland's variational principle ([22]). More precisely, [5, Theorem 2.5.2] extends Ekeland's principle to the concept of gauge-type function, that we now introduce.

**Definition 4.1.** *A map  $\Psi: ([0, T] \times C([0, T]; \mathbb{R}^d))^2 \rightarrow [0, +\infty)$  is called a ***gauge-type function*** if it satisfies the following properties.*

a)  $\Psi$  is continuous on  $([0, T] \times C([0, T]; \mathbb{R}^d))^2$ .



b)  $\Psi((t, x), (t, x)) = 0$ , for every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ .

c) For every  $\varepsilon > 0$  there exists  $\eta > 0$  such that, for all  $(t', x'), (t'', x'') \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , the inequality  $\Psi((t', x'), (t'', x'')) \leq \eta$  implies  $d_\infty((t', x'), (t'', x'')) < \varepsilon$ .

In the proof of the comparison theorem we need a gauge-type function  $\Psi$  such that  $(t, x) \mapsto \Psi((t, x), (t_0, x_0))$  is smooth, for every fixed  $(t_0, x_0)$ . Notice that  $d_\infty$  is obviously a gauge-type function, however it is not smooth enough. The following result provide a smooth gauge-type function and it is due to [11, Section 3] (see Remark 4.3 for some insights on the expression of  $\rho_\infty$  in (4.1)).

**Lemma 4.2.** Define the map  $\rho_\infty: ([0, T] \times C([0, T]; \mathbb{R}^d))^2 \rightarrow \mathbb{R}$  as

$$\begin{aligned} \rho_\infty((t, x), (t_0, x_0)) &= |t - t_0|^2 \\ &+ \int_0^{+\infty} \frac{\kappa_\infty(((t+s) \wedge T, x(\cdot \wedge t)), (t_0, x_0))}{1 + C_\zeta + \kappa_\infty(((t+s) \wedge T, x(\cdot \wedge t)), (t_0, x_0))} \eta(s) ds, \end{aligned} \quad (4.1)$$

for all  $(t, x), (t_0, x_0) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , where

$$\begin{aligned} \kappa_\infty((t, x), (t_0, x_0)) &= \int_{\mathbb{R}^d} \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0) - z \mathbf{1}_{[t, T]}\|_T \zeta(z) dz - \int_{\mathbb{R}^d} |z| \zeta(z) dz, \\ \zeta(z) &= \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}|z|^2}, \\ \eta(s) &= s e^{-s}, \\ C_\zeta &= \int_{\mathbb{R}^d} |z| \zeta(z) dz = \sqrt{2} \frac{\Gamma((d+1)/2)}{\Gamma(d/2)}, \end{aligned} \quad (4.2)$$

for all  $(t, x), (t_0, x_0) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ ,  $z \in \mathbb{R}^d$ ,  $s \in \mathbb{R}$ , with  $\Gamma(\cdot)$  being the Gamma function. Then, the following holds.

1)  $\rho_\infty$  is a gauge-type function.

2) There exists a constant  $\alpha_d > 0$  (depending only on the dimension  $d$ ) such that

$$\begin{aligned} |t - t_0|^2 + \alpha_d \frac{\|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T^{d+1} \wedge \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T}{1 + C_\zeta + \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T} \\ \leq \rho_\infty((t, x), (t_0, x_0)) \leq |t - t_0|^2 + \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T \wedge 1. \end{aligned} \quad (4.3)$$

3) For every fixed  $(t_0, x_0) \in C([0, T]; \mathbb{R}^d)$ , the map  $(t, x) \mapsto \rho_\infty((t, x), (t_0, x_0))$  satisfies the following properties.

- a) the map  $(t, x) \mapsto \rho_\infty((t, x), (t_0, x_0))$  belongs to  $C^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  and is bounded;
- b) its horizontal derivative is bounded by  $2T + \frac{2}{e}$ ;
- c) its vertical derivatives of first-order are bounded by  $1 + C_\zeta$ ;

d) its vertical derivatives of second-order are bounded by  $(1 + C_\zeta)(\sqrt{2/\pi} + 2)$ .

**Proof.** The claim follows from Lemma 3.1 and Lemma 3.2 in [11].  $\square$

**Remark 4.3.** Lemma 4.2 is based on the idea of performing a smoothing of the pseudometric  $((t, x), (t_0, x_0)) \mapsto |t - t_0| + \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T$  in order to construct a smooth gauge-type function. In particular, the objective is to get a regular map of  $(t, x)$  starting from  $(t, x) \mapsto |t - t_0| + \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T$ , for every fixed  $(t_0, x_0)$ . In order to do it, the main issue is to perform a smoothing of the map  $(t, x) \mapsto \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T$ .

Firstly, notice that  $(t, x) \mapsto \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T$  is already smooth in the horizontal direction (its horizontal derivative is identically equal to zero), while it is not smooth in the vertical direction. The smoothing of  $(t, x) \mapsto \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T$  is then achieved as in (4.2), namely through a convolution along the vertical direction (identified by  $z1_{[t, T]}$ ). Notice indeed that  $k_\infty$  is infinitely continuously differentiable in the vertical direction. Moreover, thanks to the presence of the term  $-\int_{\mathbb{R}^d} |z|\zeta(z)dz$ , we have  $k_\infty((t, x), (t, x)) = 0$ , which corresponds to property b) of Definition 4.1 of gauge-type function. We also remark that the particular choice of  $\zeta$  (and, similarly, of  $\eta$ ) is only useful in order to get precise constants in item 3) of Lemma 4.2, but clearly any other mollifier can be used.

Finally, we observe that unfortunately the map  $(t, x) \mapsto k_\infty((t, x), (t_0, x_0))$  is no more smooth in the horizontal direction (despite  $(t, x) \mapsto \|x(\cdot \wedge t) - x_0(\cdot \wedge t_0)\|_T$  was horizontally smooth), because of the presence of the term  $z1_{[t, T]}$ , which depends on the time variable  $t$ . To recover the horizontal regularity, we then perform a smoothing with respect to the time variable through a convolution with the mollifier  $\eta$  as in (4.1). The presence of  $1 + C_\zeta + \kappa_\infty$  in the denominator of (4.1) is only to get a map  $\rho_\infty$  which is bounded with bounded derivatives. For another example of smooth gauge-type function on  $[0, T] \times C([0, T]; \mathbb{R}^d)$ , see [47, Lemma 2.4].

We can finally state the smooth variational principle on  $[0, T] \times C([0, T]; \mathbb{R}^d)$ .

**Theorem 4.4.** Let  $\delta > 0$  and  $G: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  be an upper semicontinuous map, bounded from above. For every  $\varepsilon > 0$ , let  $(t_0, x_0) \in [0, T] \times C([0, T]; \mathbb{R}^d)$  satisfy

$$\sup G - \varepsilon \leq G(t_0, x_0).$$

There exist  $(\bar{t}, \bar{x}) \in [0, T] \times C([0, T]; \mathbb{R}^d)$  and  $\varphi: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow [0, +\infty)$  fulfilling the following properties.

- i)  $\rho_\infty((\bar{t}, \bar{x}), (t_0, x_0)) \leq \frac{\varepsilon}{\delta}$ .
- ii)  $G(t_0, x_0) \leq G(\bar{t}, \bar{x}) - \delta\varphi(\bar{t}, \bar{x})$ .
- iii) For every  $(t, x) \neq (\bar{t}, \bar{x})$ ,  $G(t, x) - \delta\varphi(t, x) < G(\bar{t}, \bar{x}) - \delta\varphi(\bar{t}, \bar{x})$ .

In addition, the following holds.

- 1)  $\varphi \in C^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  and is bounded;

- 2) its horizontal derivative is bounded by  $2(2T + \frac{2}{\epsilon})$ ;
- 3) its vertical derivatives of first-order are bounded by  $2(1 + C_\zeta)$ ;
- 4) its vertical derivatives of second-order are bounded by  $2(1 + C_\zeta)(\sqrt{2/\pi} + 2)$ .

**Proof.** This results follows directly from [11, Theorem 3.2]. Here we simply notice that items i)-ii)-iii) follow from the variational principle [5, Theorem 2.5.2], which applies to a generic gauge-type function  $\Psi$ , not only to  $\rho_\infty$  (just observe that [5, Theorem 2.5.2] is formulated on a complete *metric* space, while  $[0, T] \times C([0, T]; \mathbb{R}^d)$  is a complete *pseudometric* space; however, this does not affect the result). On the other hand, items 1)-2)-3)-4) follow from the regularity properties of  $\rho_\infty$  stated in Lemma 4.2 and from the definition of  $\varphi$ , which we recall from the statement of [11, Theorem 3.2] that is given by

$$\varphi(t, x) := \sum_{k=0}^{+\infty} \frac{1}{2^k} \rho_\infty((t, x), (t_k, x_k)), \quad \forall (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d),$$

for some sequence  $\{(t_k, x_k)\} \subset [0, T] \times C([0, T]; \mathbb{R}^d)$  converging to  $(\bar{t}, \bar{x})$  and satisfying  $\rho_\infty((t_k, x_k), (\bar{t}, \bar{x})) \leq \varepsilon/(2^k \delta)$ , for every  $k \geq 1$ .  $\square$

## 4.2 Comparison theorem and uniqueness

**Theorem 4.5.** *Suppose that Assumptions (A) and (B) hold. Let  $u_1, u_2: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  be respectively upper and lower semicontinuous bounded functions. Suppose that  $u_1$  (resp.  $u_2$ ) is a (path-dependent) viscosity subsolution (resp. supersolution) of equation (3.5). Then  $u_1 \leq u_2$  on  $[0, T] \times C([0, T]; \mathbb{R}^d)$ .*

**Proof.** The proof consists in showing that  $u_1 \leq v$  and  $v \leq u_2$  on  $[0, T] \times C([0, T]; \mathbb{R}^d)$ , with  $v$  given by (2.5). We only prove the inequality  $u_1 \leq v$ , as  $v \leq u_2$  can be proved proceeding along the same lines. We split the proof of  $u_1 \leq v$  into five steps.

STEP I. We proceed by contradiction and assume that  $\sup(u_1 - v) > 0$ . Then, there exists  $(t_0, x_0) \in [0, T] \times C([0, T]; \mathbb{R}^d)$  such that

$$(u_1 - v)(t_0, x_0) > 0. \quad (4.4)$$

Notice that  $t_0 < T$ , since  $u_1(T, x_0) \leq g(x_0) = v(T, x_0)$ . Now, consider the sequences  $\{b_n\}_n$ ,  $\{\sigma_n\}_n$ ,  $\{f_n\}_n$ ,  $\{g_n\}_n$  in (B.16). Moreover, for every  $n$  and any  $\varepsilon \in (0, 1)$ , consider the functions  $v_{n,\varepsilon} \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  and  $\bar{v}_{n,\varepsilon} \in C^{1,2}([0, T] \times \mathbb{R}^d)$  introduced in Theorem B.6, with  $v_{n,\varepsilon}$  classical solution of the following equation:

$$\begin{cases} \partial_t^H v_{n,\varepsilon}(t, x) + \frac{1}{2} \varepsilon^2 \text{tr}[\partial_{yy} \bar{v}_{n,\varepsilon}(t, y_n^{t,x})] + \sup_{a \in A} \left\{ \langle b_n(t, x, a), \partial_x^V v_{n,\varepsilon}(t, x) \rangle \right. \\ \left. + \frac{1}{2} \text{tr}[(\sigma_n \sigma_n^T)(t, x, a) \partial_{xx}^V v_{n,\varepsilon}(t, x)] + f_n(t, x, a) \right\} = 0, & (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d), \\ v_{n,\varepsilon}(T, x) = g_n(x), & x \in C([0, T]; \mathbb{R}^d), \end{cases}$$

where

$$y_n^{t,x} := \left( \int_{[0,t]} \phi_{n,1}(s) d^-x(s), \dots, \int_{[0,t]} \phi_{n,d_n}(s) d^-x(s) \right). \quad (4.5)$$

STEP II. Given  $\delta > 0$ , we set  $\tilde{u}_1(t, x) := e^{\sqrt{\delta}(t+1)} u_1(t, x)$ , for all  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , and we define similarly  $\tilde{v}_{n,\varepsilon}, \tilde{f}, \tilde{f}_n$ . We also define  $\tilde{g}(x) := e^{\sqrt{\delta}(T+1)} g(x)$  and  $\tilde{g}_n(x) := e^{\sqrt{\delta}(T+1)} g_n(x)$ , for all  $x \in C([0, T]; \mathbb{R}^d)$ . Notice that  $\tilde{u}_1$  is a (path-dependent) viscosity sub-solution of the following path-dependent partial differential equation:

$$\begin{cases} \partial_t^H \tilde{u}_1(t, x) + \sup_{a \in A} \left\{ \tilde{f}(t, x, a) + \langle b(t, x, a), \partial_x^V \tilde{u}_1(t, x) \rangle \right. \\ \left. + \frac{1}{2} \text{tr}[(\sigma \sigma^\top)(t, x, a) \partial_{xx}^V \tilde{u}_1(t, x)] \right\} = \sqrt{\delta} \tilde{u}_1(t, x), & (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d), \\ \tilde{u}_1(T, x) = \tilde{g}(x), & x \in C([0, T]; \mathbb{R}^d). \end{cases} \quad (4.6)$$

Similarly,  $\tilde{v}_{n,\varepsilon} \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  and is a classical solution of the following equation:

$$\begin{cases} \partial_t^H \tilde{v}_{n,\varepsilon}(t, x) + \frac{1}{2} \varepsilon^2 e^{\sqrt{\delta}(t+1)} \text{tr}[\partial_{yy} \bar{v}_{n,\varepsilon}(t, y_n^{t,x})] + \sup_{a \in A} \left\{ \tilde{f}_n(t, x, a) + \langle b_n(t, x, a), \partial_x^V \tilde{v}_{n,\varepsilon}(t, x) \rangle \right. \\ \left. + \frac{1}{2} \text{tr}[(\sigma_n \sigma_n^\top)(t, x, a) \partial_{xx}^V \tilde{v}_{n,\varepsilon}(t, x)] \right\} = \sqrt{\delta} \tilde{v}_{n,\varepsilon}(t, x), & (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d), \\ \tilde{v}_{n,\varepsilon}(T, x) = \tilde{g}_n(x), & x \in C([0, T]; \mathbb{R}^d). \end{cases} \quad (4.7)$$

Now, recall from the definition of  $f_n$  and  $g_n$  in Theorem B.6 that  $f_n$  and  $g_n$  are bounded by the same constant  $K$  appearing in Assumption (A). As a consequence,  $v_{n,\varepsilon}$  is bounded by a constant independent of  $n, \varepsilon$ . So, in particular, there exists a constant  $\lambda \geq 0$ , independent of  $n, \varepsilon$ , such that

$$\sup(\tilde{u}_1 - \tilde{v}_{n,\varepsilon}) \leq (\tilde{u}_1 - \tilde{v}_{n,\varepsilon})(t_0, x_0) + \lambda. \quad (4.8)$$

STEP III. Notice that  $\tilde{u}_1 - \tilde{v}_{n,\varepsilon}$  is upper semicontinuous and bounded. Then, by (4.8) and the smooth variational principle (Theorem 4.4) with  $G = \tilde{u}_1 - \tilde{v}_{n,\varepsilon}$ , we deduce that for every  $\delta > 0$  there exist  $(\bar{t}, \bar{x}) \in [0, T] \times C([0, T]; \mathbb{R}^d)$  and  $\varphi: [0, T] \times C([0, T] \times \mathbb{R}^d) \rightarrow \mathbb{R}$  (with  $\bar{t}, \bar{x}, \varphi$  possibly depending on  $n, \varepsilon, \delta$ ), such that the following holds.

i)  $\rho_\infty((\bar{t}, \bar{x}), (t_0, x_0)) \leq \frac{\lambda}{\delta}$ , where  $\rho_\infty$  is the smooth gauge-type function defined by (4.1).

ii)  $(\tilde{u}_1 - \tilde{v}_{n,\varepsilon})(t_0, x_0) \leq (\tilde{u}_1 - (\tilde{v}_{n,\varepsilon} + \delta\varphi))(\bar{t}, \bar{x})$ .

iii) It holds that

$$(\tilde{u}_1 - (\tilde{v}_{n,\varepsilon} + \delta\varphi))(\bar{t}, \bar{x}) = \sup_{(t,x) \in [0,T] \times C([0,T]; \mathbb{R}^d)} (\tilde{u}_1 - (\tilde{v}_{n,\varepsilon} + \delta\varphi))(t, x). \quad (4.9)$$

We also recall from Theorem 4.4 that  $\varphi \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$ , it is bounded and it has bounded derivatives.

By item i), we get

$$\sqrt{\delta} |\bar{t} - t_0| \leq \lambda, \quad \forall \delta > 0, \quad (4.10)$$

Moreover, by (4.3), we obtain

$$\|\bar{x}(\cdot \wedge \bar{t}) - x_0(\cdot \wedge t_0)\|_T^{d+1} \leq \frac{\lambda}{\delta} \frac{3 + C_\zeta}{\alpha_d}, \quad \forall \delta \geq \lambda \frac{3 + C_\zeta}{\alpha_d}, \quad (4.11)$$

with  $\alpha_d$  and  $C_\zeta$  being the same constants appearing in Lemma 4.2. As a matter of fact, let  $h_1, h_2: [0, \infty) \rightarrow [0, \infty)$  be given by

$$h_1(r) = \alpha_d \frac{r^{d+1} \wedge r}{1 + C_\zeta + r}, \quad h_2(r) = \alpha_d \frac{r^{d+1}}{1 + C_\zeta + r} \quad r \geq 0.$$

Observe that  $h_1(r) = h_2(r)$ , whenever  $r \leq 1$ . Now, by (4.3) and item i), we have

$$h_1(\|\bar{x}(\cdot \wedge \bar{t}) - x_0(\cdot \wedge t_0)\|_T) \leq \rho_\infty((t_0, x_0), (\bar{t}, \bar{x})) \leq \frac{\lambda}{\delta}.$$

Notice that  $h_1$  is non-decreasing and  $h_1(1) = \alpha_d/(2 + C_\zeta)$ . If  $\delta \geq \lambda \frac{3 + C_\zeta}{\alpha_d}$ , or equivalently  $\lambda/\delta \leq \alpha_d/(3 + C_\zeta) < h_1(1)$ , then  $\|\bar{x}(\cdot \wedge \bar{t}) - x_0(\cdot \wedge t_0)\|_T \leq 1$ . Recalling that  $h_1(r) = h_2(r)$ ,  $\forall r \leq 1$ , we obtain

$$h_2(\|\bar{x}(\cdot \wedge \bar{t}) - x_0(\cdot \wedge t_0)\|_T) \leq \frac{\lambda}{\delta},$$

which can be rewritten as

$$\|\bar{x}(\cdot \wedge \bar{t}) - x_0(\cdot \wedge t_0)\|_T^{d+1} \leq \frac{\lambda}{\delta} \frac{2 + C_\zeta + \|\bar{x}(\cdot \wedge \bar{t}) - x_0(\cdot \wedge t_0)\|_T}{\alpha_d}.$$

Since  $\|\bar{x}(\cdot \wedge \bar{t}) - x_0(\cdot \wedge t_0)\|_T \leq 1$ , we end up with (4.11).

STEP IV. Let us prove that  $\bar{t} < T$ , whenever  $\delta$  is large enough. As a matter of fact, letting  $\delta \rightarrow +\infty$ , it follows from item i) of STEP III and (4.3) that  $d_\infty((\bar{t}, \bar{x}), (t_0, x_0)) \rightarrow 0$ . Then, the claim follows from the fact that  $t_0 < T$ .

STEP V. By (4.9) and the definition of viscosity subsolution of (4.6) applied to  $\tilde{u}_1$  at the point  $(\bar{t}, \bar{x})$  with test function  $\tilde{v}_{n,\varepsilon} + \delta\varphi$ , we obtain

$$\begin{aligned} & -\partial_t^H(\tilde{v}_{n,\varepsilon} + \delta\varphi)(\bar{t}, \bar{x}) - \sup_{a \in A} \left\{ \langle b(\bar{t}, \bar{x}, a), \partial_x^V(\tilde{v}_{n,\varepsilon} + \delta\varphi)(\bar{t}, \bar{x}) \rangle \right. \\ & \left. + \frac{1}{2} \text{tr}[(\sigma\sigma^\top)(\bar{t}, \bar{x}, a) \partial_{xx}^V(\tilde{v}_{n,\varepsilon} + \delta\varphi)(\bar{t}, \bar{x})] + \tilde{f}(\bar{t}, \bar{x}, a) \right\} + \sqrt{\delta} \tilde{u}_1(\bar{t}, \bar{x}) \leq 0. \end{aligned}$$

This implies that

$$\begin{aligned} \sqrt{\delta} \tilde{u}_1(\bar{t}, \bar{x}) & \leq \delta \partial_t^H \varphi(\bar{t}, \bar{x}) + \delta \sup_{a \in A} \left\{ \langle b(\bar{t}, \bar{x}, a), \partial_x^V \varphi(\bar{t}, \bar{x}) \rangle + \frac{1}{2} \text{tr}[(\sigma\sigma^\top)(\bar{t}, \bar{x}, a) \partial_{xx}^V \varphi(\bar{t}, \bar{x})] \right\} \\ & + \partial_t^H \tilde{v}_{n,\varepsilon}(\bar{t}, \bar{x}) + \sup_{a \in A} \left\{ \langle b(\bar{t}, \bar{x}, a), \partial_x^V \tilde{v}_{n,\varepsilon}(\bar{t}, \bar{x}) \rangle + \frac{1}{2} \text{tr}[(\sigma\sigma^\top)(\bar{t}, \bar{x}, a) \partial_{xx}^V \tilde{v}_{n,\varepsilon}(\bar{t}, \bar{x})] + \tilde{f}(\bar{t}, \bar{x}, a) \right\}. \end{aligned}$$

Recalling that  $\tilde{v}_{n,\varepsilon}$  is a classical solution of equation (4.7), we find

$$\begin{aligned}
& \sqrt{\delta} (\tilde{u}_1 - \tilde{v}_{n,\varepsilon})(\bar{t}, \bar{x}) \\
& \leq \delta \partial_t^H \varphi(\bar{t}, \bar{x}) + \delta \sup_{a \in A} \left\{ \langle b(\bar{t}, \bar{x}, a), \partial_x^V \varphi(\bar{t}, \bar{x}) \rangle + \frac{1}{2} \text{tr} [(\sigma \sigma^\top)(\bar{t}, \bar{x}, a) \partial_{xx}^V \varphi(\bar{t}, \bar{x})] \right\} \\
& \quad + \sup_{a \in A} \left\{ \langle b(\bar{t}, \bar{x}, a) - b_n(\bar{t}, \bar{x}, a), \partial_x^V \tilde{v}_{n,\varepsilon}(\bar{t}, \bar{x}) \rangle + \tilde{f}(\bar{t}, \bar{x}, a) - \tilde{f}_n(\bar{t}, \bar{x}, a) \right. \\
& \quad \left. + \frac{1}{2} \text{tr} [(\sigma \sigma^\top - \sigma_n \sigma_n^\top)(\bar{t}, \bar{x}, a) \partial_{xx}^V \tilde{v}_{n,\varepsilon}(\bar{t}, \bar{x})] \right\} - \frac{1}{2} \varepsilon^2 e^{\sqrt{\delta}(\bar{t}+1)} \text{tr} [\partial_{yy} \bar{v}_{n,\varepsilon}(\bar{t}, y_n^{\bar{t}, \bar{x}})],
\end{aligned}$$

where  $y_n^{\bar{t}, \bar{x}}$  is given by (4.5) with  $t$  and  $x$  replaced respectively by  $\bar{t}$  and  $\bar{x}$ . By item ii) of STEP III, subtracting  $\delta^{3/2} \varphi(\bar{t}, \bar{x})$  from both sides of the above inequality and using that  $\varphi \geq 0$ , we obtain

$$\begin{aligned}
& \sqrt{\delta} e^{\sqrt{\delta}(t_0+1)} (u_1 - v_{n,\varepsilon})(t_0, x_0) = \sqrt{\delta} (\tilde{u}_1 - \tilde{v}_{n,\varepsilon})(t_0, x_0) \leq \sqrt{\delta} (\tilde{u}_1 - (\tilde{v}_{n,\varepsilon} + \delta \varphi))(\bar{t}, \bar{x}) \\
& \leq \delta \partial_t^H \varphi(\bar{t}, \bar{x}) + \delta \sup_{a \in A} \left\{ \langle b(\bar{t}, \bar{x}, a), \partial_x^V \varphi(\bar{t}, \bar{x}) \rangle + \frac{1}{2} \text{tr} [(\sigma \sigma^\top)(\bar{t}, \bar{x}, a) \partial_{xx}^V \varphi(\bar{t}, \bar{x})] \right\} \\
& \quad + \sup_{a \in A} \left\{ \langle b(\bar{t}, \bar{x}, a) - b_n(\bar{t}, \bar{x}, a), \partial_x^V \tilde{v}_{n,\varepsilon}(\bar{t}, \bar{x}) \rangle + \tilde{f}(\bar{t}, \bar{x}, a) - \tilde{f}_n(\bar{t}, \bar{x}, a) \right. \\
& \quad \left. + \frac{1}{2} \text{tr} [(\sigma \sigma^\top - \sigma_n \sigma_n^\top)(\bar{t}, \bar{x}, a) \partial_{xx}^V \tilde{v}_{n,\varepsilon}(\bar{t}, \bar{x})] \right\} - \frac{1}{2} \varepsilon^2 e^{\sqrt{\delta}(\bar{t}+1)} \text{tr} [\partial_{yy} \bar{v}_{n,\varepsilon}(\bar{t}, y_n^{\bar{t}, \bar{x}})].
\end{aligned}$$

By Theorem 4.4 we know that  $\varphi$  has bounded derivatives, moreover  $b$  and  $\sigma$  have linear growth, hence the quantity

$$\partial_t^H \varphi(\bar{t}, \bar{x}) + \sup_{a \in A} \left\{ \langle b(\bar{t}, \bar{x}, a), \partial_x^V \varphi(\bar{t}, \bar{x}) \rangle + \frac{1}{2} \text{tr} [(\sigma \sigma^\top)(\bar{t}, \bar{x}, a) \partial_{xx}^V \varphi(\bar{t}, \bar{x})] \right\}$$

is bounded by  $M(1 + \|\bar{x}\|_t)$ , for some constant  $M \geq 0$ , independent of  $n, \varepsilon, \delta$ . Therefore, we get

$$\begin{aligned}
& \sqrt{\delta} e^{\sqrt{\delta}(t_0+1)} (u_1 - v_{n,\varepsilon})(t_0, x_0) \leq \delta M (1 + \|\bar{x}\|_t) \\
& \quad + \sup_{a \in A} \left\{ \langle b(\bar{t}, \bar{x}, a) - b_n(\bar{t}, \bar{x}, a), \partial_x^V \tilde{v}_{n,\varepsilon}(\bar{t}, \bar{x}) \rangle + \tilde{f}(\bar{t}, \bar{x}, a) - \tilde{f}_n(\bar{t}, \bar{x}, a) \right. \\
& \quad \left. + \frac{1}{2} \text{tr} [(\sigma \sigma^\top - \sigma_n \sigma_n^\top)(\bar{t}, \bar{x}, a) \partial_{xx}^V \tilde{v}_{n,\varepsilon}(\bar{t}, \bar{x})] \right\} - \frac{1}{2} \varepsilon^2 e^{\sqrt{\delta}(\bar{t}+1)} \text{tr} [\partial_{yy} \bar{v}_{n,\varepsilon}(\bar{t}, y_n^{\bar{t}, \bar{x}})].
\end{aligned}$$

From estimates (B.25) and (B.31), we obtain

$$\begin{aligned}
& \sqrt{\delta} e^{\sqrt{\delta}(t_0+1)} (u_1 - v_{n,\varepsilon})(t_0, x_0) \leq \delta M (1 + \|\bar{x}\|_t) \\
& \quad + e^{\sqrt{\delta}(\bar{t}+1)} C_{n,\varepsilon} (1 + \|\bar{x}\|_t)^q \sup_{a \in A} \left\{ |b(\bar{t}, \bar{x}, a) - b_n(\bar{t}, \bar{x}, a)| + |(\sigma \sigma^\top - \sigma_n \sigma_n^\top)(\bar{t}, \bar{x}, a)| \right\} \\
& \quad + e^{\sqrt{\delta}(\bar{t}+1)} \sup_{a \in A} \left\{ |f(\bar{t}, \bar{x}, a) - f_n(\bar{t}, \bar{x}, a)| \right\} - \frac{1}{2} \varepsilon^2 e^{\sqrt{\delta}(\bar{t}+1)} d d_n \bar{C}_n e^{\bar{C}_n(T-t)} (1 + |y_n^{\bar{t}, \bar{x}}|)^3.
\end{aligned}$$

From the linear growth condition of  $b, b_n, \sigma, \sigma_n$  and the boundedness of  $f, f_n$ , we get (possibly enlarging  $C_{n,\varepsilon}$ )

$$\begin{aligned} \sqrt{\delta} e^{\sqrt{\delta}(t_0+1)}(u_1 - v_{n,\varepsilon})(t_0, x_0) &\leq \delta M (1 + \|\bar{x}\|_t) + e^{\sqrt{\delta}(\bar{t}+1)} C_{n,\varepsilon} (1 + \|\bar{x}\|_t)^q \\ &\quad - \frac{1}{2} \varepsilon^2 e^{\sqrt{\delta}(\bar{t}+1)} d d_n \bar{C}_n e^{\bar{C}_n(T-t)} (1 + |y_n^{\bar{t},\bar{x}}|)^3. \end{aligned}$$

Hence

$$\begin{aligned} (u_1 - v_{n,\varepsilon})(t_0, x_0) &\leq \sqrt{\delta} e^{-\sqrt{\delta}(t_0+1)} M (1 + \|\bar{x}\|_t) + \frac{1}{\sqrt{\delta}} e^{\sqrt{\delta}(\bar{t}-t_0)} C_{n,\varepsilon} (1 + \|\bar{x}\|_t)^q \\ &\quad - \frac{1}{2\sqrt{\delta}} \varepsilon^2 e^{\sqrt{\delta}(\bar{t}-t_0)} d d_n \bar{C}_n e^{\bar{C}_n(T-t)} (1 + |y_n^{\bar{t},\bar{x}}|)^3. \end{aligned}$$

Recalling the bounds (4.10) and (4.11), we get, letting  $\delta \rightarrow +\infty$ ,

$$(u_1 - v_{n,\varepsilon})(t_0, x_0) \leq 0.$$

By items 4) and 5) of Theorem B.6, sending  $\varepsilon \rightarrow 0^+$  and then  $n \rightarrow +\infty$ , we obtain

$$(u_1 - v)(t_0, x_0) \leq 0,$$

which gives a contradiction to (4.4).  $\square$

Finally, we can state the following uniqueness result.

**Corollary 4.6.** *Suppose that Assumptions (A) and (B) hold. Then, the value function  $v$  in (2.5) is the unique bounded and continuous (path-dependent) viscosity solution of equation (3.5).*

**Proof.** If  $u$  is another bounded and continuous (path-dependent) viscosity solution of equation (3.5), then, by Theorem 4.5, we get the two following inequalities:

$$u \leq v, \quad v \leq u,$$

from which the claim follows.  $\square$

## Appendix A Pathwise derivatives and functional Itô's formula

In the present appendix, we briefly recall the definitions of pathwise (or functional) derivatives following [11, Section 2], for which we refer for more details.

As we follow the standard approach (as it was introduced in the seminal paper [21]), in order to introduce the pathwise derivatives for a map  $u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$ , we firstly define them for a map  $\hat{u}: [0, T] \times D([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$ , with  $D([0, T]; \mathbb{R}^d)$  being the set of càdlàg paths, endowed with the supremum norm  $\|\hat{x}\|_T = \sup_{s \in [0, T]} |\hat{x}(s)|$ , for every  $\hat{x} \in D([0, T]; \mathbb{R}^d)$ . We also define on  $[0, T] \times D([0, T]; \mathbb{R}^d)$  the pseudometric  $\hat{d}_\infty: ([0, T] \times D([0, T]; \mathbb{R}^d))^2 \rightarrow [0, \infty)$  as

$$\hat{d}_\infty((t, \hat{x}), (t', \hat{x}')) := |t - t'| + \|\hat{x}(\cdot \wedge t) - \hat{x}'(\cdot \wedge t')\|_T.$$

**Definition A.1.** Consider a map  $\hat{u}: [0, T] \times D([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$ .

- (i) For every  $(t, \hat{x}) \in [0, T] \times D([0, T]; \mathbb{R}^d)$ , with  $t < T$ , the horizontal derivative of  $\hat{u}$  at  $(t, \hat{x})$  is defined as (if the limit exists)

$$\partial_t^H \hat{u}(t, \hat{x}) := \lim_{\delta \rightarrow 0^+} \frac{\hat{u}(t + \delta, \hat{x}(\cdot \wedge t)) - \hat{u}(t, \hat{x}(\cdot \wedge t))}{\delta}.$$

At  $t = T$ , it is defined as

$$\partial_t^H \hat{u}(T, \hat{x}) := \lim_{t \rightarrow T^-} \partial_t^H \hat{u}(t, \hat{x}).$$

- (ii) For every  $(t, \hat{x}) \in [0, T] \times D([0, T]; \mathbb{R}^d)$ , the vertical derivatives of first and second-order of  $\hat{u}$  at  $(t, \hat{x})$  are defined as (if the limits exist)

$$\begin{aligned} \partial_{x_i}^V \hat{u}(t, \hat{x}) &:= \lim_{h \rightarrow 0} \frac{\hat{u}(t, \hat{x} + h e_i 1_{[t, T]}) - \hat{u}(t, \hat{x})}{h}, \\ \partial_{x_i x_j}^V \hat{u}(t, \hat{x}) &:= \partial_{x_j}^V (\partial_{x_i}^V \hat{u})(t, \hat{x}), \end{aligned}$$

where  $e_1, \dots, e_d$  is the standard orthonormal basis of  $\mathbb{R}^d$ . We also denote  $\partial_x^V \hat{u} = (\partial_{x_1}^V \hat{u}, \dots, \partial_{x_d}^V \hat{u})$  and  $\partial_{xx}^V \hat{u} = (\partial_{x_i x_j}^V \hat{u})_{i, j=1, \dots, d}$ .

**Definition A.2.**  $C^{1,2}([0, T] \times D([0, T]; \mathbb{R}^d))$  is the set of continuous real-valued maps  $\hat{u}$  defined on  $([0, T] \times D([0, T]; \mathbb{R}^d), \hat{d}_\infty)$ , such that  $\partial_t^H \hat{u}$ ,  $\partial_x^V \hat{u}$ ,  $\partial_{xx}^V \hat{u}$  exist everywhere on  $([0, T] \times D([0, T]; \mathbb{R}^d), \hat{d}_\infty)$  and are continuous.

We can now define the pathwise derivatives for a map  $u: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$ . To this end, the following consistency property plays a crucial role.

**Lemma A.3.** If  $\hat{u}_1, \hat{u}_2 \in C^{1,2}([0, T] \times D([0, T]; \mathbb{R}^d))$  coincide on continuous paths, namely

$$\hat{u}_1(t, x) = \hat{u}_2(t, x), \quad \forall (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d),$$

then the same holds for their pathwise derivatives: for every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ ,

$$\begin{aligned} \partial_t^H \hat{u}_1(t, x) &= \partial_t^H \hat{u}_2(t, x), \\ \partial_x^V \hat{u}_1(t, x) &= \partial_x^V \hat{u}_2(t, x), \\ \partial_{xx}^V \hat{u}_1(t, x) &= \partial_{xx}^V \hat{u}_2(t, x). \end{aligned}$$

**Proof.** See [11, Lemma 2.1]. □

We can now give the following definition.



**Definition A.4.**  $C^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  is the set of continuous real-valued maps  $u$  defined on  $([0, T] \times C([0, T]; \mathbb{R}^d), d_\infty)$ , for which there exists  $\hat{u} \in C^{1,2}([0, T] \times D([0, T]; \mathbb{R}^d))$  such that

$$u(t, x) = \hat{u}(t, x), \quad \forall (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d).$$

We also define, for every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ ,

$$\begin{aligned} \partial_t^H u(t, x) &:= \partial_t^H \hat{u}(t, x), \\ \partial_x^V u(t, x) &:= \partial_x^V \hat{u}(t, x), \\ \partial_{xx}^V u(t, x) &:= \partial_{xx}^V \hat{u}(t, x). \end{aligned}$$

**Remark A.5.** Thanks to the consistency property stated in Lemma A.3, the definition of pathwise derivatives of  $u$  does not depend on the map  $\hat{u}$  appearing in Definition A.4.

In the present paper we also need to consider the following subset of  $C^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$ .

**Definition A.6.** We denote by  $C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  the set of  $u \in C^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  such that  $u, \partial_t^H u, \partial_x^V u, \partial_{xx}^V u$  satisfy a polynomial growth condition: there exist constants  $M \geq 0$  and  $q \geq 0$  such that

$$|\partial_t^H u(t, x)| + |\partial_x^V u(t, x)| + |\partial_{xx}^V u(t, x)| \leq M (1 + \|x\|_t^q),$$

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

Finally, we state the so-called functional Itô formula.

**Theorem A.7** (Functional Itô's formula). *Let  $u \in C^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$ . Let also  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$  be a filtered probability space, with  $(\mathcal{F}_t)_{t \in [0, T]}$  satisfying the usual conditions, on which a  $d$ -dimensional continuous semimartingale  $X = (X_t)_{t \in [0, T]}$  is defined, with  $X = (X^1, \dots, X^d)$ . Then, it holds that*

$$\begin{aligned} u(t, X) &= u(0, X) + \int_0^t \partial_t^H u(s, X) ds + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \partial_{x_i x_j}^V u(s, X) d[X^i, X^j]_s \\ &\quad + \sum_{i=1}^d \int_0^t \partial_{x_i}^V u(s, X) dX_s^i, \end{aligned} \quad \text{for all } 0 \leq t \leq T, \mathbb{P}\text{-a.s.} \quad (\text{A.1})$$

*Proof.* See [11, Theorem 2.2]. □

## Appendix B Cylindrical approximations

### B.1 The deterministic calculus via regularization

In the present appendix, we need to consider “cylindrical” maps defined on  $C([0, T]; \mathbb{R}^d)$ , namely maps depending on a path  $x \in C([0, T]; \mathbb{R}^d)$  only through a finite number of integrals

with respect to  $x$ . An integral with respect to  $x$  can be formally written as “ $\int_{[0,T]} \varphi(s) dx(s)$ ”. In order to give a meaning to the latter notation, it is useful to notice that we look for a deterministic integral which coincides with the Itô integral when  $x$  is replaced by an Itô process (such a property will be exploited in the sequel). This is the case if we interpret “ $\int_{[0,T]} \varphi(s) dx(s)$ ” as the deterministic version of the forward integral, which we now introduce and denote by  $\int_{[0,T]} \varphi(s) d^-x(s)$ . For more details on such an integral and, more generally, on the deterministic calculus via regularization we refer to [18, Section 3.2] and [12, Section 2.2]. The only difference with respect to [18] and [12] being that here we consider  $d$ -dimensional paths (with  $d$  possibly greater than 1), even though, as usual, we work component by component, therefore relying on the one-dimensional theory.

We fix some notations. For every path  $x: [0, T] \rightarrow \mathbb{R}^d$  we consider the following extensions from  $[0, T]$  to  $\mathbb{R}$ :

$$x_{(0,T]}(s) := \begin{cases} x(0), & s < 0, \\ x(s), & s \in [0, T], \\ 0, & s > T, \end{cases} \quad x_{[0,T]}(s) := \begin{cases} 0, & s < 0, \\ x(s), & s \in [0, T], \\ x(T), & s > T. \end{cases}$$

**Definition B.1.** Let  $x: [0, T] \rightarrow \mathbb{R}^d$  and  $\varphi: [0, T] \rightarrow \mathbb{R}$  be càdlàg functions. When the following limit

$$\int_{[0,T]} \varphi(s) d^-x(s) := \lim_{\varepsilon \rightarrow 0^+} \int_{-\infty}^{+\infty} \varphi_{(0,T]}(s) \frac{x_{[0,T]}(s + \varepsilon) - x_{[0,T]}(s)}{\varepsilon} ds$$

exists and it is finite, we denote it by  $\int_{[0,T]} \varphi(s) d^-x(s)$  and call it **forward integral of  $\varphi$  with respect to  $x$** .

When  $\varphi$  is of bounded variation, an integration by parts formula provides an explicit representation of the forward integral of  $\varphi$  with respect to  $x$ . Its simple proof is omitted.

**Proposition B.2.** Let  $x: [0, T] \rightarrow \mathbb{R}^d$  and  $\varphi: [0, T] \rightarrow \mathbb{R}$  be càdlàg functions, with  $\varphi$  being of bounded variation. The following **integration by parts formula** holds:

$$\int_{[0,T]} \varphi(s) d^-x(s) = \varphi(T) x(T) - \int_{(0,T]} x(s) d\varphi(s), \quad (\text{B.1})$$

where  $\int_{(0,T]} x(s) d\varphi(s)$  is a Lebesgue-Stieltjes integral on  $(0, T]$ .

## B.2 Cylindrical approximations

**Lemma B.3.** Let  $h: [0, T] \times C([0, T]; \mathbb{R}^d) \times A \rightarrow \mathbb{R}$  be continuous and satisfying, for some constant  $K \geq 0$ , items a)-b)-d) or, alternatively, items a)-c)-d):

- a)  $|h(t, x, a) - h(t, x', a)| \leq K \|x - x'\|_t$ , for all  $t \in [0, T]$ ,  $x, x' \in C([0, T]; \mathbb{R}^d)$ ,  $a \in A$ ;
- b)  $|h(t, 0, a)| \leq K$ , for all  $(t, a) \in [0, T] \times A$ ;

- c)  $|h(t, x, a)| \leq K$ , for all  $(t, x, a) \in [0, T] \times C([0, T]; \mathbb{R}^d) \times A$ ;
- d)  $h$  is uniformly continuous in  $t$ , uniformly with respect to the other variables, namely there exists a modulus of continuity  $w: [0, \infty) \rightarrow [0, \infty)$  such that

$$|h(t, x, a) - h(s, x, a)| \leq w(|t - s|), \quad (\text{B.2})$$

for all  $t, s \in [0, T], x \in C([0, T]; \mathbb{R}^d), a \in A$ . The map  $w$  is continuous, increasing, subadditive, and  $w(0) = 0$ .

Then, there exists a sequence  $\{h_n\}_n$  with  $h_n: [0, T] \times C([0, T]; \mathbb{R}^d) \times A \rightarrow \mathbb{R}$  continuous and satisfying the following.

- 1)  $h_n$  converges pointwise to  $h$  uniformly with respect to  $a$ , namely: for every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , it holds that

$$\lim_{n \rightarrow +\infty} \sup_{a \in A} |h_n(t, x, a) - h(t, x, a)| = 0.$$

- 2) If  $h$  satisfies items a) and b) (resp. a) and c)) then  $h_n$  also satisfies the same items. In particular,  $h_n$  satisfies item a) with constant  $2K$  and item b) with a constant  $K_d \geq 0$ , depending only on  $K$  and  $d$  (resp. item c) with the same constant  $K$ ).
- 3) For every  $n$ , there exist  $d_n \in \mathbb{N}$ , a continuous function  $\bar{h}_n: [0, T] \times \mathbb{R}^{dd_n} \times A \rightarrow \mathbb{R}$ , and some continuously differentiable functions  $\phi_{n,1}, \dots, \phi_{n,d_n}: [0, T] \rightarrow \mathbb{R}$  such that

$$h_n(t, x, a) = \bar{h}_n\left(t, \int_{[0,t]} \phi_{n,1}(s) d^-x(s), \dots, \int_{[0,t]} \phi_{n,d_n}(s) d^-x(s), a\right),$$

for every  $(t, x, a) \in [0, T] \times C([0, T]; \mathbb{R}^d) \times A$ . Moreover,  $d_n$  and  $\phi_{n,1}, \dots, \phi_{n,d_n}$  do not depend on  $h$ .

- 4) If  $h$  satisfies items a) and b) (resp. a) and c)) then  $\bar{h}_n$  satisfies items i) and ii) (resp. i) and iii)) below:
- i)  $|\bar{h}_n(t, y, a) - \bar{h}_n(t, y', a)| \leq K |y - y'|$ , for all  $t \in [0, T], y, y' \in \mathbb{R}^{dd_n}, a \in A$ ;
  - ii)  $|\bar{h}_n(t, 0, a)| \leq K_d$ , for all  $(t, a) \in [0, T] \times A$ , for some constant  $K_d \geq 0$ , depending only on  $K$  and  $d$ ;
  - iii)  $|\bar{h}_n(t, y, a)| \leq K$ , for all  $(t, y, a) \in [0, T] \times \mathbb{R}^{dd_n} \times A$ .
- 5) For every  $n$  and any  $a \in A$ , the function  $\bar{h}_n(\cdot, \cdot, a)$  is  $C^{1,2}([0, T] \times \mathbb{R}^{dd_n})$ . Moreover, there exist constants  $K_n \geq 0$  and  $q \in \{0, 1\}$  such that

$$|\partial_t \bar{h}_n(t, y, a)| + |\partial_y \bar{h}_n(t, y, a)| + |\partial_{yy}^2 \bar{h}_n(t, y, a)| \leq K_n (1 + |y|)^q, \quad (\text{B.3})$$

for all  $(t, y, a) \in [0, T] \times \mathbb{R}^{dd_n} \times A$ . The constant  $q$  is equal to 1 if  $h$  satisfies item b), while it is equal to 0 if  $h$  satisfies item c).

**Proof.** We split the proof into six steps.

**Step I. Definitions of  $x_{n,y}^{\text{pol}}$  and  $x_n^{t,\text{pol}}$ .** For every  $n \in \mathbb{N}$ , consider the  $n$ -th dyadic mesh of the time interval  $[0, T]$ , that is

$$0 = t_0^n < t_1^n < \dots < t_{2^n}^n = T, \quad \text{with } t_j^n := \frac{j}{2^n}T, \text{ for every } j = 0, \dots, 2^n.$$

For every  $y = (y_0, \dots, y_{2^n}) \in \mathbb{R}^{d \cdot (2^n+1)}$ , we consider the corresponding  $n$ -th polygonal, denoted  $x_{n,y}^{\text{pol}}$ , which is an element of  $C([0, T]; \mathbb{R}^d)$  and is characterized by the following properties:

- $x_{n,y}^{\text{pol}}(t_j^n) = y_j$ , for every  $j = 0, \dots, 2^n$ ;
- $x_{n,y}^{\text{pol}}$  is linear on every interval  $[t_{j-1}^n, t_j^n]$ , for any  $j = 1, \dots, 2^n$ .

So, in particular,  $x_{n,y}^{\text{pol}}$  is given by the following formula:

$$x_{n,y}^{\text{pol}}(s) = \frac{y_j - y_{j-1}}{t_j^n - t_{j-1}^n} s + \frac{t_j^n y_{j-1} - t_{j-1}^n y_j}{t_j^n - t_{j-1}^n},$$

for every  $s \in [t_{j-1}^n, t_j^n]$  and any  $j = 1, \dots, 2^n$ .

Now, given  $t \in [0, T]$  and  $x \in C([0, T]; \mathbb{R}^d)$ , we denote

$$x_n^{t,\text{pol}} := x_{n,\tilde{y}_n}^{\text{pol}},$$

with

$$\tilde{y}_n^{t,x} := (x(t_0^n \wedge t), \dots, x(t_{2^n}^n \wedge t)) = \left( \int_{[0,t]} 1_{[0,t_0^n]}(s) d^-x(s), \dots, \int_{[0,t]} 1_{[0,t_{2^n}^n]}(s) d^-x(s) \right),$$

where the second inequality follows from the integration by parts formula (B.1).

It is easy to see that (in the following formulae we use the same symbol, that is  $|\cdot|$ , to denote the Euclidean norms on  $\mathbb{R}^d$  and  $\mathbb{R}^{d \cdot (2^n+1)}$ )

$$\|x_{n,y}^{\text{pol}}\|_T \leq \max_j |y_j| \leq |y|, \quad \|x_{n,y}^{\text{pol}} - x_{n,\tilde{y}}^{\text{pol}}\|_T \leq \max_j |y_j - \tilde{y}_j| \leq |y - \tilde{y}|, \quad (\text{B.4})$$

for every  $y = (y_0, \dots, y_{2^n}), \tilde{y} = (\tilde{y}_0, \dots, \tilde{y}_{2^n}) \in \mathbb{R}^{d \cdot (2^n+1)}$ . Similarly, we have

$$\|x_n^{t,\text{pol}}\|_t \leq \|x\|_t, \quad \|x_n^{t,\text{pol}} - x\|_t \leq \sup_{|r-s| \leq 2^{-n}} |x(r \wedge t) - x(s \wedge t)| \leq \sup_{|r-s| \leq 2^{-n}} |x(r) - x(s)|. \quad (\text{B.5})$$

**Step II. Definitions of  $\phi_{n,j}$  and  $y_n^{t,x}$ .** Let  $\chi: \mathbb{R} \rightarrow \mathbb{R}$  be given by

$$\chi(r) = \begin{cases} 0, & r \leq 0, \\ \frac{1}{1 + \frac{e^{1/r}}{e^{1/(1-r)}}}, & 0 < r < 1, \\ 1, & r \geq 1. \end{cases}$$

Notice that  $\chi$  belongs to  $C^\infty(\mathbb{R})$  and is strictly increasing on  $[0, 1]$ . Then, for every  $n \in \mathbb{N}$ , we define the functions  $\phi_{n,0}, \dots, \phi_{n,2^n}: [0, T] \rightarrow \mathbb{R}$  as follows:

$$\phi_{n,j}(r) = \begin{cases} 1, & 0 \leq r < t_j^n, \\ 1 - \chi(2^{2n}(r - t_j^n)), & t_j^n \leq r \leq T, \end{cases} \quad (\text{B.6})$$

for  $j = 0, \dots, 2^n$ ; so, in particular,

$$\phi_{n,0}(r) = 1 - \chi(2^{2n}r), \quad 0 \leq r \leq T.$$

Moreover, for every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , let  $y_n^{t,x} \in \mathbb{R}^{d \cdot (2^n+1)}$  be defined as

$$y_n^{t,x} := \left( \int_{[0,t]} \phi_{n,0}(s) d^-x(s), \dots, \int_{[0,t]} \phi_{n,2^n}(s) d^-x(s) \right). \quad (\text{B.7})$$

In the rest of this step we prove that the following estimate holds:

$$\|x_{n,y_n^{t,x}}^{\text{pol}} - x_n^{t,\text{pol}}\|_t \leq 2 \sup_{|r-s| \leq 2^{-2n}} |x(r \wedge t) - x(s \wedge t)| \leq 2 \sup_{|r-s| \leq 2^{-2n}} |x(r) - x(s)|. \quad (\text{B.8})$$

We begin noting that

$$\|x_{n,y_n^{t,x}}^{\text{pol}} - x_n^{t,\text{pol}}\|_t = \max_{j=0, \dots, 2^n} |(y_n^{t,x})_j - x(t_j^n \wedge t)|.$$

where  $y_n^{t,x} = ((y_n^{t,x})_0, \dots, (y_n^{t,x})_{2^n}) \in \mathbb{R}^{d \cdot (2^n+1)}$ . By the integration by parts formula (B.1), we have

$$\begin{aligned} (y_n^{t,x})_j - x(t_j^n \wedge t) &= \int_{[0,t]} \phi_{n,j}(s) d^-x(s) - x(t_j^n \wedge t) \\ &= \int_{[0, (t_j^n + 2^{-2n}) \wedge t]} \phi_{n,j}(s) d^-x(s) - x(t_j^n \wedge t) \\ &= \phi_{n,j}((t_j^n + 2^{-2n}) \wedge t) x((t_j^n + 2^{-2n}) \wedge t) - x(t_j^n \wedge t) - \int_0^{(t_j^n + 2^{-2n}) \wedge t} x(s) \phi'_{n,j}(s) ds. \end{aligned}$$

Notice that  $\phi_{n,j}(s) = 1$  and  $\phi'_{n,j}(s) = 0$ , for  $0 \leq s \leq t_j^n$ . Now, we distinguish two cases.

1. If  $t \leq t_j^n$ , we have

$$(y_n^{t,x})_j - x(t_j^n \wedge t) = 0. \quad (\text{B.9})$$

2. If  $t > t_j^n$ , we have

$$\begin{aligned} (y_n^{t,x})_j - x(t_j^n \wedge t) &= \phi_{n,j}((t_j^n + 2^{-2n}) \wedge t) x((t_j^n + 2^{-2n}) \wedge t) - x(t_j^n) \\ &\quad - \int_{t_j^n}^{(t_j^n + 2^{-2n}) \wedge t} x(s) \phi'_{n,j}(s) ds. \end{aligned}$$

Observe that  $\int_{t_j^n}^{(t_j^n+2^{-2n})\wedge t} \phi'_{n,j}(s) ds = \phi_{n,j}((t_j^n+2^{-2n})\wedge t) - 1$ , then

$$\begin{aligned} (y_n^{t,x})_j - x(t_j^n \wedge t) &= x((t_j^n+2^{-2n})\wedge t) - x(t_j^n) \\ &\quad + \int_{t_j^n}^{(t_j^n+2^{-2n})\wedge t} (x((t_j^n+2^{-2n})\wedge t) - x(s)) \phi'_{n,j}(s) ds \\ &\leq \sup_{|r-s|\leq 2^{-2n}} |x(r\wedge t) - x(s\wedge t)| \left( 1 + \int_{t_j^n}^{(t_j^n+2^{-2n})\wedge t} |\phi'_{n,j}(s)| ds \right) \\ &\leq \sup_{|r-s|\leq 2^{-n}} |x(r\wedge t) - x(s\wedge t)| \left( 1 + \int_{t_j^n}^{(t_j^n+2^{-2n})\wedge t} |\phi'_{n,j}(s)| ds \right). \end{aligned}$$

Since  $\int_{t_j^n}^{(t_j^n+2^{-2n})\wedge t} |\phi'_{n,j}(s)| ds = -\int_{t_j^n}^{(t_j^n+2^{-2n})\wedge t} \phi'_{n,j}(s) ds = 1 - \phi_{n,j}((t_j^n+2^{-2n})\wedge t)$  and  $1 - \phi_{n,j}((t_j^n+2^{-2n})\wedge t) \leq 1$ , we get

$$|(y_n^{t,x})_j - x(t_j^n \wedge t)| \leq 2 \sup_{|r-s|\leq 2^{-n}} |x(r\wedge t) - x(s\wedge t)| \leq 2 \sup_{|r-s|\leq 2^{-n}} |x(r) - x(s)|. \quad (\text{B.10})$$

From (B.9) and (B.10) we conclude that (B.8) holds.

**Step III. Definitions of  $\bar{h}_n, h_n$  and proof of item 3).** For every  $n \in \mathbb{N}$ , let  $\eta_n: \mathbb{R} \rightarrow \mathbb{R}$  be given by

$$\eta_n(s) = \frac{2n}{\sqrt{2\pi}} e^{-\frac{n^2}{2}s^2}, \quad \forall s \in \mathbb{R}.$$

Notice that  $\int_0^\infty \eta_n(s) ds = 1$ . Moreover, for every  $n \in \mathbb{N}$ , let  $\zeta_n: \mathbb{R}^{d(2^n+1)} \rightarrow \mathbb{R}$  be the probability density function of the multivariate normal distribution  $\mathcal{N}(0, (d(2^n+1))^{-2} I_{d(2^n+1)})$ , where  $I_{d(2^n+1)}$  denotes the identity matrix of order  $d(2^n+1)$ :

$$\zeta_n(z) = \frac{(d(2^n+1))^{d(2^n+1)}}{(2\pi)^{d(2^n+1)/2}} e^{-\frac{(d(2^n+1))^2}{2}|z|^2}, \quad \forall z \in \mathbb{R}^{d(2^n+1)}.$$

Now, define  $\bar{h}_n: [0, T] \times \mathbb{R}^{d(2^n+1)} \times A \rightarrow \mathbb{R}$  as follows:

$$\bar{h}_n(t, y, a) = \int_0^\infty \int_{\mathbb{R}^{d(2^n+1)}} \eta_n(s) \zeta_n(z) h((t+s)\wedge T, x_{n,y+z}^{\text{pol}}, a) ds dz,$$

for all  $(t, y, a) \in [0, T] \times \mathbb{R}^{d(2^n+1)} \times A$ . Finally, let  $h_n: [0, T] \times C([0, T]; \mathbb{R}^d) \times A \rightarrow \mathbb{R}$  be given by

$$h_n(t, x, a) = \bar{h}_n\left(t, \int_{[0,t]} \phi_{n,0}(s) d^-x(s), \dots, \int_{[0,t]} \phi_{n,2^n}(s) d^-x(s), a\right),$$

with  $\phi_{n,j}$  as in (B.6). From the continuity of  $h$ , we see that both  $h_n$  and  $\bar{h}_n$  are continuous.

**Step IV. Proof of item 1).** For every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , let  $y_n^{t,x} \in \mathbb{R}^{d \cdot (2^n+1)}$  be given by (B.7). Then, we have (using also the equality  $h(t, x, a) = h(t, x_{\cdot \wedge t}, a)$ )

$$\begin{aligned}
& |h_n(t, x, a) - h(t, x, a)| \\
& \leq \int_0^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(s) \zeta_n(z) |h((t+s) \wedge T, x_{n, y_n^{t,x}+z}^{\text{pol}}, a) - h(t, x_{\cdot \wedge t}, a)| ds dz \\
& \leq \int_0^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(s) \zeta_n(z) |h((t+s) \wedge T, x_{n, y_n^{t,x}+z}^{\text{pol}}, a) - h((t+s) \wedge T, x_{\cdot \wedge t}, a)| ds dz \\
& \quad + \int_0^\infty \eta_n(s) |h((t+s) \wedge T, x_{\cdot \wedge t}, a) - h(t, x_{\cdot \wedge t}, a)| ds \\
& \leq K \int_0^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(s) \zeta_n(z) \|x_{n, y_n^{t,x}+z}^{\text{pol}} - x_{\cdot \wedge t}\|_{(t+s) \wedge T} ds dz \\
& \quad + \int_0^\infty \eta_n(s) |h((t+s) \wedge T, x_{\cdot \wedge t}, a) - h(t, x_{\cdot \wedge t}, a)| ds.
\end{aligned}$$

Since both paths  $x_{n, y_n^{t,x}+z}^{\text{pol}}$  and  $x_{\cdot \wedge t}$  are constant after time  $t$ , moreover recalling (B.4), (B.5) and the linearity of the map  $y \mapsto x_{n, y}^{\text{pol}}$ , we obtain

$$\begin{aligned}
\|x_{n, y_n^{t,x}+z}^{\text{pol}} - x_{\cdot \wedge t}\|_{(t+s) \wedge T} &= \|x_{n, y_n^{t,x}}^{\text{pol}} + x_{n, z}^{\text{pol}} - x_{\cdot \wedge t}\|_t \leq \|x_{n, y_n^{t,x}}^{\text{pol}} - x_{\cdot \wedge t}\|_t + |z| \\
&\leq \|x_{n, y_n^{t,x}}^{\text{pol}} - x_n^{\text{pol}}\|_t + \|x_n^{\text{pol}} - x\|_t + |z|.
\end{aligned}$$

Then, by (B.5) and (B.8), we get

$$\begin{aligned}
|h_n(t, x, a) - h(t, x, a)| &\leq 3K \sup_{|r-s| \leq 2^{-n}} |x(r) - x(s)| + K \int_{\mathbb{R}^{d \cdot (2^n+1)}} \zeta_n(z) |z| dz \quad (\text{B.11}) \\
&\quad + \int_0^\infty \eta_n(s) |h((t+s) \wedge T, x_{\cdot \wedge t}, a) - h(t, x_{\cdot \wedge t}, a)| ds.
\end{aligned}$$

Now, by (B.2) we have

$$\begin{aligned}
& \int_0^\infty \eta_n(s) |h((t+s) \wedge T, x_{\cdot \wedge t}, a) - h(t, x_{\cdot \wedge t}, a)| ds \\
& \leq \int_0^\infty \eta_n(s) w((t+s) \wedge T - t) ds = \int_0^{T-t} \eta_n(s) w(s) ds + w(T-t) \int_{T-t}^\infty \eta_n(s) ds \\
& = \int_0^{n(T-t)} \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2} w(r/n) dr + w(T-t) \int_{n(T-t)}^\infty \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2} dr. \quad (\text{B.12})
\end{aligned}$$

Finally, consider the integral  $\int_{\mathbb{R}^{d \cdot (2^n+1)}} \zeta_n(z) |z| dz$ . Since the integrand is a radial function, it is more convenient to rewrite it in terms of spherical coordinates (see for instance [28, Appendix C.3]). In particular, denoting by  $S_{d(2^n+1)-1}(R)$  the surface area of the sphere  $\{|z| = R\}$ , which is equal to  $2\pi^{d(2^n+1)/2} R^{d(2^n+1)-1} / \Gamma(d(2^n+1)/2)$ , with  $\Gamma(\cdot)$  being the Gamma

function, we get

$$\begin{aligned}
\int_{\mathbb{R}^{d \cdot (2^n + 1)}} \zeta_n(z) |z| dz &= \int_{\mathbb{R}^{d \cdot (2^n + 1)}} \frac{(d(2^n + 1))^{d(2^n + 1)}}{(2\pi)^{d(2^n + 1)/2}} e^{-\frac{(d(2^n + 1))^2}{2} |z|^2} |z| dz \\
&= \frac{1}{d(2^n + 1)} \int_{\mathbb{R}^{d \cdot (2^n + 1)}} \frac{1}{(2\pi)^{d(2^n + 1)/2}} e^{-\frac{1}{2}|y|^2} |y| dy \\
&= \frac{1}{d(2^n + 1)} \frac{1}{(2\pi)^{(d(2^n + 1) - 1)/2}} \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}R^2} R S_{d(2^n + 1) - 1}(R) dR \\
&= \frac{1}{d(2^n + 1)} \frac{2\pi^{d(2^n + 1)/2}}{(2\pi)^{(d(2^n + 1) - 1)/2}} \frac{1}{\Gamma(d(2^n + 1)/2)} \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}R^2} R^{d(2^n + 1)} dR \\
&= \frac{1}{d(2^n + 1)} \frac{\pi^{d(2^n + 1)/2}}{(2\pi)^{(d(2^n + 1) - 1)/2}} \frac{1}{\Gamma(d(2^n + 1)/2)} \frac{2^{d(2^n + 1)/2}}{\sqrt{\pi}} \Gamma((d(2^n + 1) + 1)/2) \\
&= \sqrt{2} \frac{\Gamma((d(2^n + 1) + 1)/2)}{d(2^n + 1) \Gamma(d(2^n + 1)/2)} = \frac{(d(2^n + 1) - 1)!!}{(d(2^n + 1))!!}. \tag{B.13}
\end{aligned}$$

In conclusion, plugging (B.12) and (B.13) into (B.11), we obtain

$$\begin{aligned}
|h_n(t, x, a) - h(t, x, a)| &\leq 3K \sup_{|r-s| \leq 2^{-n}} |x(r) - x(s)| + K \frac{(d(2^n + 1) - 1)!!}{(d(2^n + 1))!!} \\
&\quad + \int_0^{n(T-t)} \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2} w(r/n) dr + w(T-t) \int_{n(T-t)}^\infty \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2} dr. \tag{B.14}
\end{aligned}$$

By Stirling's formula, we see that  $\frac{(d(2^n + 1) - 1)!!}{(d(2^n + 1))!!} = O(1/\sqrt{d(2^n + 1)})$  as  $n \rightarrow \infty$ . Then, letting  $n \rightarrow \infty$  in (B.14), using Lebesgue's dominated convergence theorem together with the fact that  $w$  has linear growth (see [29, Remark D.4]), we deduce that item 1) holds.

**Step V. Proof of items 2) and 4).** It is clear that  $h_n$  (resp.  $\bar{h}_n$ ) satisfies item c) (resp. 4)-iii)) with the same constant  $K$ . If  $h$  satisfies item b) then  $|h(t, x, a)| \leq K(1 + \|x\|_t)$ , therefore, by (B.4),

$$\begin{aligned}
|\bar{h}_n(t, y, a)| &\leq K \int_{\mathbb{R}^{d \cdot (2^n + 1)}} \zeta_n(z) (1 + \|x_{n, y+z}^{\text{pol}}\|_t) dz \\
&\leq K(1 + |y|) + K \int_{\mathbb{R}^{d \cdot (2^n + 1)}} \zeta_n(z) |z| dz \\
&= K(1 + |y|) + 2K \frac{(d(2^n + 1) - 1)!!}{(d(2^n + 1))!!},
\end{aligned}$$

where the last equality follows from (B.13). Recall that, by Stirling's formula,  $\frac{(d(2^n + 1) - 1)!!}{(d(2^n + 1))!!} = O(1/\sqrt{d(2^n + 1)})$  as  $n \rightarrow \infty$ . So, in particular, the sequence  $\left\{ \frac{(d(2^n + 1) - 1)!!}{(d(2^n + 1))!!} \right\}_n$  is bounded, as a consequence there exists a constant  $C_d \geq 0$  such that

$$|\bar{h}_n(t, y, a)| \leq K(1 + |y|) + 2K C_d,$$



which proves item 4)-ii). Concerning  $h_n$ , we have

$$\begin{aligned} |h_n(t, x, a)| &\leq K (1 + \|y_n^{t,x}\|_t) + 2K C_d \\ &\leq K (1 + \|y_n^{t,x} - x_n^{t,\text{pol}}\|_t + \|x_n^{t,\text{pol}}\|_t) + 2K C_d. \end{aligned}$$

By (B.8) and (B.5), we obtain

$$\begin{aligned} |h_n(t, x, a)| &\leq K \left( 1 + 2 \sup_{|r-s| \leq 2^{-n}} |x(r \wedge t) - x(s \wedge t)| + \|x\|_t \right) + 2K C_d \\ &\leq K (1 + 3\|x\|_t) + 2K C_d, \end{aligned}$$

which proves that  $h_n$  satisfies item b) with a constant  $K_d$ , depending only on  $K$  and  $d$ . Let us now prove that  $h_n$  and  $\bar{h}_n$  satisfy respectively item a) and item 4)-i). We have

$$\begin{aligned} &|\bar{h}_n(t, y, a) - \bar{h}_n(t, \tilde{y}, a)| \\ &\leq \int_0^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(s) \zeta_n(z) |h((t+s) \wedge T, x_{n,y+z}^{\text{pol}}, a) - h((t+s) \wedge T, x_{n,\tilde{y}+z}^{\text{pol}}, a)| ds dz \\ &\leq K \int_0^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(s) \zeta_n(z) \|x_{n,y+z}^{\text{pol}} - x_{n,\tilde{y}+z}^{\text{pol}}\|_{(t+s) \wedge T} ds dz \\ &= K \int_0^\infty \eta_n(s) \|x_{n,y}^{\text{pol}} - x_{n,\tilde{y}}^{\text{pol}}\|_{(t+s) \wedge T} ds, \end{aligned}$$

where the last equality follows from the linearity of the map  $y \mapsto x_{n,y}^{\text{pol}}$ . Hence, recalling (B.4), we obtain

$$|\bar{h}_n(t, y, a) - \bar{h}_n(t, \tilde{y}, a)| \leq K \max_j |y_j - \tilde{y}_j| \leq K |y - \tilde{y}|, \quad (\text{B.15})$$

which proves that  $\bar{h}_n$  satisfies item 4)-i). Let us now prove that  $h_n$  satisfies item a). From (B.15) we have

$$|h_n(t, x, a) - h_n(t, \tilde{x}, a)| \leq K \max_j |(y_n^{t,x})_j - (y_n^{t,\tilde{x}})_j|.$$

By the integration by parts formula (B.1), we get

$$\begin{aligned} (y_n^{t,x})_j - (y_n^{t,\tilde{x}})_j &= \int_{[0,t]} \phi_{n,j}(s) d^- x(s) - \int_{[0,t]} \phi_{n,j}(s) d^- \tilde{x}(s) \\ &= \phi_{n,j}(t) (x(t) - \tilde{x}(t)) - \int_0^t (x(s) - \tilde{x}(s)) \phi'_{n,j}(s) ds. \end{aligned}$$

Hence

$$|(y_n^{t,x})_j - (y_n^{t,\tilde{x}})_j| \leq |x(t) - \tilde{x}(t)| + \|x - \tilde{x}\|_t \int_0^t |\phi'_{n,j}(s)| ds.$$

Since  $\int_0^t |\phi'_{n,j}(s)| ds = -\int_0^t \phi'_{n,j}(s) ds = 1 - \phi_{n,j}(t)$  and  $1 - \phi_{n,j}(t) \leq 1$ , we conclude that

$$|h_n(t, x, a) - h_n(t, \tilde{x}, a)| \leq 2K \|x - \tilde{x}\|_t,$$

which proves that  $h_n$  satisfies item a) with constant  $2K$ .

**Step VI. Proof of item 5).** Recall that

$$\begin{aligned}\bar{h}_n(t, y, a) &= \int_0^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(s) \zeta_n(z) h((t+s) \wedge T, x_{n,y+z}^{\text{pol}}, a) ds dz \\ &= \int_t^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(s-t) \zeta_n(z-y) h(s \wedge T, x_{n,z}^{\text{pol}}, a) ds dz,\end{aligned}$$

for all  $(t, y, a) \in [0, T] \times \mathbb{R}^{d \cdot (2^n+1)} \times A$ . Then, it is clear that, for every  $a \in A$ , the function  $\bar{h}_n(\cdot, \cdot, a) \in C^{1,2}([0, T] \times \mathbb{R}^{dd_n})$ . Moreover, by direct calculation, we have

$$\begin{aligned}\partial_t \bar{h}_n(t, y, a) &= - \int_t^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta'_n(s-t) \zeta_n(z-y) h(s \wedge T, x_{n,z}^{\text{pol}}, a) ds dz \\ &\quad - \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(0) \zeta_n(z-y) h(t \wedge T, x_{n,z}^{\text{pol}}, a) dz, \\ \partial_y \bar{h}_n(t, y, a) &= - \int_t^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(s-t) \partial_y \zeta_n(z-y) h(s \wedge T, x_{n,z}^{\text{pol}}, a) ds dz, \\ \partial_{yy}^2 \bar{h}_n(t, y, a) &= \int_t^\infty \int_{\mathbb{R}^{d \cdot (2^n+1)}} \eta_n(s-t) \partial_{yy}^2 \zeta_n(z-y) h(s \wedge T, x_{n,z}^{\text{pol}}, a) ds dz.\end{aligned}$$

By item b) or, alternatively, item c), we conclude that (B.3) holds.  $\square$

Under Assumptions (A) and (B), we can apply Lemma B.3 to  $b, \sigma, f, g$ , from which we get sequences  $\{b_n\}_n, \{\sigma_n\}_n, \{f_n\}_n, \{g_n\}_n$ , with

$$b_n, \sigma_n, f_n: [0, T] \times C([0, T]; \mathbb{R}^d) \times A \longrightarrow \mathbb{R}^d, \mathbb{R}^{d \times m}, \mathbb{R}, \quad g_n: C([0, T]; \mathbb{R}^d) \longrightarrow \mathbb{R}, \quad (\text{B.16})$$

satisfying items 1)-2)-3)-4)-5) of Lemma B.3. We also recall from Lemma B.3 that  $d_n$  and  $\phi_{n,1}, \dots, \phi_{n,d_n}$  are the same for  $b, \sigma, f, g$ .

Finally, let

$$v_n: [0, T] \times C([0, T]; \mathbb{R}^d) \longrightarrow \mathbb{R} \quad (\text{B.17})$$

denote the value function of the optimal control problem with coefficients  $b_n, \sigma_n, f_n, g_n$ .

**Lemma B.4.** *Let Assumptions (A) and (B) hold. Consider the sequences  $\{b_n\}_n, \{\sigma_n\}_n, \{f_n\}_n, \{g_n\}_n, \{v_n\}_n$  in (B.16)-(B.17). Then,  $v_n$  converges pointwise to  $v$  in (2.5) as  $n \rightarrow +\infty$ .*

**Proof.** For every  $n \in \mathbb{N}, t \in [0, T], x \in C([0, T]; \mathbb{R}^d), \alpha \in \mathcal{A}$ , let  $X^{n,t,x,\alpha} \in \mathbf{S}_2(\mathbb{F})$  be the unique solution to the following system of controlled stochastic differential equations:

$$\begin{cases} dX_s = b_n(s, X, \alpha_s) ds + \sigma_n(s, X, \alpha_s) dB_s, & s \in (t, T], \\ X_s = x(s), & s \in [0, t]. \end{cases}$$

Then

$$\begin{aligned}
& |v_n(t, x) - v(t, x)| \\
& \leq \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \int_t^T |f_n(s, X^{n,t,x,\alpha}, \alpha_s) - f(s, X^{t,x,\alpha}, \alpha_s)| ds + |g_n(X^{n,t,x,\alpha}) - g(X^{t,x,\alpha})| \right] \\
& \leq \sup_{\alpha \in \mathcal{A}} \left\{ K(T+1) \|X^{n,t,x,\alpha} - X^{t,x,\alpha}\|_{\mathbf{S}_2} + \mathbb{E} \left[ \int_t^T |f_n(s, X^{t,x,\alpha}, \alpha_s) - f(s, X^{t,x,\alpha}, \alpha_s)| ds \right] \right. \\
& \quad \left. + \mathbb{E} [|g_n(X^{t,x,\alpha}) - g(X^{t,x,\alpha})|] \right\}.
\end{aligned}$$

By standard calculations (as for instance in [35, Theorem 2.5.9]), we find

$$\begin{aligned}
\|X^{n,t,x,\alpha} - X^{t,x,\alpha}\|_{\mathbf{S}_2}^2 & \leq C_K T e^{C_K T} \mathbb{E} \left[ \int_t^T |b_n(s, X^{t,x,\alpha}, \alpha_s) - b(s, X^{t,x,\alpha}, \alpha_s)|^2 ds \right. \\
& \quad \left. + \int_t^T |\sigma_n(s, X^{t,x,\alpha}, \alpha_s) - \sigma(s, X^{t,x,\alpha}, \alpha_s)|^2 ds \right],
\end{aligned}$$

for some constant  $C_K$ , depending only on the constant  $K$ . It remains to prove that

$$\begin{aligned}
& \sup_{\alpha \in \mathcal{A}} \left\{ \mathbb{E} \left[ \int_t^T |b_n(s, X^{t,x,\alpha}, \alpha_s) - b(s, X^{t,x,\alpha}, \alpha_s)|^2 ds \right] \right. \\
& + \mathbb{E} \left[ \int_t^T |\sigma_n(s, X^{t,x,\alpha}, \alpha_s) - \sigma(s, X^{t,x,\alpha}, \alpha_s)|^2 ds \right] \\
& \left. + \mathbb{E} \left[ \int_t^T |f_n(s, X^{t,x,\alpha}, \alpha_s) - f(s, X^{t,x,\alpha}, \alpha_s)| ds \right] + \mathbb{E} [|g_n(X^{t,x,\alpha}) - g(X^{t,x,\alpha})|] \right\} \xrightarrow{n \rightarrow +\infty} 0.
\end{aligned}$$

Let us address the term with  $f$  and  $f_n$ , the other three terms can be treated in a similar way. From the proof of Lemma B.3, and in particular from estimate (B.14) with  $h$  and  $h_n$  replaced respectively by  $f$  and  $f_n$ , we get

$$\begin{aligned}
& \mathbb{E} \left[ \int_t^T |f_n(s, X^{t,x,\alpha}, \alpha_s) - f(s, X^{t,x,\alpha}, \alpha_s)| ds \right] \\
& \leq 3KT \mathbb{E} \left[ \sup_{|r-s| \leq 2^{-n}} |X_r^{t,x,\alpha} - X_s^{t,x,\alpha}| \right] + KT \frac{(d(2^n+1)-1)!!}{(d(2^n+1))!!} \tag{B.18} \\
& + \int_t^T \left( \int_0^{n(T-s)} \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2} w(r/n) dr \right) ds + \int_t^T w(T-s) \left( \int_{n(T-s)}^\infty \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2}r^2} dr \right) ds,
\end{aligned}$$

where we recall that  $w$  is the modulus of continuity of  $f$  with respect to the time variable. By Lebesgue's dominated convergence theorem, we see that the last two integrals in (B.18) go to zero as  $n \rightarrow +\infty$ . Moreover, we recall that from the proof of Lemma B.3 that by Stirling's formula we get  $\frac{(d(2^n+1)-1)!!}{(d(2^n+1))!!} = O(1/\sqrt{d(2^n+1)})$  as  $n \rightarrow \infty$ , therefore the second

term on the right-hand side of (B.18) also goes to zero. Finally, by standard calculations, we get

$$\mathbb{E} \left[ \sup_{|r-s| \leq 2^{-n}} |X_r^{t,x,\alpha} - X_s^{t,x,\alpha}|^2 \right] \leq C_{K,T} 2^{-n} (1 + \|x\|_t^2),$$

for some constant  $C_{K,T} \geq 0$ , depending only on  $K$  and  $T$ . This allows to prove that right-hand side of (B.18) goes to zero as  $n \rightarrow +\infty$  and concludes the proof.  $\square$

**Lemma B.5.** *Let Assumptions (A) and (B) hold. Suppose also that there exist  $\hat{d} \in \mathbb{N}$  and functions*

$$\bar{b}, \bar{\sigma}, \bar{f}: [0, T] \times \mathbb{R}^{\hat{d}} \times A \longrightarrow \mathbb{R}^{\hat{d}}, \mathbb{R}^{\hat{d} \times m}, \mathbb{R}, \quad \bar{g}: \mathbb{R}^{\hat{d}} \longrightarrow \mathbb{R},$$

satisfying the following conditions.

i) *There exist some continuously differentiable functions  $\phi_1, \dots, \phi_{\hat{d}}: [0, T] \rightarrow \mathbb{R}$  such that:*

$$\begin{aligned} b(t, x, a) &= \bar{b} \left( t, \int_{[0,t]} \phi_1(s) d^-x(s), \dots, \int_{[0,t]} \phi_{\hat{d}}(s) d^-x(s), a \right), \\ \sigma(t, x, a) &= \bar{\sigma} \left( t, \int_{[0,t]} \phi_1(s) d^-x(s), \dots, \int_{[0,t]} \phi_{\hat{d}}(s) d^-x(s), a \right), \\ f(t, x, a) &= \bar{f} \left( t, \int_{[0,t]} \phi_1(s) d^-x(s), \dots, \int_{[0,t]} \phi_{\hat{d}}(s) d^-x(s), a \right), \\ g(x) &= \bar{g} \left( \int_{[0,T]} \phi_1(s) d^-x(s), \dots, \int_{[0,T]} \phi_{\hat{d}}(s) d^-x(s) \right), \end{aligned}$$

for every  $(t, x, a) \in [0, T] \times C([0, T]; \mathbb{R}^{\hat{d}}) \times A$ .

ii) *The functions  $\bar{b}, \bar{\sigma}, \bar{f}, \bar{g}$  are continuous.*

iii) *There exist a constant  $\bar{K} \geq 0$  such that*

$$\begin{aligned} &|\bar{b}(t, y, a) - \bar{b}(t, y', a)| + |\bar{\sigma}(t, y, a) - \bar{\sigma}(t, y', a)| + \\ &\quad + |\bar{f}(t, y, a) - \bar{f}(t, y', a)| + |\bar{g}(y) - \bar{g}(y')| \leq \bar{K} |y - y'|, \\ &|\bar{b}(t, 0, a)| + |\bar{\sigma}(t, 0, a)| + |\bar{f}(t, y, a)| + |\bar{g}(y)| \leq \bar{K}, \end{aligned}$$

for all  $(t, a) \in [0, T] \times A$ ,  $y, y' \in \mathbb{R}^{\hat{d}}$ .

iv) *For every  $a \in A$ , the functions  $\bar{b}(\cdot, \cdot, a), \bar{\sigma}(\cdot, \cdot, a), \bar{f}(\cdot, \cdot, a), \bar{g}(\cdot)$  are  $C^{1,2}([0, T] \times \mathbb{R}^{\hat{d}})$ . Moreover, there exist constants  $\bar{K} \geq 0$  and  $\bar{q} \geq 0$  such that*

$$\begin{aligned} &|\partial_t \bar{b}(t, y, a)| + |\partial_t \bar{\sigma}(t, y, a)| + |\partial_t \bar{f}(t, y, a)| \leq \bar{K} (1 + |y|)^{\bar{q}}, \\ &|\partial_y \bar{b}(t, y, a)| + |\partial_y \bar{\sigma}(t, y, a)| + |\partial_y \bar{f}(t, y, a)| + |\partial_y \bar{g}(y)| \leq \bar{K} (1 + |y|)^{\bar{q}}, \\ &|\partial_{yy}^2 \bar{b}(t, y, a)| + |\partial_{yy}^2 \bar{\sigma}(t, y, a)| + |\partial_{yy}^2 \bar{f}(t, y, a)| + |\partial_{yy}^2 \bar{g}(y)| \leq \bar{K} (1 + |y|)^{\bar{q}}, \end{aligned}$$

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

Then, for every  $\varepsilon \in (0, 1)$ , there exist  $v_\varepsilon: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  and  $\bar{v}_\varepsilon: [0, T] \times \mathbb{R}^{d\hat{d}} \rightarrow \mathbb{R}$ , with

$$v_\varepsilon(t, x) = \bar{v}_\varepsilon\left(t, \int_{[0,t]} \phi_1(s) d^-x(s), \dots, \int_{[0,t]} \phi_{\hat{d}}(s) d^-x(s)\right),$$

for all  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , such that the following holds.

- 1)  $v_\varepsilon \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  and  $\bar{v}_\varepsilon \in C^{1,2}([0, T] \times \mathbb{R}^{d\hat{d}})$ .
- 2)  $v_\varepsilon$  is a classical solution of the following equation:

$$\begin{cases} \partial_t^H v_\varepsilon(t, x) + \frac{1}{2} \varepsilon^2 \text{tr}[\partial_{yy} \bar{v}_\varepsilon(t, y^{t,x})] + \sup_{a \in A} \left\{ \langle b(t, x, a), \partial_x^V v_\varepsilon(t, x) \rangle \right. \\ \left. + \frac{1}{2} \text{tr}[(\sigma \sigma^\top)(t, x, a) \partial_{xx}^V v_\varepsilon(t, x)] + f(t, x, a) \right\} = 0, & (t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d), \\ v_\varepsilon(T, x) = g(x), & x \in C([0, T]; \mathbb{R}^d), \end{cases} \quad (\text{B.19})$$

where, for every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ ,  $y^{t,x} \in \mathbb{R}^{d\hat{d}}$  is defined as

$$y^{t,x} := \left( \int_{[0,t]} \phi_1(s) d^-x(s), \dots, \int_{[0,t]} \phi_{\hat{d}}(s) d^-x(s) \right). \quad (\text{B.20})$$

- 3) There exist constants  $\hat{C} = \hat{C}(\bar{K}, \bar{q}, \hat{d}, \varepsilon) \geq 0$  and  $\hat{q} = \hat{q}(\bar{q}) \geq 0$ , with  $\hat{C}$  depending on  $\bar{K}, \bar{q}, \hat{d}, \varepsilon$  and  $\hat{q}$  depending on  $\bar{q}$ , such that

$$|\partial_x^V v_\varepsilon(t, x)| + |\partial_{xx}^V v_\varepsilon(t, x)| \leq \hat{C} (1 + \|x\|_t)^{\hat{q}}, \quad (\text{B.21})$$

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

- 4) Finally,  $v_\varepsilon$  converges pointwise to  $v$  in (2.5) as  $\varepsilon \rightarrow 0^+$ .

**Proof.** Let  $\phi: [0, T] \rightarrow \mathbb{R}^{(d\hat{d}) \times d}$  be given by

$$\phi(t) = \begin{bmatrix} \phi_1(t) I_d \\ \vdots \\ \phi_{\hat{d}}(t) I_d \end{bmatrix}$$

for all  $t \in [0, T]$ , where  $I_d$  denotes the  $d \times d$  identity matrix. Let  $\bar{b}_\phi: [0, T] \times \mathbb{R}^{d\hat{d}} \times A \rightarrow \mathbb{R}^{d\hat{d}}$  and  $\bar{\sigma}_\phi: [0, T] \times \mathbb{R}^{d\hat{d}} \times A \rightarrow \mathbb{R}^{(d\hat{d}) \times m}$  be given by

$$\bar{b}_\phi(t, y, a) = \phi(t) \bar{b}(t, y, a), \quad \bar{\sigma}_\phi(t, y, a) = \phi(t) \bar{\sigma}(t, y, a),$$

for all  $(t, y, a) \in [0, T] \times \mathbb{R}^{d\hat{d}} \times A$ , with  $\bar{b}(t, y, a)$  being a column vector of dimension  $d$ . For every  $\varepsilon \in (0, 1)$ , consider the following Hamilton-Jacobi-Bellman equation on  $[0, T] \times \mathbb{R}^{d\hat{d}}$ :

$$\begin{cases} \partial_t \bar{u}(t, y) + \sup_{a \in A} \left\{ \langle \bar{b}_\phi(t, y, a), \partial_y \bar{u}(t, y) \rangle + \frac{1}{2} \varepsilon^2 \text{tr} [\partial_{yy}^2 \bar{u}(t, y)] \right. \\ \left. + \frac{1}{2} \text{tr} [(\bar{\sigma}_\phi \bar{\sigma}_\phi^\top)(t, y, a) \partial_{yy}^2 \bar{u}(t, y)] + \bar{f}(t, y, a) \right\} = 0, & (t, y) \in [0, T] \times \mathbb{R}^{d\hat{d}}, \\ \bar{u}(T, y) = \bar{g}(y), & y \in \mathbb{R}^{d\hat{d}}. \end{cases} \quad (\text{B.22})$$

*Proof of item 3).* From the assumptions on  $\bar{b}, \bar{\sigma}, \bar{f}, \bar{g}$ , it follows that there exists a unique classical solution  $\bar{v}_\varepsilon \in C^{1,2}([0, T] \times \mathbb{R}^{d\hat{d}})$  of equation (B.22) (see for instance [36, Theorem 14.15] and, in particular, the comments after Theorem 14.15 concerning the case when the operators “ $L_\nu$ ” are linear). Moreover, by [35, Theorems 4.1.1 and 4.7.4] we have that there exist constants  $\bar{C}_1 = \bar{C}_1(\hat{K}) \geq 0$  and  $\bar{C}_2 = \bar{C}_2(\bar{K}, \bar{q}) \geq 0$ , with  $\bar{C}_1$  depending only on  $\hat{K}$  and  $\bar{C}_2$  depending only on  $\bar{K}, \bar{q}$ , such that

$$|\partial_y \bar{v}_\varepsilon(t, y)| \leq \bar{C}_1, \quad (\text{B.23})$$

$$|\partial_t \bar{v}_\varepsilon(t, y)| \leq \bar{C}_2 e^{\bar{C}_2(T-t)} (1 + |y|)^{3\bar{q}}, \quad (\text{B.24})$$

$$-\bar{C}_2 e^{\bar{C}_2(T-t)} (1 + |y|)^{3\bar{q}} \leq \partial_{y_i y_j}^2 \bar{v}_\varepsilon(t, y) \leq \frac{1}{\varepsilon^2} \bar{C}_2 e^{\bar{C}_2(T-t)} (1 + |y|)^{3\bar{q}+2}, \quad (\text{B.25})$$

for all  $(t, y) \in [0, T] \times \mathbb{R}^{d\hat{d}}$  and every  $i, j = 1, \dots, d\hat{d}$ . Furthermore, by [35, Theorem 4.6.2] there exists some constant  $\hat{C} = \hat{C}(\hat{K}) \geq 0$ , depending only on  $\hat{K}$ , such that

$$|\bar{v}_\varepsilon(t, y) - \bar{v}(t, y)| \leq \varepsilon \hat{C} e^{\hat{C}(T-t)}, \quad (\text{B.26})$$

for every  $(t, y) \in [0, T] \times \mathbb{R}^{d\hat{d}}$ , where  $\bar{v}: [0, T] \times \mathbb{R}^{d\hat{d}} \rightarrow \mathbb{R}$  is defined as

$$\bar{v}(t, y) = \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \int_t^T \bar{f}(s, Y_s^{t,y,\alpha}, \alpha_s) ds + \bar{g}(Y_T^{t,y,\alpha}) \right],$$

with  $Y^{t,y,\alpha} = (Y_s^{t,y,\alpha})_{s \in [t, T]}$  solving the following system of controlled stochastic differential equations:

$$\begin{cases} dY_s^{t,y,\alpha} = \bar{b}_\phi(s, Y_s^{t,y,\alpha}, \alpha_s) ds + \bar{\sigma}_\phi(s, Y_s^{t,y,\alpha}, \alpha_s) dB_s, & s \in (t, T], \\ Y_t^{t,y,\alpha} = y. \end{cases}$$

*Proof of item 1).* For every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , consider  $y^{t,x} \in \mathbb{R}^{d\hat{d}}$  given by (B.20). Then, proceeding as in the proof of [12, Theorem 3.15] (see, in particular, equalities (3.16)), we obtain

$$\begin{aligned} Y_r^{t,y^{t,x},\alpha} &= \left( \int_{[0,t]} \phi_1(s) d^- x(s) + \int_t^r \phi_1(s) dX_s^{t,x,\alpha}, \dots, \int_{[0,t]} \phi_{\hat{d}}(s) d^- x(s) + \int_t^r \phi_{\hat{d}}(s) dX_s^{t,x,\alpha} \right) \\ &= \left( \int_{[0,r]} \phi_1(s) d^- X_s^{t,x,\alpha}, \dots, \int_{[0,r]} \phi_{\hat{d}}(s) d^- X_s^{t,x,\alpha} \right), \end{aligned} \quad (\text{B.27})$$

for all  $r \in [t, T]$ ,  $\mathbb{P}$ -a.s., where, for each  $i = 1, \dots, d$ ,  $\int_{[0,r]} \phi_i(s) d^- X_s^{t,x,\alpha}$  is intended  $\mathbb{P}$ -a.s. as a deterministic forward integral. From (B.27) we get

$$\begin{aligned} \bar{v} \left( t, \int_{[0,t]} \phi_1(s) d^- x(s), \dots, \int_{[0,t]} \phi_d(s) d^- x(s) \right) &= \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \int_t^T \bar{f}(s, Y_s^{t,y^{t,x},\alpha}, \alpha_s) ds + \bar{g}(Y_T^{t,y^{t,x},\alpha}) \right] \\ &= \sup_{\alpha \in \mathcal{A}} \mathbb{E} \left[ \int_t^T f(s, X_s^{t,x,\alpha}, \alpha_s) ds + g(X_T^{t,x,\alpha}) \right] = v(t, x), \end{aligned}$$

where  $v$  is the value function defined in (2.5).

Now, let  $v_\varepsilon: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  be defined as

$$v_\varepsilon(t, x) := \bar{v}_\varepsilon \left( t, \int_{[0,t]} \phi_1(s) d^- x(s), \dots, \int_{[0,t]} \phi_d(s) d^- x(s) \right),$$

for every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ . Then, by direct calculations (proceeding as in the proof of [11, Lemma D.1]), we deduce that  $v_\varepsilon \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  and that the following equalities hold:

$$\partial_t^H v_\varepsilon(t, x) = \partial_t \bar{v}_\varepsilon \left( t, \int_{[0,t]} \phi_1(s) d^- x(s), \dots, \int_{[0,t]} \phi_d(s) d^- x(s) \right), \quad (\text{B.28})$$

$$\partial_{x_i}^V v_\varepsilon(t, x) = \left\langle \partial_y \bar{v}_\varepsilon \left( t, \int_{[0,t]} \phi_1(s) d^- x(s), \dots, \int_{[0,t]} \phi_d(s) d^- x(s) \right), \phi_i(t) \right\rangle, \quad (\text{B.29})$$

$$\partial_{xx}^V v_\varepsilon(t, x) = \phi^\top(t) \partial_{yy}^2 \bar{v}_\varepsilon \left( t, \int_{[0,t]} \phi_1(s) d^- x(s), \dots, \int_{[0,t]} \phi_d(s) d^- x(s) \right) \phi(t), \quad (\text{B.30})$$

for every  $i = 1, \dots, d$ , where  $\phi_i(t)$  denotes the  $i$ -th column of the matrix  $\phi(t)$ .

*Proof of items 2) and 3).* Since  $\bar{v}_\varepsilon$  is a classical solution of equation (B.22), it follows from equalities (B.28)-(B.29)-(B.30) that  $v_\varepsilon$  is a classical solution of equation (B.19). In addition, from equalities (B.28)-(B.29)-(B.30) and estimates (B.23)-(B.24)-(B.25) we deduce that (B.21) holds.

*Proof of item 4).* From (B.26), we have

$$|v_\varepsilon(t, x) - v(t, x)| = |\bar{v}_\varepsilon(t, y^{t,x}) - \bar{v}(t, y^{t,x})| \leq \varepsilon \hat{C} e^{\hat{C}(T-t)},$$

for every  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ . This shows the validity of item 4).  $\square$

We can now state the following result, which plays a crucial role in the proof the comparison theorem (Theorem 4.5), in order to show that  $u_1 \leq v$ .

**Theorem B.6.** *Let Assumptions (A) and (B) hold. Consider the sequences  $\{b_n\}_n, \{\sigma_n\}_n, \{f_n\}_n, \{g_n\}_n, \{v_n\}_n$  in (B.16)-(B.17) (recall from Lemma B.3 that  $d_n$  and  $\phi_{n,1}, \dots, \phi_{n,d_n}$  are the same for  $b, \sigma, f, g$ ). Then, for every  $n$  and any  $\varepsilon \in (0, 1)$ , there exist  $v_{n,\varepsilon}: [0, T] \times C([0, T]; \mathbb{R}^d) \rightarrow \mathbb{R}$  and  $\bar{v}_{n,\varepsilon}: [0, T] \times \mathbb{R}^{d_n} \rightarrow \mathbb{R}$ , with*

$$v_{n,\varepsilon}(t, x) = \bar{v}_{n,\varepsilon} \left( t, \int_{[0,t]} \phi_{n,1}(s) d^- x(s), \dots, \int_{[0,t]} \phi_{n,d_n}(s) d^- x(s) \right),$$

for all  $(t, x) \in [0, T] \times C([0, T]; \mathbb{R}^d)$ , such that the following holds.

- 1)  $v_{n,\varepsilon} \in C_{\text{pol}}^{1,2}([0, T] \times C([0, T]; \mathbb{R}^d))$  and  $\bar{v}_{n,\varepsilon} \in C^{1,2}([0, T] \times \mathbb{R}^{dd_n})$ .
- 2)  $v_{n,\varepsilon}$  is a classical solution of equation (B.19) with  $b, \sigma, f, g, \bar{v}_\varepsilon, y^{t,x}$  replaced respectively by  $b_n, \sigma_n, f_n, g_n, \bar{v}_{n,\varepsilon}, y_n^{t,x}$ , where  $y_n^{t,x}$  is given by (B.7).
- 3) There exist constants  $C_{n,\varepsilon} \geq 0$  and  $q \geq 0$ , with  $C_{n,\varepsilon}$  depending on  $n, \varepsilon, K, d, T$ , such that

$$|\partial_x^V v_{n,\varepsilon}(t, x)| + |\partial_{xx}^V v_{n,\varepsilon}(t, x)| \leq C_{n,\varepsilon} (1 + \|x\|_t)^q, \quad (\text{B.31})$$

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

- 4) For every  $n$ ,  $v_{n,\varepsilon}$  converges pointwise to  $v_n$  in (B.17) as  $\varepsilon \rightarrow 0^+$ .
- 5)  $v_n$  converges pointwise to  $v$  in (2.5) as  $n \rightarrow +\infty$ .

**Proof.** Items 1)-2)-3)-4) follow directly from Lemma B.5 with  $b, \sigma, f, g$  replaced respectively by  $b_n, \sigma_n, f_n, g_n$ . In particular, concerning the constant  $C_{n,\varepsilon}$  appearing in items 3), it can also depend on  $d$  as a consequence of the fact that  $b_n, \sigma_n, \bar{b}_n, \bar{\sigma}_n$  satisfy a linear growth condition with a constant  $K_d$ , depending on  $K$  and also on  $d$ . Finally, item 5) follows from Lemma B.4.  $\square$

## References

- [1] J.-P. Aubin and G. Haddad. History path dependent optimal control and portfolio valuation and management. *Positivity*, 6(3):331–358, 2002. Special issue of the mathematical economics.
- [2] A. Barrasso and F. Russo. Decoupled mild solutions of path-dependent PDEs and integro PDEs represented by BSDEs driven by cadlag martingales. *Potential Anal.*, 53(2):449–481, 2020.
- [3] E. Bayraktar and C. Keller. Path-dependent Hamilton-Jacobi equations in infinite dimensions. *J. Funct. Anal.*, 275(8):2096–2161, 2018.
- [4] J. M. Borwein and D. Preiss. A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions. *Trans. Amer. Math. Soc.*, 303(2):517–527, 1987.
- [5] J. M. Borwein and Q. J. Zhu. *Techniques of variational analysis*, volume 20 of *CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC*. Springer-Verlag, New York, 2005.
- [6] R. Cont and D.-A. Fournié. Change of variable formulas for non-anticipative functionals on path space. *J. Funct. Anal.*, 259(4):1043–1072, 2010.
- [7] R. Cont and D.-A. Fournié. Functional Itô calculus and stochastic integral representation of martingales. *Ann. Probab.*, 41(1):109–133, 2013.
- [8] R. Cont and D.-A. Fournié. Functional Itô calculus and functional Kolmogorov equations. In *Stochastic integration by parts and functional Itô calculus*, Adv. Courses Math. CRM Barcelona, pages 115–207. Birkhäuser/Springer, [Cham], 2016.



- [9] A. Cosso, S. Federico, F. Gozzi, M. Rosestolato, and N. Touzi. Path-dependent equations and viscosity solutions in infinite dimension. *Ann. Probab.*, 46(1):126–174, 2018.
- [10] A. Cosso, I. Kharroubi, F. Gozzi, H. Pham, and M. Rosestolato. Optimal control of path-dependent McKean-Vlasov SDEs in infinite dimension. *Preprint arXiv:2012.14772*, 2020.
- [11] A. Cosso and F. Russo. Crandall-Lions viscosity solutions for path-dependent PDEs: The case of heat equation. To appear in *Bernoulli*. DOI:10.3150/21-BEJ1353. Preprint arXiv:1911.13095.
- [12] A. Cosso and F. Russo. Functional Itô versus Banach space stochastic calculus and strict solutions of semilinear path-dependent equations. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 19(4):1650024, 44, 2016.
- [13] A. Cosso and F. Russo. Strong-viscosity solutions: classical and path-dependent PDEs. *Osaka J. Math.*, 56(2):323–373, 2019.
- [14] M. G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.
- [15] M. G. Crandall and P.-L. Lions. Condition d’unicité pour les solutions généralisées des équations de Hamilton-Jacobi du premier ordre. *C. R. Acad. Sci. Paris Sér. I Math.*, 292(3):183–186, 1981.
- [16] M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277(1):1–42, 1983.
- [17] G. Da Prato and J. Zabczyk. *Second Order Partial Differential Equations in Hilbert Spaces*. London Mathematical Society Lecture Note Series. Cambridge University Press, 2002.
- [18] C. Di Girolami, G. Fabbri, and F. Russo. The covariation for Banach space valued processes and applications. *Metrika*, 77(1):51–104, 2014.
- [19] C. Di Girolami and F. Russo. Infinite dimensional stochastic calculus via regularization and applications. *Preprint HAL-INRIA*, inria-00473947 version 1, (Unpublished), 2010.
- [20] C. Di Girolami and F. Russo. About classical solutions of the path-dependent heat equation. *Random Oper. Stoch. Equ.*, 28(1):35–62, 2020.
- [21] B. Dupire. Functional Itô calculus. *Portfolio Research Paper*, Bloomberg, 2009.
- [22] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.
- [23] I. Ekren, C. Keller, N. Touzi, and J. Zhang. On viscosity solutions of path dependent PDEs. *Ann. Probab.*, 42(1):204–236, 2014.
- [24] I. Ekren, N. Touzi, and J. Zhang. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I. *Ann. Probab.*, 44(2):1212–1253, 2016.
- [25] I. Ekren, N. Touzi, and J. Zhang. Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II. *Ann. Probab.*, 44(4):2507–2553, 2016.

- [26] N. El Karoui and X. Tan. Capacities, Measurable Selection and Dynamic Programming Principle. Part I: Abstract Framework. *Preprint arXiv:1310.3363*, 2013.
- [27] N. El Karoui and X. Tan. Capacities, Measurable Selection and Dynamic Programming Principle. Part II: Application in Stochastic Control Problems. *Preprint arXiv:1310.3364*, 2013.
- [28] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.
- [29] G. Fabbri, F. Gozzi, and A. Swiech. *Stochastic optimal control in infinite dimension: dynamic programming and HJB equations, with a contribution by M. Fuhrman and G. Tessitore*, volume 82 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2017.
- [30] S. Federico. A pension fund in the accumulation phase: a stochastic control approach. In *Advances in mathematics of finance*, volume 83 of *Banach Center Publ.*, pages 61–83. Polish Acad. Sci. Inst. Math., Warsaw, 2008.
- [31] S. Federico. A stochastic control problem with delay arising in a pension fund model. *Finance Stoch.*, 15(3):421–459, 2011.
- [32] F. Flandoli and G. Zanco. An infinite-dimensional approach to path-dependent Kolmogorov equations. *Ann. Probab.*, 44(4):2643–2693, 2016.
- [33] M. I. Gomoyunov, N. Y. Lukoyanov, and A. R. Plaskin. Path-dependent Hamilton-Jacobi equations: the minimax solutions revised. *Appl. Math. Optim.*, 83(3):2327–2374, 2021.
- [34] O. Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.
- [35] N. V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2009. Translated from the 1977 Russian original by A. B. Aries, Reprint of the 1980 edition.
- [36] G. M. Lieberman. *Second order parabolic differential equations*. World Scientific Publishing Co., Inc., River Edge, NJ, 1996.
- [37] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. II. Viscosity solutions and uniqueness. *Comm. Partial Differential Equations*, 8(11):1229–1276, 1983.
- [38] N. Y. Lukoyanov. Viscosity solution of nonanticipating equations of Hamilton-Jacobi type. *Differ. Equ.*, 43(12):1715–1723, 2007.
- [39] S. Peng and Y. Song.  $G$ -expectation weighted Sobolev spaces, backward SDE and path dependent PDE. *J. Math. Soc. Japan*, 67(4):1725–1757, 2015.
- [40] S. Peng and F. Wang. BSDE, path-dependent PDE and nonlinear Feynman-Kac formula. *Sci. China Math.*, 59(1):19–36, 2016.
- [41] Z. Ren. Viscosity solutions of fully nonlinear elliptic path dependent partial differential equations. *Ann. Appl. Probab.*, 26(6):3381–3414, 2016.

- [42] Z. Ren and M. Rosestolato. Viscosity solutions of path-dependent PDEs with randomized time. *SIAM J. Math. Anal.*, 52(2):1943–1979, 2020.
- [43] Z. Ren, N. Touzi, and J. Zhang. Comparison of viscosity solutions of fully nonlinear degenerate parabolic path-dependent PDEs. *SIAM J. Math. Anal.*, 49(5):4093–4116, 2017.
- [44] Z. Ren, N. Touzi, and J. Zhang. Comparison of viscosity solutions of semilinear path-dependent PDEs. *SIAM J. Control Optim.*, 58(1):277–302, 2020.
- [45] M. Rosestolato and A. Swiech. Partial regularity of viscosity solutions for a class of kolmogorov equations arising from mathematical finance. *Journal of Differential Equations*, 262(3):1897–1930, 2017.
- [46] S. Tang and F. Zhang. Path-dependent optimal stochastic control and viscosity solution of associated Bellman equations. *Discrete Contin. Dyn. Syst.*, 35(11):5521–5553, 2015.
- [47] J. Zhou. Viscosity solutions to second order path-dependent Hamilton-Jacobi-Bellman equations and applications. *Preprint arXiv:2005.05309v1*, 2020.