Crandall-Lions Viscosity Solutions for Path-Dependent PDEs: The Case of Heat Equation

Andrea Cosso and Francesco Russo

Abstract. We address our interest to the development of a theory of viscosity solutions à la Crandall-Lions for path-dependent partial differential equations (PDEs), namely PDEs in the space of continuous paths $C([0,T];\mathbb{R}^d)$. Path-dependent PDEs can play a central role in the study of certain classes of optimal control problems, as for instance optimal control problems with delay. Typically, they do not admit a smooth solution satisfying the corresponding HJB equation in a classical sense, it is therefore natural to search for a weaker notion of solution. While other notions of generalized solution have been proposed in the literature, the extension of the Crandall-Lions framework to the path-dependent setting is still an open problem. The question of uniqueness of the solutions, which is the more delicate issue, will be based on early ideas from the theory of viscosity solutions and a suitable variant of Ekeland's variational principle. This latter is based on the construction of a smooth gauge-type function, where smooth is meant in the horizontal/vertical (rather than Fréchet) sense. In order to make the presentation more readable, we address the path-dependent heat equation, which in particular simplifies the smoothing of its natural "candidate" solution. Finally, concerning the existence part, we provide a new proof of the functional Itô formula under general assumptions, extending earlier results in the literature.

1. Introduction

Path-dependent heat equation refers to the following second-order partial differential equation in the space of continuous paths:

$$\begin{cases} -\partial_t^H v(t,\boldsymbol{x}) - \frac{1}{2}\mathrm{tr}\big[\partial_{\boldsymbol{x}\boldsymbol{x}}^V v(t,\boldsymbol{x})\big] \;=\; 0, & (t,\boldsymbol{x}) \in [0,T) \times C([0,T];\mathbb{R}^d), \\ v(T,\boldsymbol{x}) \;=\; \xi(\boldsymbol{x}), & \boldsymbol{x} \in C([0,T];\mathbb{R}^d). \end{cases} \tag{1.1}$$

Here $C([0,T];\mathbb{R}^d)$ denotes the Banach space of continuous paths $\boldsymbol{x}\colon [0,T] \to \mathbb{R}^d$ equipped with the supremum norm $\|\boldsymbol{x}\|_\infty := \sup_{t\in[0,T]} |\boldsymbol{x}(t)|$, with $|\cdot|$ denoting the Euclidean norm on $\mathbb{R}^d$. The terminal condition $\xi\colon C([0,T];\mathbb{R}^d) \to \mathbb{R}$ is assumed to be continuous and bounded. We refer to equation (1.1) as path-dependent *heat* equation. Similarly as for the usual heat equation, it admits the following Feynman-Kac representation formula in terms of the d-dimensional Brownian motion $\boldsymbol{W} = (\boldsymbol{W}_s)_{s\in[0,T]}$.

$$v(t,\boldsymbol{x}) \;=\; \mathbb{E}\big[\xi(\boldsymbol{W}^{t,\boldsymbol{x}})\big], \qquad \forall\,(t,\boldsymbol{x}) \in [0,T] \times C([0,T];\mathbb{R}^d), \tag{1.2}$$

where

$$\boldsymbol{W}_s^{t,\boldsymbol{x}} \;:=\; \begin{cases} \boldsymbol{x}(s), & s \leq t, \\ \boldsymbol{x}(t) + \boldsymbol{W}_s - \boldsymbol{W}_t, & s > t. \end{cases}$$

In the case of the classical heat equation ξ only depends on the terminal value $\boldsymbol{W}_T^{t,\boldsymbol{x}}$.

Date: November 26th 2019.

2010 *Mathematics Subject Classification.* 35D40, 35R15, 60H30.

Key words and phrases. Path-dependent partial differential equations; viscosity solutions; functional Itô formula.

The peculiarity of equation (1.1) is the presence of the so-called functional or pathwise derivatives $\partial_t^H v$, $\partial_{\boldsymbol{xx}}^V v$, where $\partial_t^H v$ is known as horizontal derivative, while $\partial_{\boldsymbol{xx}}^V v$ is the matrix of second-order vertical derivatives. Those derivatives appeared in [54, 55] (under the name of coinvariant derivatives) as building block of the so-called i-smooth analysis, and independently in [1], where they were denoted Clio derivatives; later, they were rediscovered by [30] (from which we borrow terminology and definitions), who adopted a slightly different definition based on the space of càdlàg paths and in addition developed a related stochastic calculus, known as functional Itô calculus, including in particular the so-called functional Itô formula. Differently from the classical Fréchet derivative on $C([0,T];\mathbb{R}^d)$, the distinguished features of the pathwise derivatives are their finite-dimensional nature and the property of being non-anticipative, which follow from the interpretation of t in $\boldsymbol{x}(t)$ as time variable. This means that $v(t,\boldsymbol{x})$ only depends on the values of the path $\boldsymbol{x}$ up to time t; moreover, the horizontal and vertical derivatives at time t are computed keeping the past values frozen, while only the present value of the path (that is $\boldsymbol{x}(t)$) can vary. The related functional Itô calculus was rigorously investigated in [11, 12, 13]. In [14, 17] we also gave a contribution in this direction, exploring the relation between pathwise derivatives and Banach space stochastic calculus, built on an appropriate notion of Fréchet type derivative and firstly conceived in [25], see also [26, 27, 28, 24].

Partial differential equations in the space of continuous paths (also known as functional or Clio or path-dependent partial differential equations) are mostly motivated by optimal control problems of deterministic and stochastic systems with delay (or path-dependence) in the state variable. Such control systems arise in many fields, as for instance optimal advertising theory [47, 48], chemical engineering [45], financial management [38, 72], economic growth theory [2], mean field game theory [5], biomedicine [46, 81], systemic risk [9]. The underlying deterministic or stochastic controlled differential equations with delay may be studied in two ways: first using a direct approach (see for instance [50, 84, 54, 51, 56]), second by lifting them into a suitable infinite-dimensional framework, leading to evolution equations in Hilbert (as in [10, 23, 41]) or Banach spaces (as in [69, 70, 25]). The latter methodology turned out to be preferable to address general optimal control problems with delay (see for instance [90, 49, 47, 39, 40, 44, 37]), although such an infinite-dimensional reformulation may require some additional artificial assumptions to be imposed on the original control problem. On the other hand, the direct approach was adopted for special problems where the Hamilton-Jacobi-Bellman equation reduces to a finite-dimensional differential equation, as in [35, 57]. This approach can now regain relevance thanks to a well-grounded theory of path-dependent partial differential equations. To this regard, the path-dependent heat equation represents the primary test for such a theory, it indeed requires the main building blocks of the methodology, without overloading the proofs with additional technicalities.

Path-dependent partial differential equations represent a quite recent area of research. Typically, they do not admit a smooth solution satisfying the equation in a classical sense, mainly because of the awkward nature of the underlying space $C([0,T];\mathbb{R}^d)$. This happens also for the path-dependent heat equation, which in particular does not have the smoothing effect characterizing the classical heat equation, except in some specific cases (as shown in [25, 29]) with ξ belonging to the class of so-called cylinder or tame functions (therefore depending specifically on a finite number of integrals with respect to the path) or ξ being smoothly Fréchet differentiable. It is indeed quite easy, relying on the probabilistic representation formula (1.2), to see that the function v is not smooth (in the horizontal/vertical sense mentioned above) for terminal conditions of the form

$$\xi(\boldsymbol{x}) \;=\; \sup_{0\leq t\leq T} \boldsymbol{x}(t), \qquad\qquad \xi(\boldsymbol{x}) \;=\; \boldsymbol{x}(t_0),$$

for some fixed $t_0 \in (0,T)$. For a detailed analysis of the first case above we refer to Section 3.2 in [16] (see also Remark 3.8 in [17]). It is however worth mentioning that some positive results on smooth solutions were obtained in [17, 76]. We also refer to Chapter 9 of [25] and [29], where smooth solutions were investigated using a Fréchet type derivative formulation.

It is therefore natural to search for a weaker notion of solution, as the notion of viscosity solution, commonly used in the standard finite-dimensional case. The theory of viscosity solutions, firstly introduced in [20, 21] for first-order equations in finite dimension and later extended to the second-order case in [59, 60, 61], provides a well-suited framework guaranteeing the desired existence, uniqueness, and stability properties (for a comprehensive account see [19]). The extension of such a theory to equations in infinite dimension was initiated by [22, 62, 63, 64, 85, 88]. One of the structural assumption is that the state space has to be a Hilbert space or, slightly more general, certain Banach space with smooth norm, not including for instance the Banach space $C([0,T];\mathbb{R}^d)$ (notice however that in this paper we do not directly generalize those results to $C([0,T];\mathbb{R}^d)$, as we adopt horizontal/vertical, rather than Fréchet, derivatives on $C([0,T];\mathbb{R}^d)$).

First-order path-dependent partial differential equations were deeply investigated in [68] using a viscosity type notion of solution, which differs from the Crandall-Lions definition as the maximum/minimum condition is formulated on the subset of absolutely continuous paths. Such a modification does not affect existence in the first-order case, however it is particularly convenient for uniqueness, which is indeed established under general conditions. Other notions of generalized solution designed for first-order equations were adopted in [1] as well as in [65, 66, 67], where the minimax framework introduced in [86, 87] was implemented. We also mention [4], where such a minimax approach was extended to first-order path-dependent Hamilton-Jacobi-Bellman equations in infinite dimension.

Concerning the second-order case, a first attempt to extend the Crandall-Lions framework to the path-dependent case was carried out in [74], even though a technical condition on the semijets was imposed, namely condition (16) in [74], which narrows down the applicability of such a result. In the literature, this was perceived as an almost insurmountable obstacle, so that the Crandall-Lions definition was not further investigated, while other notions of generalized solution were devised, see [32, 75, 89, 18, 58, 3, 8]. We mention in particular the framework designed in [32] and further investigated in [33, 34, 78, 79, 80, 15], where the notion of sub/supersolution adopted differs from the Crandall-Lions definition as the tangency condition is not pointwise but in the sense of expectation with respect to an appropriate class of probability measures. On the other hand, in [18] we introduced the so-called strong-viscosity solution, which is quite similar to the notion of good solution for partial differential equations in finite dimension, that in turn is known to be equivalent to the definition of L^p-viscosity solution, see for instance [52]. We also mention [3], where the authors deal with semilinear path-dependent equations and propose the notion of decoupled mild solution, formulated in terms of generalized transition semigroups; such a notion also adapts to path-dependent equations with integro-differential terms.

In the present paper we adopt the natural generalization of the well-known definition of viscosity solution à la Crandall-Lions given in terms of test functions and, under this notion, we establish existence and uniqueness for the path-dependent heat equation (1.1). The uniqueness property is derived, as usual, from the comparison theorem. The proof of this latter, which is the most delicate issue, is known to be quite involved even in the classical finite-dimensional case (see for instance [19]), and in its latest form is based on Ishii's lemma. Here we follow instead an earlier approach (see for instance Theorem II.1 in [61] or Theorem IV.1 in [62]), which in principle can be applied to any path-dependent equation admitting a "candidate" solution v, for which a probabilistic representation formula holds. This is the case for equation (1.1), where the candidate solution is given by formula (1.2), but it is also the case for Kolmogorov type equations or, more generally, for Hamilton-Jacobi-Bellman equations. This latter is the class of equations studied in [61] and [62], whose methodology in a nutshell can be described as follows. Let u (resp. w) be a viscosity subsolution (resp. supersolution) of the same path-dependent equation. The desired inequality $u \leq w$ follows if we compare both u and w to the "candidate" solution v, that is if we prove the two inequalities $u \leq v$ and $v \leq w$. Let us consider for instance the first inequality $u \leq v$. In the non-path-dependent and finite-dimensional case (as in [61]), this is proved proceeding as follows: firstly, performing a smoothing of v through its probabilistic representation formula; secondly, taking a local maximum of $u - v_n$ (here it is used the local compactness of the finite-dimensional underlying space), with v_n being a smooth approximation of v; finally, the inequality $u \leq v_n$ is proved proceeding as in the so-called partial comparison theorem (comparison between a viscosity subsolution/supersolution and a smooth supersolution/subsolution), namely exploiting the viscosity subsolution property of u with v_n playing the role of test function. In [62], where such a methodology was extended to the infinite-dimensional case, the existence of a maximum of $u - v_n$ is achieved relying on Ekeland's variational principle, namely exploiting the completeness of the space instead of the missing local compactness.

In this paper we generalize the methodology sketched above to the path-dependent case. There are however at least two crucial mathematical issues required by such a proof, still not at disposal in the path-dependent framework.

Firstly, given a candidate solution v, it is not a priori obvious how to perform a smooth approximation of v itself starting from its probabilistic representation formula. Here we exploit the results proved in [17] (Theorem 3.5) and [18] (Theorem 3.12), which are reported and adapted to the present framework in Appendix C (Lemma C.1 and Lemma C.2, respectively). Notice that such results apply to the case of the path-dependent heat equation (1.1), where there is only the terminal condition ξ in the probabilistic representation formula (1.2) for v. More general results are at disposal in [17] and [18], which cover the case of semilinear path-dependent partial differential equations, characterized by the presence of four coefficients b, σ, F, ξ (see, in particular, Theorem 3.16 in [18] for more details). However, when those other coefficients appear in the path-dependent partial differential equation, we need more information on the sequence $\{v_n\}_n$ approximating v. For instance, we also have to estimate the derivatives of v_n in order to proceed as in [61] or [62]. Since such results are still not at disposal in the path-dependent setting, in order to make the paper more readable and not excessively lengthy, here we address the case of the path-dependent heat equation.

Secondly, concerning the existence of a maximum of $u - v_n$, we rely on a generalized version of Ekeland's variational principle for which we need a smooth gauge-type function with bounded derivatives, as explained below. Our equation is in fact formulated on the non-locally compact space $[0,T] \times C([0,T];\mathbb{R}^d)$ endowed with the pseudometric

$$d_\infty\big((t,\boldsymbol{x}),(t',\boldsymbol{x}')\big) \;:=\; |t-t'| + \|\boldsymbol{x}(\cdot \wedge t) - \boldsymbol{x}'(\cdot \wedge t')\|_\infty.$$

Recall that Ekeland's variational principle, in its original form, applied to $([0,T]\times C([0,T];\mathbb{R}^d), d_\infty)$ states that a perturbation $u - v_n - \delta d_\infty((\cdot,\cdot),(\bar t,\bar{\boldsymbol{x}}))$ of $u - v_n$ has a strict global maximum, with the perturbation being expressed in terms of the distance d_∞ (the point $(\bar t,\bar{\boldsymbol{x}})$ is fixed). As the map $(t,\boldsymbol{x}) \mapsto d_\infty((t,\boldsymbol{x}),(\bar t,\bar{\boldsymbol{x}}))$ is not smooth, it cannot be a test function. In order to have a smooth map instead of d_∞, an extended version of Ekeland's variational principle is needed. It is known (see for instance [7]) that a generalization of the so-called Borwein-Preiss smooth variant of Ekeland's variational principle works when d_∞ is replaced by a so-called gauge-type function (see Definition 3.1). For the proof of the comparison theorem, we have to construct a gauge-type function which is also smooth and with bounded derivatives, recalling that *smooth* in the present context means in the horizontal/vertical (rather than in the Fréchet) sense. In Section 3 such a gauge-type function is built through a smoothing of d_∞ itself (more precisely, of the part concerning the supremum norm). This latter smoothing is performed by convolution, firstly in the vertical direction, that is in the direction of the map $1_{[t,T]}$ (Lemma 3.1), then in the horizontal direction (Lemmata 3.2 and 3.3), the ordering of smoothings being crucial. Notice in particular that the supremum norm is already smooth in the horizontal direction; however, after the vertical smoothing, we lose in general the horizontal regularity because of the presence of the term $1_{[t,T]}$; for this reason we have also to perform the horizontal smoothing. The resulting smooth gauge-type function with bounded derivatives corresponds to the function ρ_∞ defined in (3.20).

Regarding existence, we prove that the candidate solution v in (1.2) solves in the viscosity sense equation (1.1). We proceed essentially as in the classical non-path-dependent case, relying as usual on Itô's formula, which in the present context corresponds to the functional Itô formula. Such a formula was firstly stated in [30] and then rigorously proved in [11, 12], see also [13, 42, 17, 58, 71]. In the present paper we provide a new rigorous proof of the functional Itô formula under general assumptions (Theorem 2.2). In particular, we do not require any boundedness assumption on the functional $u\colon [0,T]\times C([0,T];\mathbb{R}^d)\to\mathbb{R}$, thus improving (when the semimartingale process is continuous) the results stated in [11, 12]. The functional Itô formula is proved following a similar approach as in [83], where the classical non-path-dependent case was studied.

Finally, in the present paper we define pathwise derivatives in an alternative and self-contained manner, as we are only interested in the case of continuous paths. Such an approach, developed in detail in Section 2, is somehow minimalist compared to what is usually done in the literature, where definitions require the space of càdlàg paths. More precisely, in order to define the so-called vertical derivatives for a map $u\colon [0,T]\times C([0,T];\mathbb{R}^d)\to\mathbb{R}$ we consider a lifting of u, that is a map $\hat u$ defined on the enlarged space $[0,T]\times C([0,T];\mathbb{R}^d)\times\mathbb{R}^d$. The third argument $\mathbf{y}$ in $\hat u(t,\boldsymbol{x},\mathbf{y})$ refers to the possible jump of $\boldsymbol{x}$ at the present time t. Indeed, $\hat u$ is a lifting of u if it holds that

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

Therefore, when $\mathbf{y}=\boldsymbol{x}(t)$ then there is no jump at time t, otherwise there is a jump of size $\mathbf{y}-\boldsymbol{x}(t)$. For this reason, we refer to the product space $C([0,T];\mathbb{R}^d)\times\mathbb{R}^d$ as the space of paths with at most one jump at the present time (concerning such a space see also [41]). As already mentioned, in the literature the product space $C([0,T];\mathbb{R}^d)\times\mathbb{R}^d$ is replaced by the space of càdlàg paths $D([0,T];\mathbb{R}^d)$.

The paper is organized as follows. Section 2 (together with Appendices A and B) is devoted to pathwise derivatives and functional Itô calculus. In particular, the proof of functional Itô formula is reported in Appendix A. In Section 3 we construct the smooth gauge-type function with bounded derivatives, which corresponds to the function ρ_∞ in (3.20). In Section 4 we provide the (path-dependent) Crandall-Lions definition of viscosity solution for a general path-dependent partial differential equation. We then study in detail the path-dependent heat equation. In particular, we prove existence showing that the so-called candidate solution v solves in the viscosity sense the path-dependent heat equation (Theorem 4.1). We conclude Section 4 proving the comparison theorem (Theorem 4.2), for which we need a smooth approximation of the candidate solution v (Appendix C) as well as a suitable generalization of Ekeland's variational principle on $([0,T]\times C([0,T];\mathbb{R}^d),d_\infty)$ (Appendix D).

2. Pathwise derivatives and functional Itô calculus

In the present section we define the pathwise derivatives and state the fundamental tool of functional Itô calculus, namely the functional Itô formula. We refer to [30, 12, 42, 13] for more details. However, as already mentioned in the Introduction, we define vertical derivatives without relying on the space of càdlàg paths (as it is generally done in the literature); on the contrary, we employ the product space $C([0,T];\mathbb{R}^d)\times\mathbb{R}^d$, which can be thought as the space of paths with at most one jump at the present time (Section 2.2).

2.1. **Maps on continuous paths.** Given $T>0$ and $d\in\mathbb{N}^*$, we denote by $C([0,T];\mathbb{R}^d)$ the set of continuous functions $\boldsymbol{x}\colon[0,T]\to\mathbb{R}^d$. We denote by $\boldsymbol{x}(t)$ the value of $\boldsymbol{x}$ at $t\in[0,T]$. We also denote by $\mathbf{0}$ the function $\boldsymbol{x}\colon[0,T]\to\mathbb{R}^d$ identically equal to zero. We consider on $C([0,T];\mathbb{R}^d)$ the supremum norm $\|\cdot\|_\infty$, namely $\|\boldsymbol{x}\|_\infty:=\sup_{t\in[0,T]}|\boldsymbol{x}(t)|$, where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^d$ (we use the same symbol $|\cdot|$ to denote the Euclidean norm on $\mathbb{R}^k$, for any $k\in\mathbb{N}$).

We set $\boldsymbol{\Lambda}:=[0,T]\times C([0,T];\mathbb{R}^d)$ and define $d_\infty\colon\boldsymbol{\Lambda}\times\boldsymbol{\Lambda}\to[0,\infty)$ as

$$d_\infty\big((t,\boldsymbol{x}),(t',\boldsymbol{x}')\big) \;:=\; |t-t'|+\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}'(\cdot\wedge t')\|_\infty.$$

Notice that d_∞ is a *pseudometric* on $\boldsymbol{\Lambda}$, that is d_∞ is not a true metric because one may have $d_\infty((t,\boldsymbol{x}),(t',\boldsymbol{x}'))=0$ even if $(t,\boldsymbol{x})\neq(t',\boldsymbol{x}')$. We recall that one can construct a true metric space $(\boldsymbol{\Lambda}^*,d_\infty^*)$, called the metric space induced by the pseudometric space $(\boldsymbol{\Lambda},d_\infty)$, by means of the equivalence relation which follows from the vanishing of the pseudometric. We also observe that $(\boldsymbol{\Lambda},d_\infty)$ is a complete pseudometric space. Finally, we denote by $\mathcal{B}(\boldsymbol{\Lambda})$ the Borel σ-algebra on $\boldsymbol{\Lambda}$ induced by d_∞.

Definition 2.1. A map (or functional) $u\colon\boldsymbol{\Lambda}\to\mathbb{R}$ is said to be **non-anticipative** (on $\boldsymbol{\Lambda}$) if it satisfies

$$u(t,\boldsymbol{x}) \;=\; u(t,\boldsymbol{x}(\cdot\wedge t)),$$

for all $(t,\boldsymbol{x})\in\boldsymbol{\Lambda}$.

Remark 2.1. Whenever $u\colon\boldsymbol{\Lambda}\to\mathbb{R}$ is Borel measurable, namely u is measurable with respect to $\mathcal{B}(\boldsymbol{\Lambda})$, then u is non-anticipative on $\boldsymbol{\Lambda}$.

Definition 2.2. We denote by $\boldsymbol{C}(\boldsymbol{\Lambda})$ the set of non-anticipative maps $u\colon\boldsymbol{\Lambda}\to\mathbb{R}$ which are continuous on $\boldsymbol{\Lambda}$ with respect to d_∞.

2.2. **Lifted maps and their pathwise derivatives.** In the sequel we consider the product space $\boldsymbol{\Lambda}\times\mathbb{R}^d$, endowed with the product topology induced by d_∞ on $\boldsymbol{\Lambda}$ and $|\cdot|$ on $\mathbb{R}^d$. We denote by $\mathcal{B}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$ the corresponding Borel σ-algebra. On the subset $[0,T)\times C([0,T];\mathbb{R}^d)\times\mathbb{R}^d$ of $\boldsymbol{\Lambda}\times\mathbb{R}^d$ we consider the corresponding subspace topology.

Definition 2.3. A map (or functional) $\hat u\colon\boldsymbol{\Lambda}\times\mathbb{R}^d\to\mathbb{R}$ is said to be **non-anticipative** (on $\boldsymbol{\Lambda}\times\mathbb{R}^d$) if it satisfies

$$\hat u(t,\boldsymbol{x},\mathbf{y}) \;=\; \hat u(t,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}),$$

for all $(t,\boldsymbol{x},\mathbf{y})\in\boldsymbol{\Lambda}\times\mathbb{R}^d$.

Remark 2.2. Whenever $\hat{u}\colon \mathbf{\Lambda}\times\mathbb{R}^d \to \mathbb{R}$ is Borel measurable, namely $\hat{u}$ is measurable with respect to $\mathcal{B}(\mathbf{\Lambda}\times\mathbb{R}^d)$, then $\hat{u}$ is non-anticipative on $\mathbf{\Lambda}\times\mathbb{R}^d$.

Definition 2.4. We denote by $\boldsymbol{C}(\mathbf{\Lambda}\times\mathbb{R}^d)$ the set of non-anticipative maps $\hat{u}\colon \mathbf{\Lambda}\times\mathbb{R}^d \to \mathbb{R}$ which are continuous on $\mathbf{\Lambda}\times\mathbb{R}^d$.

Definition 2.5 (**Pathwise derivatives**)**.** Let $\hat{u}\colon \mathbf{\Lambda}\times\mathbb{R}^d \to \mathbb{R}$.

(i) Given $(t,\boldsymbol{x}) \in [0,T)\times C([0,T];\mathbb{R}^d)$, the **horizontal derivative** of $\hat{u}$ at $(t,\boldsymbol{x})$ (if the corresponding limit exists) is defined as

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

When the above limit exists for every $(t,\boldsymbol{x}) \in [0,T)\times C([0,T];\mathbb{R}^d)$, then $\partial_t^H\hat{u}$ is a real-valued map with domain $[0,T)\times C([0,T];\mathbb{R}^d)$.

(ii) Given $(t,\boldsymbol{x},\mathbf{y}) \in \mathbf{\Lambda}\times\mathbb{R}^d$, the **vertical derivatives** of first and second-order of $\hat{u}$ at $(t,\boldsymbol{x},\mathbf{y})$ (if the corresponding limits exist) are defined as

$$\begin{aligned}\partial_{x_i}^V \hat{u}(t,\boldsymbol{x},\mathbf{y}) \ &:= \ \lim_{h\to 0} \frac{\hat{u}(t,\boldsymbol{x},\mathbf{y}+h\mathbf{e}_i) - \hat{u}(t,\boldsymbol{x},\mathbf{y})}{h}, \\ \partial_{x_ix_j}^V \hat{u}(t,\boldsymbol{x},\mathbf{y}) \ &:= \ \partial_{x_j}^V(\partial_{x_i}^V\hat{u})(t,\boldsymbol{x},\mathbf{y}),\end{aligned}$$

where $\mathbf{e}_1,\ldots,\mathbf{e}_d$ is the standard orthonormal basis of $\mathbb{R}^d$.
When the above limits exist for every $(t,\boldsymbol{x},\mathbf{y}) \in \mathbf{\Lambda}\times\mathbb{R}^d$, then $\partial_{x_i}^V$, $\partial_{x_ix_j}^V$ are real-valued maps with domain $\mathbf{\Lambda}\times\mathbb{R}^d$.
Finally, we denote $\partial_{\boldsymbol{x}}^V\hat{u} = (\partial_{x_1}^V\hat{u},\ldots,\partial_{x_d}^V\hat{u})$ and $\partial_{\boldsymbol{x}\boldsymbol{x}}^V\hat{u} = (\partial_{x_ix_j}^V\hat{u})_{i,j=1,\ldots,d}$.

Remark 2.3. We recall that our aim is to define pathwise derivatives for a map $u\colon \mathbf{\Lambda}\to\mathbb{R}$. In order to do it, as it will be stated in Section 2.3 (Definitions 2.9 and 2.10), we will firstly consider a *lift* of u, that is a map $\hat{u}\colon \mathbf{\Lambda}\times\mathbb{R}^d\to\mathbb{R}$ satisfying

$$u(t,\boldsymbol{x}) \ = \ \hat{u}(t,\boldsymbol{x},\boldsymbol{x}(t)), \qquad \forall\,(t,\boldsymbol{x}) \in \mathbf{\Lambda}.$$

Then, we will define the pathwise derivatives of u at $(t,\boldsymbol{x})$ as the pathwise derivatives of $\hat{u}$ (according to Definition 2.5) at $(t,\boldsymbol{x},\mathbf{y})$ with $\mathbf{y} = \boldsymbol{x}(t)$. The lifting $\hat{u}$ is only used to define vertical derivatives (for which we need to consider *discontinuous* paths or, more precisely, paths with at most a jump at the present time t, the jump size being given by $\mathbf{y} - \boldsymbol{x}(t)$). On the other hand, in the definition of horizontal derivative only continuous paths are involved, so that the horizontal derivative could be directly defined for u (indeed, when $\hat{u}$ is a lifting of u, the horizontal derivative of $\hat{u}$ coincides with that of u). We also notice that we will not need to consider the horizontal derivative of $\hat{u}$ at a point $(t,\boldsymbol{x},\mathbf{y})$ with $\mathbf{y}\neq\boldsymbol{x}(t)$.

Definition 2.6. We denote by $\boldsymbol{C}^{1,0}(\mathbf{\Lambda}\times\mathbb{R}^d)$ the set of $\hat{u}\in\boldsymbol{C}(\mathbf{\Lambda}\times\mathbb{R}^d)$ such that $\partial_t^H\hat{u}$ exists everywhere on $[0,T)\times C([0,T];\mathbb{R}^d)$ and is continuous.

Definition 2.7. We denote by $\boldsymbol{C}^{0,2}(\mathbf{\Lambda}\times\mathbb{R}^d)$ the set of $\hat{u}\in\boldsymbol{C}(\mathbf{\Lambda}\times\mathbb{R}^d)$ such that $\partial_{\boldsymbol{x}}^V\hat{u}$, $\partial_{\boldsymbol{x}\boldsymbol{x}}^V\hat{u}$ exist everywhere on $\mathbf{\Lambda}\times\mathbb{R}^d$ and are continuous.

Definition 2.8. We denote by $\boldsymbol{C}^{1,2}(\mathbf{\Lambda}\times\mathbb{R}^d)$ the set $\boldsymbol{C}^{1,0}(\mathbf{\Lambda}\times\mathbb{R}^d)\cap\boldsymbol{C}^{0,2}(\mathbf{\Lambda}\times\mathbb{R}^d)$.

We can finally state the functional Itô formula for lifted maps, whose proof is reported in Appendix A.

Theorem 2.1. *Let $\hat{u} \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$. Then, for every d-dimensional continuous semimartingale $X = (X_t)_{t\in[0,T]}$ on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, with $X = (X^1, \dots, X^d)$, the following* ***functional Itô formula*** *holds:*

$$\hat{u}(t, X, X_t) \;=\; \hat{u}(0, X, X_0) + \int_0^t \partial_t^H \hat{u}(s, X)\, ds + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_{x_i x_j}^V \hat{u}(s, X, X_s)\, d[X^i, X^j]_s$$
$$+ \sum_{i=1}^d \int_0^t \partial_{x_i}^V \hat{u}(s, X, X_s)\, dX_s^i, \qquad \text{for all } 0 \le t \le T,\ \mathbb{P}\text{-a.s.}$$

Proof. See Appendix A. □

2.3. **Pathwise derivatives for maps on continuous paths.**

Definition 2.9. Let $\hat{u}\colon \boldsymbol{\Lambda} \times \mathbb{R}^d \to \mathbb{R}$ and $u\colon \boldsymbol{\Lambda} \to \mathbb{R}$. We say that $\hat{u}$ is a **lifting** of u if

$$u(t, \boldsymbol{x}) \;=\; \hat{u}(t, \boldsymbol{x}, \boldsymbol{x}(t)),$$

for all $(t, \boldsymbol{x}) \in \boldsymbol{\Lambda}$.

The following consistency property is crucial as it implies that, given a map u admitting two liftings $\hat{u}_1$ and $\hat{u}_2$, their pathwise derivatives coincide on continuous paths (see also Remark 2.4).

Lemma 2.1. *If $\hat{u}_1, \hat{u}_2 \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$ are non-anticipative maps satisfying*

$$\hat{u}_1(t, \boldsymbol{x}, \boldsymbol{x}(t)) \;=\; \hat{u}_2(t, \boldsymbol{x}, \boldsymbol{x}(t)), \qquad \forall\, (t, \boldsymbol{x}) \in \boldsymbol{\Lambda},$$

then

$$\begin{aligned}
\partial_t^H \hat{u}_1(t, \boldsymbol{x}) \;&=\; \partial_t^H \hat{u}_2(t, \boldsymbol{x}), && \forall\, (t, \boldsymbol{x}) \in [0, T) \times C([0, T]; \mathbb{R}^d), \\
\partial_{\boldsymbol{x}}^V \hat{u}_1(t, \boldsymbol{x}, \boldsymbol{x}(t)) \;&=\; \partial_{\boldsymbol{x}}^V \hat{u}_2(t, \boldsymbol{x}, \boldsymbol{x}(t)), && \forall\, (t, \boldsymbol{x}) \in \boldsymbol{\Lambda}, \\
\partial_{\boldsymbol{x}\boldsymbol{x}}^V \hat{u}_1(t, \boldsymbol{x}, \boldsymbol{x}(t)) \;&=\; \partial_{\boldsymbol{x}\boldsymbol{x}}^V \hat{u}_2(t, \boldsymbol{x}, \boldsymbol{x}(t)), && \forall\, (t, \boldsymbol{x}) \in \boldsymbol{\Lambda}.
\end{aligned}$$

Proof. See Appendix B. □

Thanks to Lemma 2.1 we can now give the following definition (see also Remark 2.4).

Definition 2.10. Let $u\colon \boldsymbol{\Lambda} \to \mathbb{R}$. We say that $u \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$ if it admits a lifting $\hat{u}\colon \boldsymbol{\Lambda} \times \mathbb{R}^d \to \mathbb{R}$ satisfying $\hat{u} \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$. Moreover, we define

$$\begin{aligned}
\partial_t^H u(t, \boldsymbol{x}) \;&:=\; \partial_t^H \hat{u}(t, \boldsymbol{x}), && \forall\, (t, \boldsymbol{x}) \in [0, T) \times C([0, T]; \mathbb{R}^d), \\
\partial_{\boldsymbol{x}}^V u(t, \boldsymbol{x}) \;&:=\; \partial_{\boldsymbol{x}}^V \hat{u}(t, \boldsymbol{x}, \boldsymbol{x}(t)), && \forall\, (t, \boldsymbol{x}) \in \boldsymbol{\Lambda}, \\
\partial_{\boldsymbol{x}\boldsymbol{x}}^V u(t, \boldsymbol{x}) \;&:=\; \partial_{\boldsymbol{x}\boldsymbol{x}}^V \hat{u}(t, \boldsymbol{x}, \boldsymbol{x}(t)), && \forall\, (t, \boldsymbol{x}) \in \boldsymbol{\Lambda}.
\end{aligned}$$

Remark 2.4. Notice that, by Lemma 2.1, if $u \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$ then the definition of the pathwise derivatives of u is independent of the lifting $\hat{u} \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$ of u.

Theorem 2.2. *Let $u \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$. Then, for every d-dimensional continuous semimartingale $X = (X_t)_{t\in[0,T]}$ on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, with $X = (X^1, \dots, X^d)$, the following* ***functional Itô formula*** *holds:*

$$u(t, X) \;=\; u(0, X) + \int_0^t \partial_t^H u(s, X)\, ds + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_{x_i x_j}^V u(s, X)\, d[X^i, X^j]_s \tag{2.1}$$
$$+ \sum_{i=1}^d \int_0^t \partial_{x_i}^V u(s, X)\, dX_s^i, \qquad \text{for all } 0 \le t \le T,\ \mathbb{P}\text{-a.s.}$$

Proof. Since $u \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$, by Definition 2.10 there exists a lifting $\hat{u}\colon \boldsymbol{\Lambda} \times \mathbb{R}^d \to \mathbb{R}$ satisfying $\hat{u} \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$. By Theorem 2.1, the below Itô formula holds:

$$\hat{u}(t, X, X_t) = \hat{u}(0, X, X_0) + \int_0^t \partial_t^H \hat{u}(s, X)\, ds + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \partial_{x_i x_j}^V \hat{u}(s, X, X_s)\, d[X^i, X^j]_s + \sum_{i=1}^d \int_0^t \partial_{x_i}^V \hat{u}(s, X, X_s)\, dX_s^i, \qquad \text{for all } 0 \le t \le T, \ \mathbb{P}\text{-a.s.}$$

Now, the claim follows identifying the pathwise derivatives of $\hat{u}$ with those of u. □

3. Smooth gauge-type function with bounded derivatives

The proof of the comparison Theorem (Theorem 4.2) is based on Corollary D.1, which follows from an extended Borwein-Preiss smooth variant of Ekeland's variational principle ([31]), corresponding to Theorem 2.5.2 in [7] (see Appendix D for more details). An essential tool of such a result is the notion of gauge-type function, reported below for the specific set $\boldsymbol{\Lambda}$.

Definition 3.1. We say that $\Psi\colon \boldsymbol{\Lambda} \times \boldsymbol{\Lambda} \to [0, +\infty)$ is a **gauge-type function** provided that:

a) Ψ is continuous on $\boldsymbol{\Lambda} \times \boldsymbol{\Lambda}$;

b) $\Psi((t, \boldsymbol{x}), (t, \boldsymbol{x})) = 0$, for every $(t, \boldsymbol{x}) \in \boldsymbol{\Lambda}$;

c) for every $\varepsilon > 0$ there exists $\eta > 0$ such that, for all $(t', \boldsymbol{x}'), (t'', \boldsymbol{x}'') \in \boldsymbol{\Lambda}$, the inequality $\Psi((t', \boldsymbol{x}'), (t'', \boldsymbol{x}'')) \le \eta$ implies $d_\infty((t', \boldsymbol{x}'), (t'', \boldsymbol{x}'')) < \varepsilon$.

In the proof of the comparison theorem we need such a gauge-type function to be also smooth as a map of its first pair, namely $(t, \boldsymbol{x}) \mapsto \Psi((t, \boldsymbol{x}), (t_0, \boldsymbol{x}_0))$, and with bounded derivatives. The most important example of gauge-type function is the pseudometric d_∞ itself, which unfortunately is not smooth enough.

The present section is devoted to the construction of a smooth gauge-type function with bounded derivatives, which corresponds to the function ρ_∞ in (3.20). In order to do it, we perform a smoothing of the pseudometric d_∞ itself (more precisely of the part concerning the supremum norm), first in the vertical direction, and then in the horizontal direction.

Lemma 3.1. *Let $\zeta\colon \mathbb{R}^d \to \mathbb{R}$ be the probability density function of the standard normal multivariate distribution:*

$$\zeta(\mathbf{z}) := \frac{1}{(2\pi)^{\frac{d}{2}}} \mathrm{e}^{-\frac{1}{2}|\mathbf{z}|^2}, \qquad \forall\, \mathbf{z} \in \mathbb{R}^d.$$

For every fixed $(t_0, \boldsymbol{x}_0) \in \boldsymbol{\Lambda}$, define the map $\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\colon \boldsymbol{\Lambda} \times \mathbb{R}^d \to [0, +\infty)$ as

$$\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}(t, \boldsymbol{x}, \mathbf{y}) := \int_{\mathbb{R}^d} \big\| \boldsymbol{x}(\cdot \wedge t) - \boldsymbol{x}_0(\cdot \wedge t_0) - (\mathbf{y} - \boldsymbol{x}(t) + \mathbf{z}) 1_{[t,T]} \big\|_\infty \zeta(\mathbf{z})\, d\mathbf{z} - \int_{\mathbb{R}^d} |\mathbf{z}|\, \zeta(\mathbf{z})\, d\mathbf{z},$$

for all $(t, \boldsymbol{x}, \mathbf{y}) \in \boldsymbol{\Lambda} \times \mathbb{R}^d$. Moreover, let $\kappa_\infty^{(t_0,\boldsymbol{x}_0)}\colon \boldsymbol{\Lambda} \to [0, +\infty)$ be given by

$$\begin{aligned} \kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t, \boldsymbol{x}) &:= \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}(t, \boldsymbol{x}, \boldsymbol{x}(t)) \\ &= \int_{\mathbb{R}^d} \big\| \boldsymbol{x}(\cdot \wedge t) - \boldsymbol{x}_0(\cdot \wedge t_0) - \mathbf{z} 1_{[t,T]} \big\|_\infty \zeta(\mathbf{z})\, d\mathbf{z} - \int_{\mathbb{R}^d} |\mathbf{z}|\, \zeta(\mathbf{z})\, d\mathbf{z}, \end{aligned}$$

for every $(t, \boldsymbol{x}) \in \boldsymbol{\Lambda}$. Then, the following properties hold.

1) $\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)} \in \boldsymbol{C}^{0,2}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$.

2) *For every $i, j = 1, \ldots, d$, $\partial_{x_i}^V \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}$ is bounded by the constant 1 and $\partial_{x_i x_j}^V \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}$ is bounded by the constant $\sqrt{\frac{2}{\pi}}$.*

3) $\kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x}) \geq \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty - C_\zeta$, *for every* $(t,\boldsymbol{x}) \in \boldsymbol{\Lambda}$, *with*

$$C_\zeta \;:=\; \kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t_0,\boldsymbol{x}_0) \;=\; \int_{\mathbb{R}^d} |\mathbf{z}|\zeta(\mathbf{z})\,d\mathbf{z} \;=\; \sqrt{2}\,\frac{\Gamma\big(\frac{d}{2}+\frac{1}{2}\big)}{\Gamma\big(\frac{d}{2}\big)} \;>\; 0, \tag{3.1}$$

where $\Gamma(\cdot)$ *is the Gamma function.*

4) *For every fixed d, there exists some constant* $\alpha_d > 0$ *such that*

$$\alpha_d\left(\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty^{d+1} \wedge \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty\right) \\ \leq\; \kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x}) \;\leq\; \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty, \tag{3.2}$$

for all $(t,\boldsymbol{x}) \in \boldsymbol{\Lambda}$.

Proof. We split the proof into several steps.

STEP I. *Proof of item* 1). We first notice that $\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}$ is a non-anticipative and continuous map on $\boldsymbol{\Lambda}\times\mathbb{R}^d$, namely $\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)} \in \boldsymbol{C}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$. Now, let $(t,\boldsymbol{x},\mathbf{y}) \in \boldsymbol{\Lambda}\times\mathbb{R}^d$, $h \in \mathbb{R}\backslash\{0\}$, and $i = 1,\ldots,d$ (recall that $\mathbf{e}_1,\ldots,\mathbf{e}_n$ denotes the standard orthonormal basis of $\mathbb{R}^d$), then we have

$$\begin{aligned}
&\frac{\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y}+h\mathbf{e}_i) - \hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y})}{h} \\
&= \frac{1}{h}\bigg(\int_{\mathbb{R}^d}\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - (\mathbf{y} - \boldsymbol{x}(t) + h\mathbf{e}_i + \mathbf{z})1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&\quad - \int_{\mathbb{R}^d}\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - (\mathbf{y} - \boldsymbol{x}(t) + \mathbf{z})1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z}\bigg) \\
&= \frac{1}{h}\bigg(\int_{\mathbb{R}^d}\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - (\mathbf{y} - \boldsymbol{x}(t) + \mathbf{z})1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z}-h\mathbf{e}_i)\,d\mathbf{z} \\
&\quad - \int_{\mathbb{R}^d}\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - (\mathbf{y} - \boldsymbol{x}(t) + \mathbf{z})1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z}\bigg) \\
&= \int_{\mathbb{R}^d}\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - (\mathbf{y} - \boldsymbol{x}(t) + \mathbf{z})1_{[t,T]}\big\|_\infty\,\frac{\zeta(\mathbf{z}-h\mathbf{e}_i) - \zeta(\mathbf{z})}{h}\,d\mathbf{z} \\
&\qquad\qquad \overset{h\to 0}{\longrightarrow}\; -\int_{\mathbb{R}^d}\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - (\mathbf{y} - \boldsymbol{x}(t) + \mathbf{z})1_{[t,T]}\big\|_\infty\,\partial_{z_i}\zeta(\mathbf{z})\,d\mathbf{z},
\end{aligned}$$

where $\partial_{z_i}\zeta(\mathbf{z})$ denotes the partial derivative of ζ in the $\mathbf{e}_i$-direction at the point $\mathbf{z}$, which is given by $-z_i\,\zeta(\mathbf{z})$. This proves that $\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}$ admits continuous first-order vertical derivatives. In a similar way we can prove that $\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}$ also admits continuous second-order vertical derivatives.

STEP II. *Proof of item* 2). We begin noting that $\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}$ is Lipschitz in $\mathbf{y}$ uniformly with respect to $(t,\boldsymbol{x})$:

$$\begin{aligned}
|\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y}) - \hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y}')| &= \bigg|\int_{\mathbb{R}^d}\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - (\mathbf{y} - \boldsymbol{x}(t) + \mathbf{z})1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&\quad - \int_{\mathbb{R}^d}\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - (\mathbf{y}' - \boldsymbol{x}(t) + \mathbf{z})1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z}\bigg| \\
&\leq |\mathbf{y}-\mathbf{y}'|,
\end{aligned}$$

where we have used the fact that $\int_{\mathbb{R}^d}\zeta(\mathbf{z})\,d\mathbf{z} = 1$. It is then easy to see that, for every $i = 1,\ldots,d$, $\partial_{x_i}^V\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}$ is bounded by the constant 1. Proceeding along the same lines as for $\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}$, we deduce that $\partial_{x_i}^V\hat\kappa_\infty^{(t_0,\boldsymbol{x}_0)}$ is Lipschitz in $\mathbf{y}$ uniformly with respect to $(t,\boldsymbol{x})$, with Lipschitz constant less than or equal to $\int_{\mathbb{R}^d} |\partial_{z_i}\zeta(\mathbf{z})|\,d\mathbf{z} = \int_{\mathbb{R}^d} |z_i|\,\zeta(\mathbf{z})\,d\mathbf{z} = \sqrt{\frac{2}{\pi}}$. This allows to prove that, for every $i,j = 1,\ldots,d$, $\partial^V_{x_ix_j}\hat{\kappa}^{(t_0,\boldsymbol{x}_0)}_\infty$ is bounded by $\sqrt{\frac{2}{\pi}}$.

Step III. *Proof of item* 3). We begin noting that (using the fact that ζ is a radial function)

$$
\begin{aligned}
&\kappa^{(t_0,\boldsymbol{x}_0)}_\infty(t,\boldsymbol{x}) + \kappa^{(t_0,\boldsymbol{x}_0)}_\infty(t_0,\boldsymbol{x}_0) \\
&= \int_{[0,+\infty)\times\mathbb{R}^{d-1}} \big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - \mathbf{z}1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&\quad + \int_{(-\infty,0]\times\mathbb{R}^{d-1}} \big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - \mathbf{z}1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&= \int_{[0,+\infty)\times\mathbb{R}^{d-1}} \big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - \mathbf{z}1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&\quad + \int_{[0,+\infty)\times\mathbb{R}^{d-1}} \big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) + \mathbf{z}1_{[t,T]}\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z} \qquad (3.3)\\
&= \int_{[0,+\infty)\times\mathbb{R}^{d-1}} \Big(\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - \mathbf{z}1_{[t,T]}\big\|_\infty + \big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) + \mathbf{z}1_{[t,T]}\big\|_\infty\Big)\,\zeta(\mathbf{z})\,d\mathbf{z}.
\end{aligned}
$$

Now, we observe that, for every $\mathbf{z}\in\mathbb{R}^d$, we have

$$
\begin{aligned}
&\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - \mathbf{z}1_{[t,T]}\big\|_\infty \\
&= \max\Big\{\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t\wedge t_0)\big\|_\infty, \max_{t\le s\le t\vee t_0}\big|\boldsymbol{x}(t) - \boldsymbol{x}_0(s\wedge t_0) - \mathbf{z}\big|\Big\} \qquad (3.4)
\end{aligned}
$$

and similarly for $\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) + \mathbf{z}1_{[t,T]}\|_\infty$. Moreover, by the elementary inequality $|\mathbf{x}-\mathbf{z}| + |\mathbf{x}+\mathbf{z}| \ge 2|\mathbf{x}|$, valid for every $\mathbf{x},\mathbf{z}\in\mathbb{R}^d$, we have

$$
\begin{aligned}
&\max_{t\le s\le t\vee t_0}\big|\boldsymbol{x}(t) - \boldsymbol{x}_0(s\wedge t_0) - \mathbf{z}\big| + \max_{t\le s\le t\vee t_0}\big|\boldsymbol{x}(t) - \boldsymbol{x}_0(s\wedge t_0) + \mathbf{z}\big| \\
&\ge \max_{t\le s\le t\vee t_0}\Big\{\big|\boldsymbol{x}(t) - \boldsymbol{x}_0(s\wedge t_0) - \mathbf{z}\big| + \big|\boldsymbol{x}(t) - \boldsymbol{x}_0(s\wedge t_0) + \mathbf{z}\big|\Big\} \ge 2\max_{t\le s\le t\vee t_0}\big|\boldsymbol{x}(t) - \boldsymbol{x}_0(s\wedge t_0)\big|.
\end{aligned}
$$

Then, using the elementary fact that if $a+b\ge 2c$ it holds that $\max\{\ell,a\} + \max\{\ell,b\} \ge 2\max\{\ell,c\}$, valid for all $a,b,c,\ell\in\mathbb{R}$, we find

$$
\begin{aligned}
&\max\Big\{\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t\wedge t_0)\big\|_\infty, \max_{t\le s\le t\vee t_0}\big|\boldsymbol{x}(t) - \boldsymbol{x}_0(s\wedge t_0) - \mathbf{z}\big|\Big\} \\
&+ \max\Big\{\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t\wedge t_0)\big\|_\infty, \max_{t\le s\le t\vee t_0}\big|\boldsymbol{x}(t) - \boldsymbol{x}_0(s\wedge t_0) + \mathbf{z}\big|\Big\} \\
&\ge 2\max\Big\{\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t\wedge t_0)\big\|_\infty, \max_{t\le s\le t\vee t_0}\big|\boldsymbol{x}(t) - \boldsymbol{x}_0(s\wedge t_0)\big|\Big\} = 2\big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\big\|_\infty,
\end{aligned}
$$

where the last equality follows from (3.4). Therefore, by (3.3) we obtain (also recalling that $C_\zeta := \kappa^{(t_0,\boldsymbol{x}_0)}_\infty(t_0,\boldsymbol{x}_0)$)

$$
\kappa^{(t_0,\boldsymbol{x}_0)}_\infty(t,\boldsymbol{x}) + C_\zeta \ge 2\int_{[0,+\infty)\times\mathbb{R}^{d-1}} \big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\big\|_\infty\,\zeta(\mathbf{z})\,d\mathbf{z} = \big\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\big\|_\infty,
$$

which proves item 3). The explicit expression of the constant C_ζ, reported in (3.1), will be derived in Step VI-4.

Step IV. *Proof of the second inequality in* (3.2). The second inequality in (3.2) follows easily from an application of the triangular inequality, namely noting that $\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0) - \mathbf{z}1_{[t,T]}\|_\infty \le \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty + |\mathbf{z}|$.

STEP V. *Proof of the first inequality in* (3.2) *for the case* $\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty > 2C_\zeta$.
When $\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty > 2C_\zeta$, we have, by item 3),

$$\begin{aligned}
\kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x}) &\geq \|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty - C_\zeta \\
&= \|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty \bigg(1-\frac{C_\zeta}{\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty}\bigg) \\
&\geq \frac{1}{2}\,\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty \\
&\geq \frac{1}{2}\big(\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty^{d+1}\wedge\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty\big),
\end{aligned}$$

which proves the first inequality in (3.2) with $\alpha_d := \frac{1}{2}$, for the case $\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty > 2C_\zeta$.

STEP VI. *Proof of the first inequality in* (3.2) *for the case* $\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty \leq 2C_\zeta$.

STEP VI-1. Our aim is to prove that for every fixed d there exists some constant $\alpha_d > 0$ such that

$$\begin{aligned}
&\int_{\mathbb{R}^d}\max\big\{a,|\mathbf{y}-\mathbf{z}|\big\}\,\zeta(\mathbf{z})\,d\mathbf{z} - \int_{\mathbb{R}^d}|\mathbf{z}|\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&\qquad \geq\ \alpha_d\min\{a^{d+1},a\} + \alpha_d\min\{|\mathbf{y}|^{d+1},|\mathbf{y}|\}, \qquad \forall\,(a,\mathbf{y})\in[0,2C_\zeta]\times\mathbb{R}^d.
\end{aligned} \tag{3.5}$$

As a matter of fact, suppose for a moment that (3.5) holds true. Then, applying (3.5) with $a := \|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t\wedge t_0)\|_\infty$ and $\mathbf{y}_s := \boldsymbol{x}(t)-\boldsymbol{x}_0(s\wedge t_0)$, for every $s\in[t,t\vee t_0]$, and taking the maximum over $s\in[t,t\vee t_0]$, we find (using (3.4))

$$\begin{aligned}
\kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x}) &= \int_{\mathbb{R}^d}\max\Big\{a,\max_{t\leq s\leq t\vee t_0}|\mathbf{y}_s-\mathbf{z}|\Big\}\,\zeta(\mathbf{z})\,d\mathbf{z} - \int_{\mathbb{R}^d}|\mathbf{z}|\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&= \int_{\mathbb{R}^d}\max_{t\leq s\leq t\vee t_0}\big\{\max\{a,|\mathbf{y}_s-\mathbf{z}|\}\big\}\,\zeta(\mathbf{z})\,d\mathbf{z} - \int_{\mathbb{R}^d}|\mathbf{z}|\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&\geq \max_{t\leq s\leq t\vee t_0}\bigg\{\int_{\mathbb{R}^d}\max\big\{a,|\mathbf{y}_s-\mathbf{z}|\big\}\,\zeta(\mathbf{z})\,d\mathbf{z}\bigg\} - \int_{\mathbb{R}^d}|\mathbf{z}|\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&= \max_{t\leq s\leq t\vee t_0}\bigg\{\int_{\mathbb{R}^d}\max\big\{a,|\mathbf{y}_s-\mathbf{z}|\big\}\,\zeta(\mathbf{z})\,d\mathbf{z} - \int_{\mathbb{R}^d}|\mathbf{z}|\,\zeta(\mathbf{z})\,d\mathbf{z}\bigg\} \\
&\geq \max_{t\leq s\leq t\vee t_0}\Big\{\alpha_d\min\{a^{d+1},a\} + \alpha_d\min\{|\mathbf{y}_s|^{d+1},|\mathbf{y}_s|\}\Big\} \\
&= \alpha_d\min\{a^{d+1},a\} + \alpha_d\max_{t\leq s\leq t\vee t_0}\Big\{\min\{|\mathbf{y}_s|^{d+1},|\mathbf{y}_s|\}\Big\} \\
&= \alpha_d\min\{a^{d+1},a\} + \alpha_d\min\Big\{\max_{t\leq s\leq t\vee t_0}|\mathbf{y}_s|^{d+1},\max_{t\leq s\leq t\vee t_0}|\mathbf{y}_s|\Big\}.
\end{aligned}$$

Hence, by the elementary inequality

$$\min\{a^{d+1},a\}+\min\{b^{d+1},b\}\ \geq\ \min\big\{\max\{a^{d+1},b^{d+1}\},\max\{a,b\}\big\}, \qquad \forall\,a,b\geq 0,$$

we conclude that

$$\begin{aligned}
\kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x}) &\geq \alpha_d\,\min\Big\{\max\Big\{a^{d+1},\max_{t\leq s\leq t\vee t_0}|\mathbf{y}_s|^{d+1}\Big\},\max\Big\{a,\max_{t\leq s\leq t\vee t_0}|\mathbf{y}_s|\Big\}\Big\} \\
&= \alpha_d\,\min\big\{\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty^{d+1},\|\boldsymbol{x}(\cdot\wedge t)-\boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty\big\},
\end{aligned}$$

where the last equality follows from (3.4) with $\mathbf{z}=0$. This yields the first inequality in (3.2). It remains to prove (3.5).

Step VI-2. *Proof of* (3.5). For every positive integer d and $a \geq 0$, let $G_a\colon \mathbb{R}^d \to \mathbb{R}$ be given by

$$G_a(\mathbf{y}) \;:=\; \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{y}-\mathbf{z}|\big\}\,\zeta(\mathbf{z})\,d\mathbf{z} - \int_{\mathbb{R}^d} |\mathbf{z}|\,\zeta(\mathbf{z})\,d\mathbf{z}, \qquad \forall\,\mathbf{y}\in\mathbb{R}^d. \tag{3.6}$$

Moreover, let $F_d\colon [0,+\infty) \to \mathbb{R}$ be defined as (differently to the notation used for G_a, we emphasize the dependence of F_d on the dimension d; we do this because of statement (3.9) below which changes with d)

$$F_d(a) \;:=\; G_a(\mathbf{0}) \;=\; \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{z}|\big\}\,\zeta(\mathbf{z})\,d\mathbf{z} - \int_{\mathbb{R}^d} |\mathbf{z}|\,\zeta(\mathbf{z})\,d\mathbf{z}, \qquad \forall\,a\in[0,+\infty). \tag{3.7}$$

Notice that G_a and F_d are convex functions on their domains.

Let us fix some notations. We denote by $\partial_{\mathbf{y}}G_a(\mathbf{y})$ and $\partial_{\mathbf{y}\mathbf{y}}G_a(\mathbf{y})$ (resp. $F_d'(a)$, $F_d''(a)$, $\ldots$, $F_d^{(n)}(a)$) the gradient and Hessian (resp. first-order derivative, second-order derivative, $\ldots$, n-th order derivative) of G_a at $\mathbf{y}$ (resp. F_d at a). When $a=0$, $F_d'(a)$, $F_d''(a)$, $\ldots$, $F_d^{(n)}(a)$ are right-derivatives. We also denote by I the $d\times d$ identity matrix. Finally, given A and B in $\mathcal{S}(d)$, the inequality $B \leq A$ means that the symmetric matrix $A-B$ is positive semi-definite.

Our aim is to prove the following: *for every d, there exist constants $\beta_d>0$ and $L_d>0$ such that*

$$\begin{aligned} &\forall\,a\in[0,2C_\zeta],\ G_a\in C^2(\mathbb{R}^d),\ \partial_{\mathbf{y}}G_a(\mathbf{0})=0,\ \partial_{\mathbf{y}\mathbf{y}}G_a(\mathbf{0})\geq\beta_d I \ \textit{and}\ \partial_{\mathbf{y}\mathbf{y}}G_a \ \textit{satisfies} \\ &\qquad\qquad \partial_{\mathbf{y}\mathbf{y}}G_a(\mathbf{y}) - \partial_{\mathbf{y}\mathbf{y}}G_a(\mathbf{0}) \;\geq\; -L_d\,|\mathbf{y}|\,I, \qquad \forall\,\mathbf{y}\in\mathbb{R}^d \end{aligned} \tag{3.8}$$

and

$$\begin{aligned} &F_d\in C^{d+1}([0,+\infty)),\ F_d'(0)=\cdots=F_d^{(d)}(0)=0,\ F_d^{(d+1)}(0)\geq\beta_d \ \textit{and}\ F_d^{(d+1)} \ \textit{satisfies} \\ &\qquad\qquad F_d^{(d+1)}(a) - F_d^{(d+1)}(0) \;\geq\; -L_d\,a, \qquad \forall\,a\geq 0. \end{aligned} \tag{3.9}$$

Suppose for a moment that (3.8) and (3.9) hold. Then, by (3.8) we show below that there exist some constants $\delta_d, \tilde\delta_d \in (0,1]$ such that

$$G_a(\mathbf{y}) \;\geq\; G_a(\mathbf{0}) + \frac{1}{4}\beta_d\,|\mathbf{y}|^2, \qquad \forall\,|\mathbf{y}|\leq\delta_d,\ \forall\,a\in[0,2C_\zeta], \tag{3.10}$$

$$F_d(a) \;\geq\; F_d(0) + \frac{1}{2(d+1)!}\,\beta_d\,a^{d+1}, \qquad \forall\,a\in[0,\tilde\delta_d]. \tag{3.11}$$

As a matter of fact, for every fixed $\mathbf{y}\in\mathbb{R}^d$, set $\varphi_a(\lambda) := G_a(\lambda\mathbf{y})$, for every $\lambda\in\mathbb{R}$. Since $\varphi_a\in C^2(\mathbb{R})$, the Taylor expression given by

$$\varphi_a(1) \;=\; \varphi_a(0) + \varphi_a'(0) + \frac{1}{2}\varphi_a''(0) + \int_0^1 (1-\lambda)\,(\varphi_a''(\lambda)-\varphi_a''(0))\,d\lambda,$$

which written in terms of G_a becomes (denoting by $\langle\cdot,\cdot\rangle$ the scalar product in $\mathbb{R}^d$)

$$\begin{aligned} G_a(\mathbf{y}) \;&=\; G_a(\mathbf{0}) + \langle\partial_{\mathbf{y}}G_a(\mathbf{0}),\mathbf{y}\rangle + \frac{1}{2}\langle\partial_{\mathbf{y}\mathbf{y}}G_a(\mathbf{0})\mathbf{y},\mathbf{y}\rangle \\ &\quad + \int_0^1 (1-\lambda)\,\langle(\partial_{\mathbf{y}\mathbf{y}}G_a(\lambda\mathbf{y}) - \partial_{\mathbf{y}\mathbf{y}}G_a(\mathbf{0}))\mathbf{y},\mathbf{y}\rangle\,d\lambda \\ &\geq\; G_a(\mathbf{0}) + \langle\partial_{\mathbf{y}}G_a(\mathbf{0}),\mathbf{y}\rangle + \frac{1}{2}\beta_d|\mathbf{y}|^2 - L_d|\mathbf{y}|^3\int_0^1 \lambda\,(1-\lambda)\,d\lambda \\ &=\; G_a(\mathbf{0}) + \frac{1}{2}\beta_d|\mathbf{y}|^2 - \frac{1}{6}L_d|\mathbf{y}|^3. \end{aligned}$$

Hence

$$G_a(\mathbf{y}) \;\geq\; G_a(\mathbf{0}) + \frac{1}{4}\,\beta_d\,|\mathbf{y}|^2, \qquad \forall\,|\mathbf{y}| \leq \delta_d,\ \text{where } \delta_d := 1 \wedge \bigg(\frac{3}{2}\frac{\beta_d}{L_d}\bigg).$$

This proves (3.10). Similarly, we consider the following Taylor expression for F_d:

$$F_d(a) \;=\; \sum_{k=0}^{d+1} \frac{F_d^{(k)}(0)}{k!}\,a^k + \frac{1}{d!}\int_0^a \big(F_d^{(d+1)}(b) - F_d^{(d+1)}(0)\big)\,(a-b)^d\,db.$$

By (3.9) we obtain

$$\begin{aligned} F_d(a) \;&=\; F_d(0) + \frac{F_d^{(d+1)}(0)}{(d+1)!}\,a^{d+1} + \frac{1}{d!}\int_0^a \big(F_d^{(d+1)}(b) - F_d^{(d+1)}(0)\big)\,(a-b)^d\,db \\ &\geq\; F_d(0) + \frac{\beta_d}{(d+1)!}\,a^{d+1} - \frac{L_d}{d!}\int_0^a b\,(a-b)^d\,db \\ &=\; F_d(0) + \frac{\beta_d}{(d+1)!}\,a^{d+1} - \frac{L_d}{(d+2)!}\,a^{d+2}. \end{aligned}$$

Hence

$$F_d(a) \;\geq\; F_d(0) + \frac{1}{2}\frac{\beta_d}{(d+1)!}\,a^{(d+1)}, \qquad \forall\,a \in [0,\tilde\delta_d],\ \text{where } \tilde\delta_d := 1 \wedge \bigg(\frac{d+2}{2}\frac{\beta_d}{L_d}\bigg).$$

This proves (3.11). Now, we notice that from (3.10) we obtain

$$G_a(\mathbf{y}) \;\geq\; G_a(\mathbf{0}) + \frac{1}{4}\,\beta_d\,|\mathbf{y}|^{d+1}, \qquad \forall\,|\mathbf{y}| \leq \delta_d,\ \text{where } \delta_d := 1 \wedge \bigg(\frac{3}{2}\frac{\beta_d}{L_d}\bigg).$$

Moreover, since G_a is a convex function, it follows that

$$\begin{aligned} G_a(\mathbf{y}) \;&\geq\; \min\Big(G_a(\mathbf{0}) + \frac{1}{4}\,\beta_d\,|\mathbf{y}|^{d+1}, G_a(\mathbf{0}) + \frac{1}{4}\,\beta_d\,\delta_d^d\,|\mathbf{y}|\Big) \\ &\geq\; \min\Big(G_a(\mathbf{0}) + \frac{1}{4}\,\beta_d\,\delta_d^d\,|\mathbf{y}|^{d+1}, G_a(\mathbf{0}) + \frac{1}{4}\,\beta_d\,\delta_d^d\,|\mathbf{y}|\Big) \\ &=\; G_a(\mathbf{0}) + \frac{1}{4}\,\beta_d\,\delta_d^d\,\big(|\mathbf{y}|^{d+1} \wedge |\mathbf{y}|\big), \qquad \forall\,\mathbf{y} \in \mathbb{R}^d. \end{aligned} \tag{3.12}$$

Proceeding along the same lines, we deduce by (3.11) that

$$F_d(a) \;\geq\; F_d(0) + \frac{1}{4}\,\beta_d\,\tilde\delta_d^d\,\big(a^{d+1} \wedge a\big), \qquad \forall\,a \geq 0. \tag{3.13}$$

So, in particular, since $G_a(\mathbf{0}) = F_d(a)$ and $F_d(0) = 0$, we obtain, from (3.12) and (3.13),

$$G_a(\mathbf{y}) \;\geq\; \frac{1}{4}\,\beta_d\,\tilde\delta_d^d\,(a^{d+1} \wedge a) + \frac{1}{4}\,\beta_d\,\delta_d^d\,\big(|\mathbf{y}|^{d+1} \wedge |\mathbf{y}|\big), \qquad \forall\,(a,\mathbf{y}) \in [0,2C_\zeta] \times \mathbb{R}^d,$$

which proves (3.5) with $\alpha_d := \frac{1}{4}\beta_d(\tilde\delta_d^d \wedge \delta_d^d)$ for the case $a \leq 2C_\zeta$. It remains to prove (3.8) and (3.9).

STEP VI-3. *Proof of* (3.9). From the definition (3.7) of F_d we see that F_d is continuous. Moreover, by direct calculation we find

$$\begin{aligned} \lim_{h\to 0^+} \frac{F_d(a+h) - F_d(a)}{h} \;&=\; \int_{|\mathbf{z}|\leq a} \zeta(\mathbf{z})\,d\mathbf{z}, \qquad \forall\,a \geq 0, \\ \lim_{h\to 0^+} \frac{F_d(a-h) - F_d(a)}{-h} \;&=\; \int_{|\mathbf{z}|\leq a} \zeta(\mathbf{z})\,d\mathbf{z}, \qquad \forall\,a > 0. \end{aligned}$$

Hence, the first derivative of F_d exists everywhere and is given by

$$F_d'(a) \;=\; \int_{|\mathbf{z}|\leq a} \zeta(\mathbf{z})\,d\mathbf{z}, \qquad \forall\,a \geq 0.$$

Notice that $F_d'(0) = 0$. We also see that F_d' is continuous on $[0,+\infty)$. Now, for every $r > 0$ let $S_{d-1}(r)$ denote the surface area of the boundary of the ball $\{\mathbf{z} \in \mathbb{R}^d : |\mathbf{z}| \leq r\}$, which is given by $S_{d-1}(r) = \frac{2\pi^{d/2}}{\Gamma(d/2)} r^{d-1}$, where $\Gamma(\cdot)$ is the Gamma function. Then, recalling that ζ is a radial function and using d-dimensional spherical coordinates (see for instance Appendix C.3 in [36]), we get

$$F_d'(a) \;=\; \int_0^a \frac{1}{(2\pi)^{\frac{d}{2}}}\, \mathrm{e}^{-\frac{1}{2}r^2}\, S_{d-1}(r)\, dr, \qquad \forall\, a \geq 0.$$

So, in particular, the second derivative of F_d exists everywhere and is given by

$$F_d''(a) \;=\; \frac{1}{(2\pi)^{\frac{d}{2}}}\, \mathrm{e}^{-\frac{1}{2}a^2}\, S_{d-1}(a) \;=\; \frac{2^{1-\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}\, a^{d-1}\, \mathrm{e}^{-\frac{1}{2}a^2}, \qquad \forall\, a \geq 0.$$

We deduce that $F_d \in C^\infty([0,+\infty))$. We also observe that every derivative of F_d is bounded, so in particular $F_d^{(d+1)}$ is Lipschitz. As a consequence, there exists $L_d > 0$ such that

$$F_d^{(d+1)}(a) - F_d^{(d+1)}(0) \;\geq\; -L_d\, a, \qquad \forall\, a \geq 0.$$

Finally, let us prove by induction on d that $F_d'(0) = \cdots = F_d^{(d)}(0) = 0$ and $F_d^{(d+1)}(0) > 0$. For $d = 1$ we have, by direct calculation, $F_1(0) = F_1'(0) = 0$ and $F_1''(0) = 1/\sqrt{2\pi} > 0$. Let us now suppose that the claim holds true for F_d, for some $d \geq 1$, and let us prove it for F_{d+1}. By the explicit expressions of F_{d+1} and F_{d+1}' we see that $F_{d+1}(0) = F_{d+1}'(0) = 0$. Moreover

$$F_{d+1}''(a) \;=\; C_{d+1}\, a^d\, \mathrm{e}^{-\frac{1}{2}a^2},$$

where $C_{d+1} := 2^{1-(d+1)/2}/\Gamma((d+1)/2) > 0$. So, in particular, $F_{d+1}''(0) = 0$. Now, we observe that

$$F_{d+1}'''(a) \;=\; d\, C_{d+1}\, a^{d-1}\, \mathrm{e}^{-\frac{1}{2}a^2} - C_{d+1}\, a^{d+1}\, \mathrm{e}^{-\frac{1}{2}a^2} \;=\; \frac{C_{d+1}}{C_d}\, (d - a^2)\, F_d''(a),$$

where $C_d := 2^{1-d/2}/\Gamma(d/2) > 0$. Therefore

$$F_{d+1}^{\mathrm{iv}}(a) \;=\; \frac{C_{d+1}}{C_d}\, \big((d - a^2)\, F_d'''(a) - 2\, a\, F_d''(a)\big).$$

Moreover, by the general Leibniz rule, we have

$$F_{d+1}^{(3+n)}(a) \;=\; \frac{C_{d+1}}{C_d} \sum_{k=0}^n \binom{n}{k} (d - a^2)^{(n-k)}\, F_d^{(2+k)}(a), \qquad \text{for every } n \geq 2,$$

where $(d-a^2)^{(n-k)}$ denotes the $(n-k)$-th derivative of the map $a \mapsto d - a^2$. Since $(d-a^2)^{(n-k)}$ is identically equal to zero whenever $n - k \geq 3$, it follows that

$$\begin{aligned}
F_{d+1}^{(3+n)}(a) \;&=\; \frac{C_{d+1}}{C_d}\bigg(\binom{n}{n-2} (d-a^2)^{(2)}\, F_d^{(2+n-2)}(a) \\
&\quad + \binom{n}{n-1} (d-a^2)^{(1)}\, F_d^{(2+n-1)}(a) + \binom{n}{n} (d-a^2)\, F_d^{(2+n)}(a)\bigg) \\
&=\; \frac{C_{d+1}}{C_d}\big(- n\,(n-1)\, F_d^{(2+n-2)}(a) - 2\, n\, a\, F_d^{(2+n-1)}(a) + (d-a^2)\, F_d^{(2+n)}(a)\big).
\end{aligned}$$

In conclusion, we have

$$\begin{aligned}
F_{d+1}'''(0) \;&=\; \frac{C_{d+1}}{C_d}\, d\, F_d''(0), \\
F_{d+1}^{\mathrm{iv}}(0) \;&=\; \frac{C_{d+1}}{C_d}\, d\, F_d'''(0),
\end{aligned}$$

$$F_{d+1}^{(3+n)}(0) \;=\; \frac{C_{d+1}}{C_d}\big(-n\,(n-1)\,F_d^{(2+n-2)}(0) + d\,F_d^{(2+n)}(0)\big), \qquad \text{for every } n \geq 2.$$

From the formulae above it is straightforward to see that the claim holds. This concludes the proof of (3.9).

STEP VI-4. *Proof of* (3.8). From the definition (3.6) of G_a we see that $G_a \in C^\infty(\mathbb{R}^d)$. Moreover, we have, for every $i,j = 1,\ldots,d$,

$$\partial_{y_i} G_a(\mathbf{y}) = -\int_{\mathbb{R}^d} \max\big\{a, |\mathbf{y}-\mathbf{z}|\big\}\,\partial_{z_i}\zeta(\mathbf{z})\,d\mathbf{z} = \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{y}-\mathbf{z}|\big\}\,z_i\,\zeta(\mathbf{z})\,d\mathbf{z},$$

$$\partial_{y_iy_j} G_a(\mathbf{y}) = \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{y}-\mathbf{z}|\big\}\,\partial_{z_iz_j}\zeta(\mathbf{z})\,d\mathbf{z} = \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{y}-\mathbf{z}|\big\}\,\big(z_i\,z_j - \delta_{ij}\big)\,\zeta(\mathbf{z})\,d\mathbf{z},$$

where δ_{ij} is the Kronecker delta. Since ζ is a radial function, we have $\zeta(\mathbf{z}) = \zeta(-\mathbf{z})$, for every $\mathbf{z} \in \mathbb{R}^d$, therefore $\partial_{y_i} G_a(\mathbf{0}) = 0$.

We now prove that for every fixed d there exists $L_d > 0$ such that, for every $a \geq 0$, we have

$$\partial_{\mathbf{yy}} G_a(\mathbf{y}) - \partial_{\mathbf{yy}} G_a(\mathbf{0}) \;\geq\; -L_d|\mathbf{y}|I, \qquad \forall\,\mathbf{y} \in \mathbb{R}^d,$$

which can be equivalently written as

$$\sum_{i,j=1}^d \big(\partial_{y_iy_j} G_a(\mathbf{y}) - \partial_{y_iy_j} G_a(\mathbf{0})\big) w_i w_j \;\geq\; -L_d|\mathbf{y}||\mathbf{w}|^2, \qquad \forall\,\mathbf{y},\mathbf{w} \in \mathbb{R}^d, \tag{3.14}$$

where y_i (resp. w_i) denotes the i-th component of $\mathbf{y}$ (resp. $\mathbf{w}$). We begin noting that, for every $i,j = 1,\ldots,d$, we have (we use the elementary inequality $|\max\{a,b+c\} - \max\{a,c\}| \leq |b|$, valid for every $a,c \geq 0$ and $b \in \mathbb{R}$, with $b = |\mathbf{y}-\mathbf{z}| - |\mathbf{z}|$ and $c = |\mathbf{z}|$)

$$\begin{aligned}
\big|\partial_{y_iy_j} G_a(\mathbf{y}) - \partial_{y_iy_j} G_a(\mathbf{0})\big| \;&\leq\; \int_{\mathbb{R}^d} \big|\max\{a,|\mathbf{y}-\mathbf{z}|\} - \max\{a,|\mathbf{z}|\}\big|\,\big|\partial_{z_iz_j}\zeta(\mathbf{z})\big|\,d\mathbf{z} \\
&=\; \int_{\mathbb{R}^d} \big|\max\{a,|\mathbf{y}-\mathbf{z}| - |\mathbf{z}| + |\mathbf{z}|\} - \max\{a,|\mathbf{z}|\}\big|\,\big|\partial_{z_iz_j}\zeta(\mathbf{z})\big|\,d\mathbf{z} \\
&\leq\; \int_{\mathbb{R}^d} |\mathbf{y}|\,|\partial_{z_iz_j}\zeta(\mathbf{z})|\,d\mathbf{z} \;=\; \frac{L_d}{d}\,|\mathbf{y}|,
\end{aligned}$$

with $\frac{L_d}{d} = \int_{\mathbb{R}^d} |\partial_{z_iz_j}\zeta(\mathbf{z})|\,d\mathbf{z}$. Then, for every $\mathbf{w} \in \mathbb{R}^d$, we obtain

$$\bigg|\sum_{i,j=1}^d \big(\partial_{y_iy_j} G_a(\mathbf{y}) - \partial_{y_iy_j} G_a(\mathbf{0})\big) w_i w_j\bigg| \;\leq\; \frac{L_d}{d}\,|\mathbf{y}| \sum_{i,j=1}^d |w_i||w_j| \;\leq\; L_d\,|\mathbf{y}|\,|\mathbf{w}|^2,$$

which proves (3.14).

Finally, we prove that for every fixed d there exists $\hat\beta_d > 0$ such that, for every $a \in [0, 2C_\zeta]$,

$$\partial_{\mathbf{yy}} G_a(\mathbf{0}) \;\geq\; \hat\beta_d I. \tag{3.15}$$

As a matter of fact, for every $\mathbf{w} \in \mathbb{R}^d$, we have

$$\begin{aligned}
\langle \partial_{\mathbf{yy}} G_a(\mathbf{0})\mathbf{w},\mathbf{w}\rangle \;&=\; \sum_{i,j=1}^d \partial_{y_iy_j} G_a(\mathbf{0})\,w_i\,w_j \;=\; \sum_{i,j=1}^d w_i\,w_j \int_{\mathbb{R}^d} \max\big\{a,|\mathbf{z}|\big\}\,\big(z_i\,z_j - \delta_{ij}\big)\,\zeta(\mathbf{z})\,d\mathbf{z} \\
&=\; \sum_{i=1}^d w_i^2 \int_{\mathbb{R}^d} \max\big\{a,|\mathbf{z}|\big\}\,\big(z_1^2 - 1\big)\,\zeta(\mathbf{z})\,d\mathbf{z} - \sum_{i\neq j} w_i\,w_j \int_{\mathbb{R}^d} \max\big\{a,|\mathbf{z}|\big\}\,z_1\,z_2\,\zeta(\mathbf{z})\,d\mathbf{z},
\end{aligned}$$

where the latter equality comes from

$$\int_{\mathbb{R}^d} \max\big\{a,|\mathbf{z}|\big\}\,\big(1 - z_i^2\big)\,\zeta(\mathbf{z})\,d\mathbf{z} \;=\; \int_{\mathbb{R}^d} \max\big\{a,|\mathbf{z}|\big\}\,\big(1 - z_1^2\big)\,\zeta(\mathbf{z})\,d\mathbf{z}, \qquad i = 1,\ldots,d,$$

$$\int_{\mathbb{R}^d} \max\big\{a, |\mathbf{z}|\big\}\, z_i\, z_j\, \zeta(\mathbf{z})\, d\mathbf{z} \;=\; \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{z}|\big\}\, z_1\, z_2\, \zeta(\mathbf{z})\, d\mathbf{z}, \qquad i,j = 1,\dots,d, \;\; i \neq j.$$

Now, notice that

$$\int_{\mathbb{R}^d} \max\big\{a, |\mathbf{z}|\big\}\, z_1\, z_2\, \zeta(\mathbf{z})\, d\mathbf{z} \;=\; 0, \qquad \text{for every } a \geq 0.$$

Hence

$$\begin{aligned}
\langle \partial_{\mathbf{y}\mathbf{y}} G_a(\mathbf{0})\mathbf{w}, \mathbf{w}\rangle \;&=\; |\mathbf{w}|^2 \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{z}|\big\}\, \big(z_1^2 - 1\big)\, \zeta(\mathbf{z})\, d\mathbf{z} \\
&=\; |\mathbf{w}|^2 \frac{1}{d} \sum_{i=1}^d \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{z}|\big\}\, \big(z_1^2 - 1\big)\, \zeta(\mathbf{z})\, d\mathbf{z} \\
&=\; |\mathbf{w}|^2 \frac{1}{d} \sum_{i=1}^d \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{z}|\big\}\, \big(z_i^2 - 1\big)\, \zeta(\mathbf{z})\, d\mathbf{z} \\
&=\; |\mathbf{w}|^2 \frac{1}{d} \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{z}|\big\}\, \big(|\mathbf{z}|^2 - d\big)\, \zeta(\mathbf{z})\, d\mathbf{z}.
\end{aligned}$$

Let $H\colon [0,+\infty) \to \mathbb{R}$ be defined as

$$H(a) \;:=\; \int_{\mathbb{R}^d} \max\big\{a, |\mathbf{z}|\big\}\, \big(|\mathbf{z}|^2 - d\big)\, \zeta(\mathbf{z})\, d\mathbf{z}, \qquad \forall\, a \in [0,+\infty).$$

Notice that (3.15) follows if we prove the following (actually, it would be enough to require that $H(a) > 0$ for every $a \geq 0$ and H decreasing; (3.16) is a sufficient condition for this):

(3.16) $$H(0) > 0, \; \lim_{a\to+\infty} H(a) = 0, \; H \textit{ is a decreasing function.}$$

As a matter of fact, if (3.16) holds, then

$$\inf_{a \in [0, 2C_\zeta]} H(a) \;=\; H(2C_\zeta) \;>\; 0,$$

from which (3.15) follows with $\hat{\beta}_d = \frac{1}{d} H(2C_\zeta)$.

It remains to prove (3.16). Denoting by $\mu_{\chi^2(d),p}$ the moment of order $p > 0$ of a χ^2-distribution with d degrees of freedom, and recalling that $\Gamma(\cdot)$ is the Gamma function, we have

$$\begin{aligned}
H(0) \;=\; \mu_{\chi^2(d),\frac{3}{2}} - d\, \mu_{\chi^2(d),\frac{1}{2}} \;&=\; 2^{\frac{3}{2}}\, \frac{\Gamma\big(\frac{d}{2} + \frac{3}{2}\big)}{\Gamma\big(\frac{d}{2}\big)} - d\, 2^{\frac{1}{2}}\, \frac{\Gamma\big(\frac{d}{2} + \frac{1}{2}\big)}{\Gamma\big(\frac{d}{2}\big)} \\
&=\; 2^{\frac{3}{2}}\, \frac{\big(\frac{d}{2} + \frac{1}{2}\big)\Gamma\big(\frac{d}{2} + \frac{1}{2}\big)}{\Gamma\big(\frac{d}{2}\big)} - d\, 2^{\frac{1}{2}}\, \frac{\Gamma\big(\frac{d}{2} + \frac{1}{2}\big)}{\Gamma\big(\frac{d}{2}\big)} \;=\; \mu_{\chi^2(d),\frac{1}{2}} \;>\; 0.
\end{aligned}$$

Incidentally, we notice that $C_\zeta = \mu_{\chi^2(d),\frac{1}{2}}$, which proves formula (3.1). Concerning the function H, we also have that (we perform the change of variables $\mathbf{z} = a\mathbf{w}$ under the integral sign)

$$\lim_{a\to+\infty} H(a) \;=\; \lim_{a\to+\infty} \int_{\mathbb{R}^d} a^2\, \max\big\{1, |\mathbf{w}|\big\}\, \big(a^2\, |\mathbf{w}|^2 - d\big)\, \zeta(a\mathbf{w})\, d\mathbf{w} \;=\; 0,$$

where the limit follows from an application of the Lebesgue dominated convergence theorem.

Now, proceeding as in Step VI-3 for the function F, we deduce that $H \in C^\infty([0,+\infty))$ and

$$H'(a) \;=\; \int_{|\mathbf{z}|\leq a} \big(|\mathbf{z}|^2 - d\big)\, \zeta(\mathbf{z})\, d\mathbf{z}, \qquad H''(a) \;=\; \frac{a^2 - d}{(2\pi)^{\frac{d}{2}}}\, \mathrm{e}^{-\frac{1}{2}a^2}\, S_{d-1}(a), \qquad \forall\, a \in [0,+\infty),$$

where $S_{d-1}(a)$ denotes the surface area of the boundary of the ball $\{\mathbf{z} \in \mathbb{R}^d \colon |\mathbf{z}| \leq a\}$. Notice that

$$H'(0) \;=\; 0, \qquad\qquad \lim_{a\to+\infty} H'(a) \;=\; \int_{\mathbb{R}^d} \big(|\mathbf{z}|^2 - d\big)\,\zeta(\mathbf{z})\,d\mathbf{z} \;=\; \mu_{\chi^2(d),1} - d \;=\; 0.$$

Then, we deduce from the sign of H'' that $H'(a) < 0$, for every $a > 0$. This implies that H is a strictly decreasing function and concludes the proof. □

Notice that the map

$$\big((t,\boldsymbol{x}),(t_0,\boldsymbol{x}_0)\big) \;\longmapsto\; |t-t_0|^2 + \kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x})$$

is already a gauge-type function. However, as a map of the first pair, it is not smooth in the horizontal direction because of the second term $\kappa_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x})$. We now address the problem of smoothing the map $\kappa_\infty^{(t_0,\boldsymbol{x}_0)}$ with respect to the horizontal derivative. To this regard, the following lemma plays an important role.

Lemma 3.2. *Let $\eta\colon [0,+\infty) \to \mathbb{R}$ be a function satisfying the following properties.*

- $\eta \in C^1([0,+\infty))$;
- η *is non-negative;*
- $\int_0^{+\infty} \eta(s)\,ds = 1$;
- $\eta(0) = 0$.

Let $\hat u \in \boldsymbol{C}^{0,2}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$. The map $\hat v\colon \boldsymbol{\Lambda}\times\mathbb{R}^d \to \mathbb{R}$ given by

$$\hat v(t,\boldsymbol{x},\mathbf{y}) \;:=\; \int_0^{+\infty} \hat u\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)\,\eta(s)\,ds, \qquad \forall\,(t,\boldsymbol{x},\mathbf{y}) \in \boldsymbol{\Lambda}\times\mathbb{R}^d,$$

belongs to $\boldsymbol{C}^{1,2}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$.

Proof. We begin noting that $\hat v$ is a non-anticipative and continuous map on $\boldsymbol{\Lambda}\times\mathbb{R}^d$, namely $\hat v \in \boldsymbol{C}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$. Now, let $(t,\boldsymbol{x}) \in [0,T)\times C([0,T];\mathbb{R}^d)$ and $\delta > 0$, with $\delta \leq T-t$, then we have

$$\begin{aligned}
&\frac{\hat v(t+\delta,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)) - \hat v(t,\boldsymbol{x},\boldsymbol{x}(t))}{\delta} \\
&= \frac{1}{\delta}\bigg(\int_0^{+\infty} \hat u\big((t+\delta+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\eta(s)\,ds - \int_0^{+\infty} \hat u\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\eta(s)\,ds\bigg) \\
&= \frac{1}{\delta}\bigg(\int_\delta^{+\infty} \hat u\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\eta(s-\delta)\,ds - \int_0^{+\infty} \hat u\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\eta(s)\,ds\bigg) \\
&= \int_0^{+\infty} \hat u\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\frac{\eta(s-\delta)-\eta(s)}{\delta}\,ds \\
&\quad - \frac{1}{\delta}\int_0^\delta \hat u\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\eta(s-\delta)\,ds.
\end{aligned}$$

Notice that

$$\begin{aligned}
&\int_0^{+\infty} \hat u\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\frac{\eta(s-\delta)-\eta(s)}{\delta}\,ds \\
&\qquad\qquad \overset{\delta\to0^+}{\longrightarrow}\; -\int_0^{+\infty} \hat u\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\eta'(s)\,ds,
\end{aligned}$$

where $\eta'(s)$ denotes the first-order derivative of η at s, and

$$\frac{1}{\delta}\int_0^\delta \hat u\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\eta(s-\delta)\,ds \;\overset{\delta\to0^+}{\longrightarrow}\; \hat u\big(t,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t)\big)\,\eta(0) \;=\; 0,$$

where we have used that $\eta(0) = 0$. This proves that the horizontal derivative of $\hat{v}$ exists everywhere on $[0,T) \times C([0,T];\mathbb{R}^d)$ and is continuous.

Let us now consider the vertical derivatives of $\hat{v}$. Given $(t,\boldsymbol{x},\mathbf{y}) \in \boldsymbol{\Lambda} \times \mathbb{R}^d$, $h \in \mathbb{R}\backslash\{0\}$, and $i = 1,\ldots,d$, we have

$$\begin{aligned}&\frac{\hat{v}(t,\boldsymbol{x},\mathbf{y}+h\mathbf{e}_i) - \hat{v}(t,\boldsymbol{x},\mathbf{y})}{h}\\ &\quad= \int_0^{+\infty} \frac{\hat{u}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}+h\mathbf{e}_i\big) - \hat{u}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}\big)}{h}\,\eta(s)\,ds.\end{aligned}$$

Notice that the map $f_{\hat{u}}^{t,\boldsymbol{x},\mathbf{y}}\colon [0,+\infty)\times\mathbb{R}\to\mathbb{R}$, defined as

$$f_{\hat{u}}^{t,\boldsymbol{x},\mathbf{y}}(s,h) := \begin{cases} \dfrac{\hat{u}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}+h\mathbf{e}_i\big) - \hat{u}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}\big)}{h}, & h\neq 0,\\ \partial_{x_i}^V \hat{u}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}\big), & h=0,\end{cases}$$

is continuous. Then, by the Lebesgue dominated convergence theorem, we obtain

$$\begin{aligned}&\int_0^{+\infty} \frac{\hat{u}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}+h\mathbf{e}_i\big) - \hat{u}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}\big)}{h}\,\eta(s)\,ds\\ &\qquad\qquad \overset{h\to 0}{\longrightarrow} \int_0^{+\infty} \partial_{x_i}^V \hat{u}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}\big)\,\eta(s)\,ds.\end{aligned}$$

This proves that $\hat{v}$ admits continuous first-order vertical derivatives. In a similar way we can prove that $\hat{v}$ also admits continuous second-order vertical derivatives. We conclude that $\hat{v} \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$. □

We now apply Lemma 3.2 to the map $\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}/(1+\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)})$ introduced in Lemma 3.1. We apply it to such a map in order to have bounded derivatives (see item 2) of Lemma 3.3). To make things clear, we fix the function η appearing in Lemma 3.2, taking η to be the probability density function of the chi-squared distribution with three degrees of freedom. Clearly, the claim of Lemma 3.2 holds true if we replace such an η by any other function η satisfying the properties stated in Lemma 3.2.

Lemma 3.3. *Let $\eta\colon [0,+\infty)\to\mathbb{R}$ be given by*

$$\eta(s) := \sqrt{\frac{s}{2\pi}}\,\mathrm{e}^{-\frac{1}{2}s}, \qquad \forall\, s\geq 0. \tag{3.17}$$

For every fixed $(t_0,\boldsymbol{x}_0)\in\boldsymbol{\Lambda}$, let $\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}$ be the map defined in Lemma 3.1 and define the map $\hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}\colon \boldsymbol{\Lambda}\times\mathbb{R}^d\to[0,+\infty)$ as

$$\hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y}) := \int_0^{+\infty} \frac{\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}\big)}{1+\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T,\boldsymbol{x}(\cdot\wedge t),\mathbf{y}\big)}\,\eta(s)\,ds, \qquad \forall\,(t,\boldsymbol{x},\mathbf{y})\in\boldsymbol{\Lambda}\times\mathbb{R}^d.$$

Moreover, let $\chi_\infty^{(t_0,\boldsymbol{x}_0)}\colon \boldsymbol{\Lambda}\to[0,+\infty)$ be given by

$$\chi_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x}) := \hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\boldsymbol{x}(t)), \qquad \forall\,(t,\boldsymbol{x})\in\boldsymbol{\Lambda}. \tag{3.18}$$

Then, the following properties hold.

1) $\hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)} \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$.

2) *The horizontal derivative of* $\hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}$ *is bounded by the constant* $\sqrt{\frac{2}{\pi \mathrm{e}}}$ $\big(= \int_0^{+\infty} |\eta'(s)|\,ds\big)$*; the first-order vertical derivatives of* $\hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}$ *are bounded by the constant* 1*; the second-order vertical derivatives of* $\hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}$ *are bounded by the constant* $\sqrt{\frac{2}{\pi}} + 2$.
3) *For every* $(t,\boldsymbol{x}) \in \boldsymbol{\Lambda}$,

$$\alpha_d\,\frac{\|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty^{d+1} \wedge \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty}{1 + \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty} \leq \chi_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x}) \leq \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}_0(\cdot\wedge t_0)\|_\infty \wedge 1, \tag{3.19}$$

with the same constant α_d *as in* (3.2).

Proof. We begin noting that the function η given by (3.17) satisfies all the properties required in the statement of Lemma 3.2. Then, item 1) follows directly from Lemma 3.2 and the fact that, by item 1) of Lemma 3.1, the map

$$(t,\boldsymbol{x},\mathbf{y}) \mapsto \frac{\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y})}{1 + \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y})}$$

belongs to $\boldsymbol{C}^{0,2}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$.

Let us now prove item 2). By the proof of Lemma 3.2, we have

$$\partial_t^H \hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y}) = -\int_0^{+\infty} \frac{\hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)}{1 + \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)}\,\eta'(s)\,ds \leq \int_0^{+\infty} |\eta'(s)|\,ds.$$

Concerning the first-order vertical derivatives, for every $i = 1,\ldots,d$, we have

$$\partial_{x_i}^V \hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y}) = \int_0^{+\infty} \frac{\partial_{x_i}^V \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)}{\big(1 + \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)\big)^2}\,\eta(s)\,ds.$$

Recalling from item 2) in Lemma 3.1 that $|\partial_{x_i}^V \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}|$ is bounded by the constant 1, it is easy to see that $|\partial_{x_i}^V \hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y})| \leq \int_0^{+\infty}\eta(s)\,ds$, from which the claim follows. On the other hand, regarding the second-order vertical derivatives, for every $i,j = 1,\ldots,d$,

$$\partial_{x_i x_j}^V \hat{\chi}_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x},\mathbf{y}) = \int_0^{+\infty} \frac{\partial_{x_i x_j}^V \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)}{\big(1 + \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)\big)^2}\,\eta(s)\,ds$$
$$- 2\int_0^{+\infty} \frac{\partial_{x_i}^V \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)\,\partial_{x_j}^V \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)}{\big(1 + \hat{\kappa}_\infty^{(t_0,\boldsymbol{x}_0)}\big((t+s)\wedge T, \boldsymbol{x}(\cdot\wedge t), \mathbf{y}\big)\big)^3}\,\eta(s)\,ds,$$

from which it is easy to see that the claim follows. Finally, item 3) is a direct consequence of the two inequalities in (3.2). □

In conclusion, by Lemma 3.3 it follows that the map $\rho_\infty\colon \boldsymbol{\Lambda}\times\boldsymbol{\Lambda} \to [0,+\infty)$ given by

$$\rho_\infty\big((t,\boldsymbol{x}),(t_0,\boldsymbol{x}_0)\big) = |t - t_0|^2 + \chi_\infty^{(t_0,\boldsymbol{x}_0)}(t,\boldsymbol{x}), \qquad \forall\,(t,\boldsymbol{x}),(t_0,\boldsymbol{x}_0) \in \boldsymbol{\Lambda}, \tag{3.20}$$

with χ_∞ as in (3.18), is the claimed gauge-type function, smooth as a map of the first pair, namely $(t,\boldsymbol{x}) \mapsto \rho_\infty((t,\boldsymbol{x}),(t_0,\boldsymbol{x}_0))$, and with bounded derivatives.

4. Crandall-Lions (path-dependent) viscosity solutions

In the present section we consider the following second-order path-dependent partial differential equation:

$$\begin{cases} -\partial_t^H u(t,\boldsymbol{x}) + F\big(t,\boldsymbol{x},u(t,\boldsymbol{x}),\partial_{\boldsymbol{x}}^V u(t,\boldsymbol{x}),\partial_{\boldsymbol{x}\boldsymbol{x}}^V u(t,\boldsymbol{x})\big) = 0, & (t,\boldsymbol{x}) \in [0,T) \times C([0,T];\mathbb{R}^d), \\ u(T,\boldsymbol{x}) = \xi(\boldsymbol{x}), & \boldsymbol{x} \in C([0,T];\mathbb{R}^d). \end{cases} \tag{4.1}$$

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

Definition 4.1. We denote by $\boldsymbol{C}_{\mathrm{pol}}^{1,2}(\boldsymbol{\Lambda})$ the set of $\varphi \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$ such that φ, $\partial_t^H\varphi$, $\partial_{\boldsymbol{x}}^V\varphi$, $\partial_{\boldsymbol{x}\boldsymbol{x}}^V\varphi$ satisfy a polynomial growth condition.

Definition 4.2. We say that an upper semicontinuous map $u\colon \boldsymbol{\Lambda} \to \mathbb{R}$ is a (**path-dependent**) **viscosity subsolution** of equation (4.1) if the following holds.

- $u(T,\boldsymbol{x}) \le \xi(\boldsymbol{x})$, for all $\boldsymbol{x} \in C([0,T];\mathbb{R}^d)$;
- for any $(t,\boldsymbol{x}) \in [0,T) \times C([0,T];\mathbb{R}^d)$ and $\varphi \in \boldsymbol{C}_{\mathrm{pol}}^{1,2}(\boldsymbol{\Lambda})$, satisfying
$$(u-\varphi)(t,\boldsymbol{x}) = \sup_{(t',\boldsymbol{x}')\in\boldsymbol{\Lambda}} (u-\varphi)(t',\boldsymbol{x}'),$$
we have
$$-\partial_t^H\varphi(t,\boldsymbol{x}) + F\big(t,\boldsymbol{x},u(t,\boldsymbol{x}),\partial_{\boldsymbol{x}}^V\varphi(t,\boldsymbol{x}),\partial_{\boldsymbol{x}\boldsymbol{x}}^V\varphi(t,\boldsymbol{x})\big) \ \le \ 0.$$

We say that a lower semicontinuous map $u\colon \boldsymbol{\Lambda} \to \mathbb{R}$ is a (**path-dependent**) **viscosity supersolution** of equation (4.1) if:

- $u(T,\boldsymbol{x}) \ge \xi(\boldsymbol{x})$, for all $\boldsymbol{x} \in C([0,T];\mathbb{R}^d)$;
- for any $(t,\boldsymbol{x}) \in [0,T) \times C([0,T];\mathbb{R}^d)$ and $\varphi \in \boldsymbol{C}_{\mathrm{pol}}^{1,2}(\boldsymbol{\Lambda})$, satisfying:
$$(u-\varphi)(t,\boldsymbol{x}) = \inf_{(t',\boldsymbol{x}')\in\boldsymbol{\Lambda}} (u-\varphi)(t',\boldsymbol{x}'),$$
we have
$$-\partial_t^H\varphi(t,\boldsymbol{x}) + F\big(t,\boldsymbol{x},u(t,\boldsymbol{x}),\partial_{\boldsymbol{x}}^V\varphi(t,\boldsymbol{x}),\partial_{\boldsymbol{x}\boldsymbol{x}}^V\varphi(t,\boldsymbol{x})\big) \ \ge \ 0.$$

We say that a continuous map $u\colon \boldsymbol{\Lambda} \to \mathbb{R}$ is a (**path-dependent**) **viscosity solution** of equation (4.1) if u is both a (path-dependent) viscosity subsolution and a (path-dependent) viscosity supersolution of (4.1).

4.1. Path-dependent heat equation.

In the present section we focus on the path-dependent heat equation, namely when $F(t,\boldsymbol{x},r,p,M) = -\frac{1}{2}\mathrm{tr}[M]$:

$$\begin{cases} -\partial_t^H u(t,\boldsymbol{x}) - \dfrac{1}{2}\mathrm{tr}\big[\partial_{\boldsymbol{x}\boldsymbol{x}}^V u(t,\boldsymbol{x})\big] \ = \ 0, & (t,\boldsymbol{x}) \in [0,T) \times C([0,T];\mathbb{R}^d), \\ u(T,\boldsymbol{x}) \ = \ \xi(\boldsymbol{x}), & \boldsymbol{x} \in C([0,T];\mathbb{R}^d). \end{cases} \tag{4.2}$$

In the sequel we denote

$$\mathcal{L}u(t,\boldsymbol{x}) \ := \ \partial_t^H u(t,\boldsymbol{x}) + \frac{1}{2}\mathrm{tr}[\partial_{\boldsymbol{x}\boldsymbol{x}}^V u(t,\boldsymbol{x})]. \tag{4.3}$$

We impose the following assumption on the terminal condition ξ.

(A) *The function $\xi\colon C([0,T];\mathbb{R}^d) \to \mathbb{R}$ is continuous and bounded.*

Remark 4.1. The boundedness of ξ will be used in the proof of Theorem 4.2 (comparison). On the other hand, the proof that the function v in (4.4) is continuous and is a viscosity solution of equation (4.2) (see the proof of Theorem 4.1) holds under weaker growth condition on ξ (for instance, ξ having polynomial growth).

4.1.1. *Existence.* The "candidate solution" to equation (4.2) is

$$v(t,\boldsymbol{x}) \;:=\; \mathbb{E}\big[\xi(\boldsymbol{W}^{t,\boldsymbol{x}})\big], \qquad \text{for all } (t,\boldsymbol{x}) \in \boldsymbol{\Lambda}, \tag{4.4}$$

where $\boldsymbol{W} = (\boldsymbol{W}_s)_{s\in[0,T]}$ is a d-dimensional Brownian motion on some probability space $(\Omega,\mathcal{F},\mathbb{P})$, and the stochastic process $\boldsymbol{W}^{t,\boldsymbol{x}} = (\boldsymbol{W}_s^{t,\boldsymbol{x}})_{s\in[0,T]}$ is given by

$$\boldsymbol{W}_s^{t,\boldsymbol{x}} \;:=\; \begin{cases} \boldsymbol{x}(s), & s \le t, \\ \boldsymbol{x}(t) + \boldsymbol{W}_s - \boldsymbol{W}_t, & s > t. \end{cases} \tag{4.5}$$

Theorem 4.1. *Under Assumption* **(A)***, the function v in* (4.4) *is continuous and bounded. Moreover, v is a (path-dependent) viscosity solution of equation* (4.2).

Proof. STEP I. *Continuity of v.* Given $(t,\boldsymbol{x}),(t',\boldsymbol{x}') \in \boldsymbol{\Lambda}$, with $t \le t'$, we have from (4.5):

$$\boldsymbol{W}_s^{t,\boldsymbol{x}} - \boldsymbol{W}_s^{t',\boldsymbol{x}'} \;=\; \begin{cases} \boldsymbol{x}(s) - \boldsymbol{x}'(s), & s \le t, \\ \boldsymbol{x}(t) - \boldsymbol{x}'(s) + \boldsymbol{W}_s - \boldsymbol{W}_t, & t < s \le t', \\ \boldsymbol{x}(t) - \boldsymbol{x}'(t') + \boldsymbol{W}_{t'} - \boldsymbol{W}_t, & s > t'. \end{cases}$$

Hence

$$\begin{aligned} \sup_{s\in[0,T]} \big|\boldsymbol{W}_s^{t,\boldsymbol{x}} - \boldsymbol{W}_s^{t',\boldsymbol{x}'}\big| \;&\le\; \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}'(\cdot\wedge t')\|_\infty + \sup_{s\in[t,t']} \big|\boldsymbol{W}_s - \boldsymbol{W}_t\big| \\ &\le\; \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}'(\cdot\wedge t')\|_\infty + \sum_{i=1}^d \sup_{s\in[t,t']} \big|W_s^i - W_t^i\big|, \end{aligned}$$

where $\boldsymbol{W} = (W^1,\ldots,W^d)$ and the second inequality follows from the fact the Euclidean norm on $\mathbb{R}^d$ is estimated by the 1-norm. By the reflection principle, $\sup_{s\in[t,t']}|W_s^i - W_t^i|$ has the same law as $|W_{t'}^i - W_t^i|$, therefore

$$\begin{aligned} \mathbb{E}\Big[\sup_{s\in[0,T]} \big|\boldsymbol{W}_s^{t,\boldsymbol{x}} - \boldsymbol{W}_s^{t',\boldsymbol{x}'}\big|\Big] \;&\le\; \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}'(\cdot\wedge t')\|_\infty + \sum_{i=1}^d \mathbb{E}\big[|W_{t'}^i - W_t^i|\big] \\ &=\; \|\boldsymbol{x}(\cdot\wedge t) - \boldsymbol{x}'(\cdot\wedge t')\|_\infty + d\sqrt{\frac{2}{\pi}}\sqrt{|t-t'|}. \end{aligned}$$

Then, since ξ is bounded and continuous, the continuity of v follows from the above estimate together with the Lebesgue dominated convergence theorem.

STEP II. *v is a viscosity solution of equation* (4.2). For every $t \in [0,T]$, let $\mathbb{F}^t = (\mathcal{F}_s^t)_{s\in[t,T]}$ be the filtration given by: $\mathcal{F}_s^t := \sigma(\boldsymbol{W}_r - \boldsymbol{W}_t, r \in [t,s])$, for all $s \in [t,T]$. Now, fix $(t,\boldsymbol{x}) \in \boldsymbol{\Lambda}$ and $t' \in [t,T]$. We first prove that the following formula holds:

$$v(t,\boldsymbol{x}) \;=\; \mathbb{E}\big[v\big(t',\boldsymbol{W}^{t,\boldsymbol{x}}\big)\big]. \tag{4.6}$$

To this end, we begin noting that by (4.5) we have

$$\boldsymbol{W}_\cdot^{t,\boldsymbol{x}} \;=\; \boldsymbol{x}(\cdot\wedge t) + \boldsymbol{W}_{\cdot\vee t} - \boldsymbol{W}_t. \tag{4.7}$$

Therefore

$$v(t,\boldsymbol{x}) \;=\; \mathbb{E}[\xi(\boldsymbol{x}(\cdot\wedge t) + \boldsymbol{W}_{\cdot\vee t} - \boldsymbol{W}_t)]. \tag{4.8}$$

Now, notice that, by (4.7),

$$\boldsymbol{W}_\cdot^{t',\boldsymbol{W}^{t,\boldsymbol{x}}} \;=\; \boldsymbol{W}_{\cdot\wedge t'}^{t,\boldsymbol{x}} + \boldsymbol{W}_{\cdot\vee t'} - \boldsymbol{W}_{t'} \;=\; \boldsymbol{W}_\cdot^{t,\boldsymbol{x}}.$$

This proves the flow property $\boldsymbol{W}_\cdot^{t,\boldsymbol{x}} = \boldsymbol{W}_\cdot^{t',\boldsymbol{W}^{t,\boldsymbol{x}}}$. Then, by the freezing lemma for conditional expectation and formula (4.8), we obtain

$$\begin{aligned} v(t,\boldsymbol{x}) \;&=\; \mathbb{E}\big[\xi(\boldsymbol{W}^{t,\boldsymbol{x}})\big] \;=\; \mathbb{E}\big[\xi\big(\boldsymbol{W}^{t',\boldsymbol{W}^{t,\boldsymbol{x}}}\big)\big] \;=\; \mathbb{E}\big[\xi\big(\boldsymbol{W}_{\cdot\wedge t'}^{t,\boldsymbol{x}} + \boldsymbol{W}_{\cdot\vee t'} - \boldsymbol{W}_{t'}\big)\big] \\ &=\; \mathbb{E}\big[\mathbb{E}\big[\xi\big(\boldsymbol{W}_{\cdot\wedge t'}^{t,\boldsymbol{x}} + \boldsymbol{W}_{\cdot\vee t'} - \boldsymbol{W}_{t'}\big)\big|\mathcal{F}_{t'}^t\big]\big] \;=\; \mathbb{E}\big[v\big(t',\boldsymbol{W}_{\cdot\wedge t'}^{t,\boldsymbol{x}}\big)\big]. \end{aligned}$$

Finally, recalling that v is non-anticipative we deduce that $v(t',\boldsymbol{W}_{\cdot\wedge t'}^{t,\boldsymbol{x}}) = v(t',\boldsymbol{W}^{t,\boldsymbol{x}})$. This concludes the proof of formula (4.6).

Let us now prove that v is a viscosity solution of equation (4.2). We only prove the viscosity subsolution property, as the supersolution property can be proved in a similar way. We proceed along the same lines as in the proof of the subsolution property in Theorem 3.66 of [37]. Let $(t,\boldsymbol{x}) \in [0,T) \times C([0,T];\mathbb{R}^d)$ and $\varphi \in \boldsymbol{C}_{\mathrm{pol}}^{1,2}(\boldsymbol{\Lambda})$, satisfying:

$$(v-\varphi)(t,\boldsymbol{x}) = \sup_{(t',\boldsymbol{x}')\in\boldsymbol{\Lambda}} (v-\varphi)(t',\boldsymbol{x}').$$

We suppose that $(v-\varphi)(t,\boldsymbol{x}) = 0$ (if this is not the case, we replace φ by $\psi(\cdot,\cdot) := \varphi(\cdot,\cdot) + v(t,\boldsymbol{x}) - \varphi(t,\boldsymbol{x})$). Take

$$\varphi(t,\boldsymbol{x}) \;=\; v(t,\boldsymbol{x}) \;=\; \mathbb{E}\big[v\big(t+\varepsilon,\boldsymbol{W}^{t,\boldsymbol{x}}\big)\big] \;\leq\; \mathbb{E}\big[\varphi\big(t+\varepsilon,\boldsymbol{W}^{t,\boldsymbol{x}}\big)\big], \tag{4.9}$$

where the last inequality follows from the fact that $\sup(v-\varphi) = 0$, so that $v \leq \varphi$ on $\boldsymbol{\Lambda}$. Notice that the last expectation in (4.9) is finite, as φ has polynomial growth. Now, by the functional Itô formula (2.1), we have

$$\varphi(t+\varepsilon,\boldsymbol{W}^{t,\boldsymbol{x}}) \;=\; \varphi(t,\boldsymbol{x}) + \int_t^{t+\varepsilon} \mathcal{L}\varphi(s,\boldsymbol{W}^{t,\boldsymbol{x}})\,ds + \sum_{i=1}^d \int_t^{t+\varepsilon} \partial_{x_i}^V \varphi(s,\boldsymbol{W}^{t,\boldsymbol{x}})\,dW_s^i,$$

where $\mathcal{L}$ was defined in (4.3). Since $\partial_{x_i}^V\varphi$ has polynomial growth, the corresponding stochastic integral is a martingale. Then, plugging the above formula into (4.9) and dividing by ε, we find

$$-\,\mathbb{E}\bigg[\frac{1}{\varepsilon}\int_t^{t+\varepsilon} \mathcal{L}\varphi(s,\boldsymbol{W}^{t,\boldsymbol{x}})\,ds\bigg] \;\leq\; 0.$$

Letting $\varepsilon \to 0^+$, we conclude that

$$-\mathcal{L}\varphi(t,\boldsymbol{x}) \;\leq\; 0,$$

which proves the viscosity subsolution property. □

4.1.2. *Comparison theorem.*

Theorem 4.2. *Suppose that Assumption* **(A)** *holds. Let $u,w\colon \boldsymbol{\Lambda} \to \mathbb{R}$ be respectively upper and lower semicontinuous, satisfying*

$$\sup u \;<\; +\infty, \qquad\qquad \inf w \;>\; -\infty.$$

Suppose that u (resp. w) is a (path-dependent) viscosity subsolution (resp. supersolution) of equation (4.2). Then $u \leq w$ on $\boldsymbol{\Lambda}$.

Proof. The proof consists in showing that $u \leq v$ and $v \leq w$ on $\boldsymbol{\Lambda}$ (with v given by (4.4)), from which we immediately deduce the claim. In what follows, we only report the proof of the inequality $u \leq v$, as the other inequality (that is $v \leq w$) can be deduced from the first one replacing u, v, ξ with $-w$, $-v$, $-\xi$, respectively.

We proceed by contradiction and assume that $\sup(u-v)>0$. Then, there exists $(t_0,\boldsymbol{x}_0)\in\boldsymbol{\Lambda}$ such that

$$(u-v)(t_0,\boldsymbol{x}_0) \ > \ 0.$$

Notice that $t_0<T$, since $u(T,\cdot)\leq\xi(\cdot)=v(T,\cdot)$. We split the rest of the proof into five steps.

STEP I. Let $\{\xi_N\}_N$ be the sequence given by Lemma C.2. Since ξ is bounded, we have that ξ_N is bounded uniformly with respect to N. Now, denote

$$v_N(t,\boldsymbol{x}) \ := \ \mathbb{E}\big[\xi_N(\boldsymbol{W}^{t,\boldsymbol{x}})\big], \qquad \text{for all } (t,\boldsymbol{x})\in\boldsymbol{\Lambda}.$$

Then, v_N is bounded uniformly with respect to N. Moreover, by Lemma C.1 it follows that, for every N, $v_N\in\boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$ and is a classical (smooth) solution of equation (4.2) with terminal condition ξ_N. Finally, recalling from Lemma C.2 that $\{\xi_N\}_N$ converges pointwise to ξ as $N\to+\infty$, it follows from the Lebesgue dominated convergence theorem that $\{v_N\}_N$ converges pointwise to v as $N\to+\infty$. Then, we notice that there exists $N_0\in\mathbb{N}$ such that

$$(u-v_{N_0})(t_0,\boldsymbol{x}_0) \ > \ 0. \tag{4.10}$$

We also suppose that (possibly enlarging N_0)

$$|\xi(\boldsymbol{x}_0)-\xi_{N_0}(\boldsymbol{x}_0)| \ \leq \ \frac{1}{2}(u-v_{N_0})(t_0,\boldsymbol{x}_0). \tag{4.11}$$

STEP II. For every $\lambda>0$, we set

$$u^\lambda(t,\boldsymbol{x}) := \mathrm{e}^{\lambda t}\,u(t,\boldsymbol{x}), \quad \xi^\lambda(\boldsymbol{x}) := \mathrm{e}^{\lambda T}\,\xi(\boldsymbol{x}), \quad v_{N_0}^\lambda(t,\boldsymbol{x}) := \mathrm{e}^{\lambda t}\,v_{N_0}(t,\boldsymbol{x}), \quad \xi_{N_0}^\lambda(\boldsymbol{x}) := \mathrm{e}^{\lambda T}\,\xi_{N_0}(\boldsymbol{x}).$$

for all $(t,\boldsymbol{x})\in\boldsymbol{\Lambda}$. Notice that u^λ is a (path-dependent) viscosity subsolution of the path-dependent partial differential equation

$$\begin{cases}-\partial_t^H u^\lambda(t,\boldsymbol{x})-\frac{1}{2}\mathrm{tr}[\partial_{\boldsymbol{x}}^V u^\lambda(t,\boldsymbol{x})]+\lambda\,u^\lambda(t,\boldsymbol{x}) \ = \ 0, & (t,\boldsymbol{x})\in[0,T)\times C([0,T];\mathbb{R}^d),\\ u^\lambda(T,\boldsymbol{x}) \ = \ \xi^\lambda(\boldsymbol{x}), & \boldsymbol{x}\in C([0,T];\mathbb{R}^d).\end{cases} \tag{4.12}$$

Similarly, $v_{N_0}^\lambda$ is a classical (smooth) solution of equation (4.12) with ξ^λ replaced by $\xi_{N_0}^\lambda$. We finally notice that by (4.10) we have

$$(u^\lambda-v_{N_0}^\lambda)(t_0,\boldsymbol{x}_0) \ > \ 0.$$

So, in particular,

$$\sup(u^\lambda-v_{N_0}^\lambda)-\varepsilon \ = \ (u^\lambda-v_{N_0}^\lambda)(t_0,\boldsymbol{x}_0) \ \leq \ \sup(u^\lambda-v_{N_0}^\lambda), \tag{4.13}$$

where

$$\varepsilon \ := \ \sup(u^\lambda-v_{N_0}^\lambda)-(u^\lambda-v_{N_0}^\lambda)(t_0,\boldsymbol{x}_0).$$

STEP III. Notice that $u^\lambda-v_{N_0}^\lambda$ is upper semicontinuous and bounded from above. Then, by (4.13) and Corollary D.1 with $G=u^\lambda-v_{N_0}^\lambda$, we deduce that for every $\delta>0$ there exists a sequence $\{(t_n,\boldsymbol{x}_n)\}_{n\geq1}\subset\boldsymbol{\Lambda}$ converging to some $(\bar t,\bar{\boldsymbol{x}})\in\boldsymbol{\Lambda}$ (possibly depending on $\varepsilon,\delta,\lambda,N_0$) such that:

i) $\rho_\infty((t_n,\boldsymbol{x}_n),(\bar t,\bar{\boldsymbol{x}}))\leq\frac{\varepsilon}{2^n\delta}$, for every $n\geq0$, where ρ_∞ is the smooth gauge-type function with bounded derivatives defined by (3.20).
ii) $(u^\lambda-v_{N_0}^\lambda)(t_0,\boldsymbol{x}_0)\leq\big(u^\lambda-(v_{N_0}^\lambda+\delta\varphi_\varepsilon)\big)(\bar t,\bar{\boldsymbol{x}})$, where

$$\varphi_\varepsilon(t,\boldsymbol{x}) \ := \ \sum_{n=0}^{+\infty}\frac{1}{2^n}\,\rho_\infty\big((t,\boldsymbol{x}),(t_n,\boldsymbol{x}_n)\big) \qquad \forall\,(t,\boldsymbol{x})\in\boldsymbol{\Lambda}.$$

iii) It holds that

$$\big(u^\lambda - (v_{N_0}^\lambda + \delta\varphi_\varepsilon)\big)(\bar t, \bar{\boldsymbol{x}}) \;=\; \sup_{(t,\boldsymbol{x})\in\boldsymbol{\Lambda}} \big(u^\lambda - (v_{N_0}^\lambda + \delta\varphi_\varepsilon)\big)(t, \boldsymbol{x}). \tag{4.14}$$

We also recall from Corollary D.1 that φ_ε satisfies the following properties.

1) $\varphi_\varepsilon \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$ and is bounded.
2) $\big|\partial_t^H \varphi_\varepsilon(t,\boldsymbol{x})\big| \le 2\Big(2\,T + \sqrt{\frac{2}{\pi \mathrm{e}}}\Big)$, for every $(t,\boldsymbol{x}) \in [0,T)\times C([0,T];\mathbb{R}^d)$.
3) For every $i,j = 1,\ldots,d$, $\partial_{x_i}^V\varphi_\varepsilon$ is bounded by the constant 2 and $\partial_{x_ix_j}^V\varphi_\varepsilon$ is bounded by the constant $2\Big(\sqrt{\frac{2}{\pi}} + 2\Big)$.

In particular, $\varphi_\varepsilon \in \boldsymbol{C}_{\mathrm{pol}}^{1,2}(\boldsymbol{\Lambda})$.

STEP IV. We prove below that $\bar t < T$. As a matter of fact, by item ii) of STEP III we have

$$\big(u^\lambda - (v_{N_0}^\lambda + \delta\varphi_\varepsilon)\big)(\bar t, \bar{\boldsymbol{x}}) \;\ge\; (u^\lambda - v_{N_0}^\lambda)(t_0, \boldsymbol{x}_0). \tag{4.15}$$

On the other hand, if $\bar t = T$ we obtain

$$\big(u^\lambda - (v_{N_0}^\lambda + \delta\varphi_\varepsilon)\big)(\bar t, \bar{\boldsymbol{x}}) \;=\; \mathrm{e}^{\lambda T}(\xi(\bar{\boldsymbol{x}}) - \xi_{N_0}(\bar{\boldsymbol{x}})) - \delta\varphi_\varepsilon(T,\bar{\boldsymbol{x}}) \;\le\; \mathrm{e}^{\lambda T}(\xi(\bar{\boldsymbol{x}}) - \xi_{N_0}(\bar{\boldsymbol{x}})), \tag{4.16}$$

where the latter inequality comes from the fact that $\varphi_\varepsilon \ge 0$. Hence, by (4.15) and (4.16) we get

$$\mathrm{e}^{\lambda t_0}(u - v_{N_0})(t_0, \boldsymbol{x}_0) \;\le\; \mathrm{e}^{\lambda T}(\xi(\bar{\boldsymbol{x}}) - \xi_{N_0}(\bar{\boldsymbol{x}})).$$

Letting $\varepsilon \to 0$, it follows from item i) above with $n = 0$ and (3.19) that $d_\infty((\bar t, \bar{\boldsymbol{x}}), (t_0, \boldsymbol{x}_0)) \to 0$. Therefore, letting $\varepsilon \to 0$ in the previous inequality, we obtain

$$\mathrm{e}^{\lambda t_0}(u - v_{N_0})(t_0, \boldsymbol{x}_0) \;\le\; \mathrm{e}^{\lambda T}(\xi(\boldsymbol{x}_0) - \xi_{N_0}(\boldsymbol{x}_0)).$$

By (4.11), we end up with

$$\mathrm{e}^{\lambda t_0} \;\le\; \frac{1}{2}\,\mathrm{e}^{\lambda T}.$$

Letting $\lambda \to 0$, we find a contradiction.

STEP V. Here again $\lambda > 0$ is fixed. By (4.14) and the definition of viscosity subsolution of (4.12) applied to u^λ at the point $(\bar t, \bar{\boldsymbol{x}})$ with test function $v_{N_0}^\lambda + \delta\varphi_\varepsilon$, we obtain

$$-\mathcal{L}(v_{N_0}^\lambda + \delta\varphi_\varepsilon)(\bar t, \bar{\boldsymbol{x}}) + \lambda\, u^\lambda(\bar t, \bar{\boldsymbol{x}}) \;\le\; 0.$$

Recalling that $v_{N_0}^\lambda$ is a classical (smooth) solution of equation (4.12) with ξ^λ replaced by $\xi_{N_0}^\lambda$, we find

$$\lambda\,(u^\lambda - v_{N_0}^\lambda)(\bar t, \bar{\boldsymbol{x}}) \;\le\; \delta\,\mathcal{L}\varphi_\varepsilon(\bar t, \bar{\boldsymbol{x}}).$$

By item ii) in STEP III (namely (4.15)), subtracting from both sides the quantity $\lambda\delta\varphi_\varepsilon(\bar t, \bar{\boldsymbol{x}})$, we obtain

$$\lambda\,(u^\lambda - v_{N_0}^\lambda)(t_0, \boldsymbol{x}_0) \;\le\; \lambda\,\big(u^\lambda - (v_{N_0}^\lambda + \delta\varphi_\varepsilon)\big)(\bar t, \bar{\boldsymbol{x}}) \;\le\; \delta\,\mathcal{L}\varphi_\varepsilon(\bar t, \bar{\boldsymbol{x}}) - \lambda\,\delta\,\varphi_\varepsilon(\bar t, \bar{\boldsymbol{x}}).$$

Recalling that $\varphi_\varepsilon \ge 0$, we see that

$$\lambda\,(u^\lambda - v_{N_0}^\lambda)(t_0, \boldsymbol{x}_0) \;\le\; \lambda\,\big(u^\lambda - (v_{N_0}^\lambda + \delta\varphi_\varepsilon)\big)(\bar t, \bar{\boldsymbol{x}}) \;\le\; \delta\,\mathcal{L}\varphi_\varepsilon(\bar t, \bar{\boldsymbol{x}}).$$

From items 2) and 3) above, it follows that $\mathcal{L}\varphi_\varepsilon(\bar t, \bar{\boldsymbol{x}})$ is bounded by a constant (not depending on ε, δ, λ). Therefore, letting $\delta \to 0^+$, taking into account the notations of STEP II, we have

$$\lambda\,\mathrm{e}^{\lambda t_0}\,(u - v_{N_0})(t_0, \boldsymbol{x}_0) \;= \lambda\,(u^\lambda - v_{N_0}^\lambda)(t_0, \boldsymbol{x}_0) \;\le\; 0,$$

which gives a contradiction to (4.10). □

Appendix A.
Functional Itô's formula for lifted maps

We start with a technical result.

Lemma A.1. *Let* $\hat{u} \in \boldsymbol{C}^{1,0}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$. *Then, for every* $(t, \boldsymbol{x}) \in [0,T) \times C([0,T];\mathbb{R}^d) \times \mathbb{R}^d$, $\varepsilon \in (0, T-t)$, *the map* $\phi\colon [0,(T-t)/\varepsilon) \to \mathbb{R}$, *defined as*

$$\phi(a) \ := \ \hat{u}(t + a\varepsilon, \boldsymbol{x}(\cdot \wedge t), \boldsymbol{x}(t)), \qquad \forall\, a \in [0,(T-t)/\varepsilon),$$

is in $C^1([0,(T-t)/\varepsilon))$ *and*

$$\phi'(a) \ = \ \varepsilon\, \partial_t^H \hat{u}(t + a\varepsilon, \boldsymbol{x}(\cdot \wedge t)), \qquad \forall\, a \in [0,(T-t)/\varepsilon).$$

Proof. Let $a \in [0,(T-t)/\varepsilon)$. We have, for any $\delta \in (0,(T-t)/\varepsilon - a)$,

$$\begin{aligned}\frac{\phi(a+\delta) - \phi(a)}{\delta} \ &= \ \frac{\hat{u}(t + (a+\delta)\varepsilon, \boldsymbol{x}(\cdot \wedge t), \boldsymbol{x}(t)) - \hat{u}(t + a\varepsilon, \boldsymbol{x}(\cdot \wedge t), \boldsymbol{x}(t))}{\delta} \\ &\underset{\delta \to 0^+}{\longrightarrow} \ \partial_t^H \hat{u}(t + a\varepsilon, \boldsymbol{x}(\cdot \wedge t)).\end{aligned}$$

This shows that ϕ is right-differentiable on $[0,(T-t)/\varepsilon)$ and that such a right-derivative is continuous on $[0,(T-t)/\varepsilon)$. Then, it follows for instance from Corollary 1.2, Chapter 2, in [73] that $\phi \in C^1([0,(T-t)/\varepsilon))$. □

We now introduce a special class of lifted maps, for which the proof of the functional Itô formula will be easier.

Definition A.1. A **strictly non-anticipative map (on** $\boldsymbol{\Lambda} \times \mathbb{R}^d$**)** is a non-anticipative map $\hat{u}\colon \boldsymbol{\Lambda} \times \mathbb{R}^d \to \mathbb{R}$ for which there exists $\delta_0 > 0$ such that, for all $(t, \boldsymbol{x}) \in \boldsymbol{\Lambda}$,

$$\hat{u}\big((t+\delta) \wedge T, \boldsymbol{x}, \boldsymbol{x}(t)\big) \ = \ \hat{u}\big((t+\delta) \wedge T, \boldsymbol{x}(\cdot \wedge t), \boldsymbol{x}(t)\big)$$

or, equivalently,

$$\hat{u}\big(t, \boldsymbol{x}, \boldsymbol{x}((t-\delta) \vee 0)\big) \ = \ \hat{u}\big(t, \boldsymbol{x}(\cdot \wedge ((t-\delta) \vee 0)), \boldsymbol{x}((t-\delta) \vee 0)\big),$$

for all $\delta \in [0, \delta_0]$.

Remark A.1. We give an example of delayed map. Let $\hat{u}\colon [0,T] \times C([0,T];\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ be a non-anticipative map. Take $\delta_0 > 0$ and define the map $\hat{u}_{\delta_0}\colon [0,T] \times C([0,T];\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}$ as

$$\hat{u}_{\delta_0}(t, \boldsymbol{x}, \mathbf{y}) \ := \ \hat{u}\big(t, \boldsymbol{x}(\cdot \wedge ((t-\delta_0) \vee 0)), \mathbf{y}\big), \qquad \forall\, (t, \boldsymbol{x}, \mathbf{y}) \in \boldsymbol{\Lambda} \times \mathbb{R}^d. \tag{A.1}$$

Then, $\hat{u}_{\delta_0}$ is a delayed map with delay δ_0, in the sense of Definition A.1.

Lemma A.2. *Let* $\hat{u} \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$ *and* $\delta_0 > 0$. *Let also* $\hat{u}_{\delta_0}$ *be given by* (A.1). *Then,* $\hat{u}_{\delta_0} \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$. *Moreover*

$$\begin{aligned}\partial_t^H \hat{u}_{\delta_0}(t, \boldsymbol{x}) \ &= \ \partial_t^H \hat{u}\big(t, \boldsymbol{x}(\cdot \wedge ((t-\delta_0) \vee 0))\big), \qquad && \forall\, (t, \boldsymbol{x}) \in [0,T) \times C([0,T];\mathbb{R}^d), \\ \partial_{\boldsymbol{x}}^V \hat{u}_{\delta_0}(t, \boldsymbol{x}, \mathbf{y}) \ &= \ \partial_{\boldsymbol{x}}^V \hat{u}\big(t, \boldsymbol{x}(\cdot \wedge ((t-\delta_0) \vee 0)), \mathbf{y}\big), && \forall\, (t, \boldsymbol{x}, \mathbf{y}) \in \boldsymbol{\Lambda} \times \mathbb{R}^d, \\ \partial_{\boldsymbol{x}\boldsymbol{x}}^V \hat{u}_{\delta_0}(t, \boldsymbol{x}, \mathbf{y}) \ &= \ \partial_{\boldsymbol{x}\boldsymbol{x}}^V \hat{u}\big(t, \boldsymbol{x}(\cdot \wedge ((t-\delta_0) \vee 0)), \mathbf{y}\big), && \forall\, (t, \boldsymbol{x}, \mathbf{y}) \in \boldsymbol{\Lambda} \times \mathbb{R}^d.\end{aligned}$$

Proof. The claim follows easily from the definitions of horizontal and vertical derivatives, see Definition 2.5. □

The functional Itô formula is proved extending to the path-dependent setting the approach of stochastic calculus via regularization. It is therefore useful to recall the following standard definitions of stochastic calculus via regularization (for more details we refer for instance to [83]).

Definition A.2. Let $X = (X_t)_{t\in[0,T]}$ and $Y = (Y_t)_{t\in[0,T]}$ be real-valued processes on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with X continuous. For every $\varepsilon > 0$, we denote

$$I_t^\varepsilon \;:=\; \int_0^t Y_s \frac{X_{(s+\varepsilon)\wedge T} - X_s}{\varepsilon}\, ds, \qquad \forall\, t \in [0,T].$$

If I^ε converges in the ucp sense as $\varepsilon \to 0^+$, we denote its limit by $\int_0^\cdot Y_s\, d^- X_s$ and call it the **forward integral**.

Definition A.3. Let $X = (X_t)_{t\in[0,T]}$ be a d-dimensional continuous process on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with $X = (X^1, \ldots, X^d)$. For every $i, j = 1, \ldots, d$ and any $\varepsilon > 0$, we denote

$$[X^i, X^j]_t^\varepsilon \;:=\; \int_0^t \frac{\big(X^i_{(s+\varepsilon)\wedge T} - X^i_s\big)\big(X^j_{(s+\varepsilon)\wedge T} - X^j_s\big)}{\varepsilon}\, ds, \qquad \forall\, t\in[0,T]. \tag{A.2}$$

We say that X has **all its mutual brackets** if, for every $i, j = 1, \ldots, d$, $[X^i, X^j]^\varepsilon$ converges in the ucp sense as $\varepsilon \to 0^+$. When $i = j$ we denote $[X^i, X^i]^\varepsilon$ simply by $[X^i]^\varepsilon$.
We denote its limit by $[X^i, X^j]$. When $i = j$ we denote $[X^i, X^i]$ simply by $[X^i]$.

The following technical result concerning the stability of Itô integral will be useful.

Lemma A.3. *Let $X = (X_t)_{t\in[0,T]}$ be a real continuous semimartingale on some filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$. Moreover, for every $\varepsilon > 0$, let $Y^\varepsilon = (Y_t^\varepsilon)_{t\in[0,T]}$ be a family of real progressively measurable processes such that*

$$Y^\varepsilon \;\underset{\varepsilon\to 0^+}{\overset{\text{ucp}}{\longrightarrow}}\; Y,$$

where $Y = (Y_t)_{t\in[0,T]}$ is a real progressively measurable process satisfying $\sup_t |Y_t| < \infty$, $\mathbb{P}$-a.s.. Then

$$\frac{1}{\varepsilon}\int_0^t Y_s^\varepsilon \big(X_{(s+\varepsilon)\wedge T} - X_s\big)\, ds \;\underset{\varepsilon\to 0^+}{\overset{\text{ucp}}{\longrightarrow}}\; \int_0^t Y_s\, dX_s, \tag{A.3}$$

where the latter is Itô integral.

Proof. We begin noting that (A.3) holds if and only if

$$\frac{1}{\varepsilon}\int_0^t Y_s^\varepsilon \big(X_{(s+\varepsilon)\wedge t} - X_s\big)\, ds \;\underset{\varepsilon\to 0^+}{\overset{\text{ucp}}{\longrightarrow}}\; \int_0^t Y_s\, dX_s. \tag{A.4}$$

As a matter of fact, it holds that

$$\begin{aligned}
&\frac{1}{\varepsilon}\int_0^t Y_s^\varepsilon \big(X_{(s+\varepsilon)\wedge T} - X_s\big)\, ds - \frac{1}{\varepsilon}\int_0^t Y_s^\varepsilon \big(X_{(s+\varepsilon)\wedge t} - X_s\big)\, ds \\
&= \frac{1}{\varepsilon}\int_{t-\varepsilon}^t Y_s^\varepsilon \big(X_{(s+\varepsilon)\wedge T} - X_{(s+\varepsilon)\wedge t}\big)\, ds \;\underset{\varepsilon\to 0^+}{\overset{\text{ucp}}{\longrightarrow}}\; 0.
\end{aligned} \tag{A.5}$$

The validity of (A.5) follows easily reasoning on subsequences in terms of $\mathbb{P}$-a.s. pointwise convergence and using the fact that $\sup_t |Y_t| < \infty$, $\mathbb{P}$-a.s..

It remains to prove (A.4). To this effect, we first remark that, for every $t \in [0,T]$,

$$\frac{1}{\varepsilon}\int_0^t Y_s^\varepsilon \big(X_{(s+\varepsilon)\wedge t} - X_s\big)\, ds \;=\; \frac{1}{\varepsilon}\int_0^t \bigg(\int_{(r-\varepsilon)^+}^r Y_s^\varepsilon\, ds\bigg)\, dX_r, \tag{A.6}$$

where $(r-\varepsilon)^+ = (r-\varepsilon)\vee 0$ denotes the positive part of $(r-\varepsilon)$. The validity of (A.6) follows from the fact that, by stopping techniques, we can reduce to the case when the bounded variation component of X, denoted A^X, is bounded, the local martingale component, denoted M^X, is bounded and also its bracket $[M^X]_T$ is bounded. In addition, we can suppose Y^ε to be bounded.

Then, (A.6) follows by the stochastic Fubini theorem, see for instance Theorem 64, Chapter IV, of [77]. It remains to show that

$$\int_0^t \left(\frac{1}{\varepsilon}\int_{(r-\varepsilon)^+}^r Y_s^\varepsilon\,ds\right) dX_r \ \xrightarrow[\varepsilon\to0^+]{\text{ucp}}\ \int_0^t Y_r\,dX_r. \tag{A.7}$$

In order to prove (A.7), recall that $X = A^X + M^X$, where A^X is the bounded variation component and M^X is the local martingale component. We prove the validity of (A.7) separately for A^X and M^X. Firstly, the ucp convergence (A.7) with A^X in place of X is a direct consequence of the Lebesgue dominated convergence theorem. Finally, (A.7) with M^X in place of X follows by Proposition 2.26, Chapter 3, of [53] and the fact that

$$\int_0^t \left(\frac{1}{\varepsilon}\int_{(r-\varepsilon)^+}^r \big|Y_s^\varepsilon - Y_s\big|\,ds\right)^2 dM_r^X \ \xrightarrow[\varepsilon\to0^+]{\mathbb{P}}\ 0.$$

□

We can now prove the functional Itô formula for strictly non-anticipative maps.

Proposition A.1. *Let $\hat u \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$. Suppose also that $\hat u$ is a strictly non-anticipative map. Then, for every d-dimensional continuous semimartingale $X = (X_t)_{t\in[0,T]}$ on some filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,T]},\mathbb{P})$, with $X = (X^1,\ldots,X^d)$, the following* ***functional Itô formula*** *holds:*

$$\begin{aligned}\hat u(t,X,X_t) \ &= \ \hat u(0,X,X_0) + \int_0^t \partial_t^H \hat u(s,X)\,ds + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_{x_ix_j}^V \hat u(s,X,X_s)\,d[X^i,X^j]_s \\ &\quad + \sum_{i=1}^d \int_0^t \partial_{x_i}^V \hat u(s,X,X_s)\,dX_s^i, \qquad \text{for all } 0\le t\le T,\ \mathbb{P}\text{-a.s.}\end{aligned} \tag{A.8}$$

where the terms in the last summation are Itô integrals.

Proof. Let $\delta_0 > 0$ be the constant appearing in Definition A.1 of $\hat u$. Given $s\in[0,T]$, for any $\varepsilon\in[0,\delta_0]$ we have

$$\hat u((s+\varepsilon)\wedge T, X, X_{(s+\varepsilon)\wedge T}) - \hat u(s,X,X_s) \ = \ I_1(s,\varepsilon) + I_2(s,\varepsilon),$$

where

$$\begin{aligned}I_1(s,\varepsilon) \ &= \ \hat u((s+\varepsilon)\wedge T, X, X_{(s+\varepsilon)\wedge T}) - \hat u((s+\varepsilon)\wedge T, X, X_s), \\ I_2(s,\varepsilon) \ &= \ \hat u((s+\varepsilon)\wedge T, X, X_s) - \hat u(s,X,X_s).\end{aligned}$$

Notice that, by telescoping, we have

$$\int_0^\cdot \frac{\hat u((s+\varepsilon)\wedge T, X, X_{(s+\varepsilon)\wedge T}) - \hat u(s,X,X_s)}{\varepsilon}\,ds \ \xrightarrow[\varepsilon\to0^+]{\text{ucp}}\ \hat u(\cdot,X,X_\cdot) - \hat u(0,X,X_0).$$

It remains to investigate the convergence of $\frac{1}{\varepsilon}\int_0^\cdot I_1(s,\varepsilon)\,ds$ and $\frac{1}{\varepsilon}\int_0^\cdot I_2(s,\varepsilon)\,ds$ as $\varepsilon\to0^+$.

Convergence of $\frac{1}{\varepsilon}\int_0^\cdot I_1(s,\varepsilon)\,ds$ as $\varepsilon\to0^+$. We have

$$I_1(s,\varepsilon) \ = \ I_{11}(s,\varepsilon) + I_{12}(s,\varepsilon) + I_{13}(s,\varepsilon) + I_{14}(s,\varepsilon),$$

where

$$I_{11}(s,\varepsilon) \ = \ \sum_{i=1}^d \partial_{x_i}^V \hat u((s+\varepsilon)\wedge T, X, X_s)(X^i_{(s+\varepsilon)\wedge T} - X^i_s),$$

$$
\begin{aligned}
I_{12}(s,\varepsilon) &= \frac{1}{2}\sum_{i,j=1}^d \partial^V_{x_ix_j}\hat u(s,X,X_s)(X^i_{(s+\varepsilon)\wedge T}-X^i_s)(X^j_{(s+\varepsilon)\wedge T}-X^j_s),\\
I_{13}(s,\varepsilon) &= \frac{1}{2}\sum_{i,j=1}^d \big(\partial^V_{x_ix_j}\hat u((s+\varepsilon)\wedge T,X,X_s)\\
&\quad -\partial^V_{x_ix_j}\hat u(s,X,X_s)\big)(X^i_{(s+\varepsilon)\wedge T}-X^i_s)(X^j_{(s+\varepsilon)\wedge T}-X^j_s),\\
I_{14}(s,\varepsilon) &= \frac{1}{2}\sum_{i,j=1}^d \int_0^1 \big(\partial^V_{x_ix_j}\hat u((s+\varepsilon)\wedge T,X,X_s+a(X_{(s+\varepsilon)\wedge T}-X_s))\\
&\quad -\partial^V_{x_ix_j}\hat u((s+\varepsilon)\wedge T,X,X_s)\big)(X^i_{(s+\varepsilon)\wedge T}-X^i_s)(X^j_{(s+\varepsilon)\wedge T}-X^j_s)\,da.
\end{aligned}
$$

We begin noting that, by usual arguments (see for instance Proposition 1.2 of [82]), it holds that

$$
\frac{1}{\varepsilon}\int_0^\cdot I_{12}(s,\varepsilon)\,ds \ \underset{\varepsilon\to0^+}{\overset{\text{ucp}}{\longrightarrow}}\ \frac{1}{2}\sum_{i,j=1}^d\int_0^t \partial^V_{x_ix_j}\hat u(s,X,X_s)\,d[X^i,X^j]_s.
$$

On the other hand, we have

$$
\bigg|\int_0^t I_{13}(s,\varepsilon)\,ds\bigg| \le \sum_{i,j=1}^d \sup_{t\in[0,T]}\big|\partial^V_{x_ix_j}\hat u((t+\varepsilon)\wedge T,X,X_t)-\partial^V_{x_ix_j}\hat u(t,X,X_t)\big|\cdot\frac{1}{2}\Big([X^i]^\varepsilon_T+[X^j]^\varepsilon_T\Big),
$$

with $[X^i]^\varepsilon_T$ and $[X^j]^\varepsilon_T$ given by (A.2). Since X^i is a semimartingale, then $[X^i]^\varepsilon_T$ converges in probability to $[X^i]_T$ as $\varepsilon\to0^+$, see for instance Section 3.3 in [83]. In particular $[X]$ exists and it is the usual bracket. Now, for every fixed ω, since $X(\omega)$ is continuous and the map $(t,\varepsilon)\mapsto\partial^V_{x_ix_j}\hat u((t+\varepsilon)\wedge T,X(\omega),X_t(\omega))$ on $[0,T]\times[0,\delta_0]$ is continuous and therefore uniformly continuous, there exists a modulus of continuity $\rho_{13}\colon[0,+\infty)\to[0,+\infty)$ (not depending on t,ε, possibly depending on ω), which can be taken to be non-decreasing and independent of i,j, such that

$$
\bigg|\int_0^t I_{13}(s,\varepsilon,\omega)\,ds\bigg| \ \le\ \frac{1}{2}\rho_{13}(\varepsilon)\sum_{i,j=1}^d\Big([X^i,X^i]^\varepsilon_T(\omega)+[X^j,X^j]^\varepsilon_T(\omega)\Big).
$$

Since X has all its mutual brackets, it follows that

$$
\frac{1}{\varepsilon}\int_0^\cdot I_{13}(s,\varepsilon)\,ds \ \underset{\varepsilon\to0^+}{\overset{\text{ucp}}{\longrightarrow}}\ 0.
$$

Similarly, we have

$$
\begin{aligned}
\bigg|\int_0^t I_{14}(s,\varepsilon)\,ds\bigg| \ &\le\ \frac{1}{2}\sum_{i,j=1}^d\int_0^1\sup_{t\in[0,T]}\big|\partial^V_{x_ix_j}\hat u((t+\varepsilon)\wedge T,X,X_t+a(X_{(t+\varepsilon)\wedge T}-X_t))\\
&\quad -\partial^V_{x_ix_j}\hat u((t+\varepsilon)\wedge T,X,X_t)\big|\,da\cdot\frac{1}{2}\Big([X^i,X^i]^\varepsilon_T+[X^j,X^j]^\varepsilon_T\Big).
\end{aligned}
$$

For every fixed ω, since $X(\omega)$ is continuous and the map $(t,a,\varepsilon)\mapsto\partial^V_{x_ix_j}\hat u((t+\varepsilon)\wedge T,X(\omega),X_t(\omega)+a(X_{(t+\varepsilon)\wedge T}(\omega)-X_t(\omega)))$ on $[0,T]\times[0,1]\times[0,\delta_0]$ is continuous and therefore uniformly continuous, there exists a modulus of continuity $\rho_{14}\colon[0,+\infty)\to[0,+\infty)$ (not depending on t,a,ε, possibly depending on ω), which can be taken to be non-decreasing and independent of i,j, such that

$$
\bigg|\int_0^t I_{14}(s,\varepsilon,\omega)\,ds\bigg| \ \le\ \frac{1}{4}\sum_{i,j=1}^d\int_0^1\sup_{t\in[0,T]}\rho_{14}\big(a|X_{(t+\varepsilon)\wedge T}(\omega)-X_t(\omega)|\big)\,da\cdot\Big([X^i]^\varepsilon_T(\omega)+[X^j]^\varepsilon_T(\omega)\Big).
$$

Recalling that ρ_{14} is non-decreasing, we find

$$\bigg|\int_0^t I_{14}(s,\varepsilon,\omega)\,ds\bigg| \ \le\ \frac{1}{4}\sum_{i,j=1}^d \rho_{14}\Big(\sup_{t\in[0,T]}|X_{(t+\varepsilon)\wedge T}(\omega)-X_t(\omega)|\Big)\Big([X^i]_T^\varepsilon(\omega)+[X^j]_T^\varepsilon(\omega)\Big).$$

Since X has all its mutual brackets, it follows that

$$\frac{1}{\varepsilon}\int_0^\cdot I_{14}(s,\varepsilon)\,ds \ \underset{\varepsilon\to0^+}{\overset{\text{ucp}}{\longrightarrow}}\ 0.$$

Finally, it remains to investigate the ucp convergence of the term $\frac{1}{\varepsilon}\int_0^\cdot I_{11}(s,\varepsilon)\,ds$.

Convergence of $\frac{1}{\varepsilon}\int_0^\cdot I_2(s,\varepsilon)\,ds$ *as* $\varepsilon\to0^+$. Since $\varepsilon\le\delta_0$ and $\hat u$ is a delayed map with delay δ_0, we obtain

$$I_2(s,\varepsilon) \ = \ \hat u((s+\varepsilon)\wedge T, X(\cdot\wedge s), X_s) - \hat u(s,X,X_s).$$

For every fixed ω, let $\phi\colon[0,1]\to\mathbb{R}$ be given by

$$\phi(a) \ = \ \hat u((s+a\varepsilon)\wedge T, X(\cdot\wedge s)(\omega), X_s(\omega)).$$

By Lemma A.1 we deduce that $\phi\in C^1([0,1])$. Then

$$\phi(1) \ = \ \phi(0)+\int_0^1\phi'(a)\,da.$$

So, in particular,

$$I_1(s,\varepsilon) \ = \ \varepsilon\int_0^1 \partial_t^H\hat u((s+a\varepsilon)\wedge T, X(\cdot\wedge s))\,da.$$

Hence

$$\int_0^t I_1(s,\varepsilon)\,ds \ = \ \varepsilon\int_0^t\bigg(\int_0^1 \partial_t^H\hat u((s+a\varepsilon)\wedge T, X(\cdot\wedge s))\,da\bigg)ds.$$

For every fixed ω, since $X(\omega)$ is continuous and the map $(a,\varepsilon)\mapsto\partial_t^H\hat u((s+a\varepsilon)\wedge T, X(\cdot\wedge s)(\omega))$ on $[0,1]\times[0,\delta_0]$ is continuous and therefore locally bounded, we conclude that

$$\frac{1}{\varepsilon}\int_0^\cdot I_1(s,\varepsilon)\,ds \ \underset{\varepsilon\to0^+}{\overset{\text{ucp}}{\longrightarrow}}\ \int_0^\cdot \partial_t^H\hat u(s,X)\,ds.$$

□

We can finally prove the functional Itô formula, namely Theorem 2.1.

Proof (of Theorem 2.1). Let $\delta>0$ and define $\hat u_\delta\colon[0,T]\times C([0,T];\mathbb{R}^d)\times\mathbb{R}^d\to\mathbb{R}$ as

$$\hat u_\delta(t,\boldsymbol{x},\mathbf{y}) \ := \ \hat u\big(t,\boldsymbol{x}(\cdot\wedge(t-\delta)^+),\mathbf{y}\big), \qquad \forall\,(t,\boldsymbol{x},\mathbf{y})\in\boldsymbol{\Lambda}\times\mathbb{R}^d.$$

By Remark A.1 we know that $\hat u_\delta$ is a delayed map with delay δ, in the sense of Definition A.1. Moreover, by Lemma A.2 we have that $\hat u_\delta\in\boldsymbol{C}^{1,2}(\boldsymbol{\Lambda}\times\mathbb{R}^d)$ and the following equalities hold:

$$\begin{aligned}
\partial_t^H\hat u_\delta(t,\boldsymbol{x}) \ &= \ \partial_t^H\hat u\big(t,\boldsymbol{x}(\cdot\wedge(t-\delta)^+)\big), && \forall\,(t,\boldsymbol{x})\in[0,T)\times C([0,T];\mathbb{R}^d),\\
\partial_{\boldsymbol{x}}^V\hat u_\delta(t,\boldsymbol{x},\mathbf{y}) \ &= \ \partial_{\boldsymbol{x}}^V\hat u\big(t,\boldsymbol{x}(\cdot\wedge(t-\delta)^+),\mathbf{y}\big), && \forall\,(t,\boldsymbol{x},\mathbf{y})\in\boldsymbol{\Lambda}\times\mathbb{R}^d,\\
\partial_{\boldsymbol{x}\boldsymbol{x}}^V\hat u_\delta(t,\boldsymbol{x},\mathbf{y}) \ &= \ \partial_{\boldsymbol{x}\boldsymbol{x}}^V\hat u\big(t,\boldsymbol{x}(\cdot\wedge(t-\delta)^+),\mathbf{y}\big), && \forall\,(t,\boldsymbol{x},\mathbf{y})\in\boldsymbol{\Lambda}\times\mathbb{R}^d.
\end{aligned}$$

By the above equalities and Proposition A.1, we obtain

$$\begin{aligned}
\hat u_\delta(t,X,X_t) \ &= \ \hat u_\delta(0,X,X_0)+\int_0^t\partial_t^H\hat u(s,X(\cdot\wedge(t-\delta)^+))\,ds\\
&\quad+\frac{1}{2}\sum_{i,j=1}^d\int_0^t\partial_{x_ix_j}^V\hat u(s,X(\cdot\wedge(t-\delta)^+)),X_s)\,d[X^i,X^j]_s
\end{aligned}$$

$$+\sum_{i=1}^d \int_0^t \partial_{x_i}^V \hat{u}(s, X(\cdot \wedge (t-\delta)^+)), X_s)\, dX_s^i, \qquad \text{for all } 0 \le t \le T,\ \mathbb{P}\text{-a.s.}$$

Our aim is to obtain the functional Itô formula (A.8) letting $\delta \to 0^+$ in the above equality, for every fixed $t \in [0,T]$, and investigating the convergence of each term. We begin noting that $\hat{u}_\delta(0,X,0) = \hat{u}(0,X,0)$. Moreover, for any fixed ω and $t \in [0,T]$,

$$\hat{u}_\delta(t, X(\omega), X_t(\omega)) \underset{\delta\to 0^+}{\longrightarrow} \hat{u}(t, X(\omega), X_t(\omega)).$$

Now, for any fixed ω and s, we have

$$\partial_t^H \hat{u}(s, X(\cdot \wedge (s-\delta)^+)(\omega)) \underset{\delta\to 0^+}{\longrightarrow} \partial_t^H \hat{u}(s, X(\omega)), \tag{A.9}$$

$$\partial_{\boldsymbol{x}}^V \hat{u}(s, X(\cdot \wedge (s-\delta)^+)(\omega), X_s(\omega)) \underset{\delta\to 0^+}{\longrightarrow} \partial_{\boldsymbol{x}}^V \hat{u}(s, X(\omega), X_s(\omega)), \tag{A.10}$$

$$\partial_{\boldsymbol{xx}}^V \hat{u}(s, X(\cdot \wedge (s-\delta)^+)(\omega), X_s(\omega)) \underset{\delta\to 0^+}{\longrightarrow} \partial_{\boldsymbol{xx}}^V \hat{u}(s, X(\omega), X_s(\omega)). \tag{A.11}$$

Notice that, for any fixed ω, the families (parametrized by δ) of continuous maps on $[0,T]$ given by $s \mapsto \partial_t^H \hat{u}(s, X(\cdot\wedge(s-\delta)^+)(\omega))$, $s \mapsto \partial_{\boldsymbol{x}}^V \hat{u}(s, X(\cdot\wedge(s-\delta)^+)(\omega), X_s(\omega))$, $s \mapsto \partial_{\boldsymbol{xx}}^V \hat{u}(s, X(\cdot\wedge(s-\delta)^+)(\omega), X_s(\omega))$ are uniformly bounded and equicontinuous. Then, by the Arzelà-Ascoli theorem we deduce that convergences (A.9)-(A.10)-(A.11) hold uniformly with respect to $s \in [0,T]$. This implies that, for any fixed ω and $t \in [0,T]$,

$$\int_0^t \partial_t^H \hat{u}(s, X(\cdot \wedge (s-\delta)^+)(\omega))\, ds \underset{\delta\to 0^+}{\longrightarrow} \int_0^t \partial_t^H \hat{u}(s, X(\omega))\, ds,$$

$$\frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_{x_i x_j}^V \hat{u}(s, X(\cdot \wedge (s-\delta)^+)(\omega), X_s(\omega))\, d[X^i, X^j]_s(\omega)$$

$$\underset{\delta\to 0^+}{\longrightarrow} \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_{x_i x_j}^V \hat{u}(s, X(\omega), X_s(\omega))\, d[X^i, X^j]_s(\omega).$$

It remains to study the convergence of the stochastic integral. Recall that $X = M + V$, where M is a local martingale and V is bounded variation process, so that the stochastic integral can be written as the sum of two integrals with respect to M and V, respectively. The integral with respect to V can be treated along the same lines as the integral with respect to $[X^i, X^j]$. It remains to deal with the stochastic integral with respect to M. By Proposition 2.26, Chapter 3, in [53] we know that the claim follows if, for every $i = 1, \ldots, d$,

$$\int_0^T \big|\partial_{x_i}^V \hat{u}(s, X(\cdot \wedge (s-\delta)^+), X_s) - \partial_{x_i}^V \hat{u}(s, X, X_s)\big|^2\, d[X^i]_s \ \underset{\delta\to 0^+}{\overset{\mathbb{P}}{\longrightarrow}}\ 0, \tag{A.12}$$

where the convergence is understood in the probability sense. Using again convergence (A.10), uniform with respect to $s \in [0,T]$, we deduce that (A.12) holds in the $\mathbb{P}$-a.s. sense. This concludes the proof. □

Appendix B.
Consistency

Proof (of Lemma 2.1). The claim concerning the horizontal derivatives follows directly from their definition (Definition 2.5-(i)).

It remains to prove the claim concerning the vertical derivatives. To this end, let $X = (X_t)_{t\in[0,T]}$ be a d-dimensional continuous semimartingale on some filtered probability space

$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, with $X = (X^1, \ldots, X^d)$. Since $\hat{u}_1, \hat{u}_2 \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$, by Theorem 2.1 the following functional Itô formulae hold:

$$\begin{aligned}\hat{u}_1(t, X, X_t) \ &= \ \hat{u}_1(0, X, X_0) + \int_0^t \partial_t^H \hat{u}_1(s, X)\, ds + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_{x_i x_j}^V \hat{u}_1(s, X, X_s)\, d[X^i, X^j]_s \\ &\quad + \sum_{i=1}^d \int_0^t \partial_{x_i}^V \hat{u}_1(s, X, X_s)\, dX_s^i, \qquad \text{for all } 0 \le t \le T, \ \mathbb{P}\text{-a.s.}\end{aligned}$$

and

$$\begin{aligned}\hat{u}_2(t, X, X_t) \ &= \ \hat{u}_2(0, X, X_0) + \int_0^t \partial_t^H \hat{u}_2(s, X)\, ds + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \partial_{x_i x_j}^V \hat{u}_2(s, X, X_s)\, d[X^i, X^j]_s \\ &\quad + \sum_{i=1}^d \int_0^t \partial_{x_i}^V \hat{u}_2(s, X, X_s)\, dX_s^i, \qquad \text{for all } 0 \le t \le T, \ \mathbb{P}\text{-a.s.}\end{aligned}$$

Recalling that $\hat{u}_1$ and $\hat{u}_2$, together with their horizontal derivatives, coincide on continuous paths, identifying bounded variation and local martingale parts in the above formulae, we obtain that the following equalities hold (up to a $\mathbb{P}$-null set), for every $i, j = 1, \ldots, d$ and any $t \in [0, T]$,

$$\partial_{x_i}^V \hat{u}_1(t, X, X_t) \ = \ \partial_{x_i}^V \hat{u}_2(t, X, X_t), \qquad \partial_{x_i x_j}^V \hat{u}_1(t, X, X_t) \ = \ \partial_{x_i x_j}^V \hat{u}_2(t, X, X_t). \tag{B.1}$$

Now, fix $t \in [0, T]$ and consider a semimartingale X whose law has full support on the set of trajectories stopped at time t, namely $\{\boldsymbol{x} \in C([0, T]; \mathbb{R}^d) \colon \boldsymbol{x}(s) = \boldsymbol{x}(t), \text{ for every } s \in [t, T]\}$. An example of such an X is given by $X_s := \eta + \boldsymbol{W}_{s\wedge t}$, $s \in [t, T]$, where $\boldsymbol{W} = (\boldsymbol{W}_s)_{s\in[0,T]}$ is a d-dimensional Brownian motion, while $\eta \colon \Omega \to \mathbb{R}^d$ is a d-dimensional random variable with full support, independent of $\boldsymbol{W}$. Then, for such a semimartingale X, we deduce from (B.1) the following equalities at time t:

$$\partial_{\boldsymbol{x}}^V \hat{u}_1(t, \boldsymbol{x}, \boldsymbol{x}(t)) \ = \ \partial_{\boldsymbol{x}}^V \hat{u}_2(t, \boldsymbol{x}, \boldsymbol{x}(t)), \qquad \partial_{\boldsymbol{x}\boldsymbol{x}}^V \hat{u}_1(t, \boldsymbol{x}, \boldsymbol{x}(t)) \ = \ \partial_{\boldsymbol{x}\boldsymbol{x}}^V \hat{u}_2(t, \boldsymbol{x}, \boldsymbol{x}(t)),$$

for every $\boldsymbol{x} \in C([0, T]; \mathbb{R}^d)$. Since the above equalities hold for every $t \in [0, T]$, the claim follows. □

APPENDIX C.
Cylindrical approximation

In the present Appendix we state two results already proved in [17, 18], namely Theorem 3.5 in [17] and Theorem 3.12 in [18], which correspond respectively to Lemma C.1 and Lemma C.2 below. Notice however that in [17, 18] the pathwise derivatives are defined in an alternative manner. For this reason, in order to help the reader, we prefer to report the proof of these two results in the present setting.

C.1. **The deterministic calculus via regularization.** We begin recalling some results from the deterministic calculus of regularization, as developed in Section 3.2 of [24] and Section 2.2 of [17], for which we refer for all the details. The only difference with respect to [24] and [17] being that here we consider $\mathbb{R}^d$-valued paths (with d not necessarily equal to 1), even if, as usual, so that we rely on the one-dimensional theory, as we work component by component.

Firstly, for every $t \geq 0$ and any function $\boldsymbol{f}\colon [0,t] \to \mathbb{R}^d$ we define the following extensions to the entire real line:

$$\boldsymbol{f}_{(0,t]}(s) \ := \ \begin{cases} \mathbf{0}, & s > t, \\ \boldsymbol{f}(s), & s \in [0,t], \\ \boldsymbol{f}(0), & s < 0, \end{cases} \qquad\qquad \boldsymbol{f}_{[0,t]}(s) \ := \ \begin{cases} \boldsymbol{f}(t), & s > t, \\ \boldsymbol{f}(s), & s \in [0,t], \\ \mathbf{0}, & s < 0. \end{cases}$$

Definition C.1. Let $\boldsymbol{f}\colon [0,t] \to \mathbb{R}^d$ and $g\colon [0,t] \to \mathbb{R}$ be càdlàg functions. When the limit

$$\int_{[0,t]} g(s)\, d^- \boldsymbol{f}(s) \ := \ \lim_{\varepsilon \to 0^+} \int_{-\infty}^{+\infty} g_{(0,t]}(s)\, \frac{\boldsymbol{f}_{[0,t]}(s+\varepsilon) - \boldsymbol{f}_{[0,t]}(s)}{\varepsilon}\, ds$$

exists and it is finite, we denote it by $\int_{[0,t]} g\, d^- \boldsymbol{f}$ and call it **forward integral of g with respect to $\boldsymbol{f}$**.

We recall from [17], Proposition 2.11, the following integration by parts formula, which will be used several times in this Appendix.

Proposition C.1. *Let $\boldsymbol{f}\colon [0,t] \to \mathbb{R}^d$ and $g\colon [0,t] \to \mathbb{R}$ be càdlàg functions, with g being of bounded variation. The following* ***integration by parts formula*** *holds:*

$$\int_{[0,t]} g(s)\, d^- \boldsymbol{f}(s) \ = \ g(t)\, \boldsymbol{f}(t) - \int_{(0,t]} \boldsymbol{f}(s)\, dg(s),$$

where $\int_{(0,t]} \boldsymbol{f}(s)\, dg(s)$ is a Lebesgue-Stieltjes integral on $(0,t]$.

C.2. **Cylinder terminal condition ξ.**

Lemma C.1. *Suppose that ξ is a cylinder (or tame) function, in the sense that it admits the representation*

$$\xi(\boldsymbol{x}) \ = \ g\bigg(\int_{[0,T]} \psi_0(t)\, d^- \boldsymbol{x}(t), \ldots, \int_{[0,T]} \psi_n(t)\, d^- \boldsymbol{x}(t) \bigg), \qquad \boldsymbol{x} \in C([0,T];\mathbb{R}^d),$$

for some non-negative integer n, where we have the following.

- *$g\colon \mathbb{R}^{d(n+1)} \to \mathbb{R}$ is of class $C^2(\mathbb{R}^{d(n+1)})$ and, together with its first and second-order partial derivatives, satisfies a polynomial growth condition;*
- *$\psi_0, \ldots, \psi_n\colon [0,T] \to \mathbb{R}$ are continuous.*

Then, the function v defined by (4.4) is in $\boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$ and is a classical (smooth) solution of equation (4.2).

Proof. Let $(t,\boldsymbol{x}) \in \boldsymbol{\Lambda}$ and consider $\boldsymbol{W}^{t,\boldsymbol{x}} = (\boldsymbol{W}_s^{t,\boldsymbol{x}})_{s\in[0,T]}$ given by (4.5). From the definition of v in (4.4), we have

$$\begin{aligned} v(t,\boldsymbol{x}) \ &= \ \mathbb{E}\bigg[g\bigg(\int_{[0,T]} \psi_0(s)\, d^- \boldsymbol{W}_s^{t,\boldsymbol{x}}, \ldots, \int_{[0,T]} \psi_n(s)\, d^- \boldsymbol{W}_s^{t,\boldsymbol{x}} \bigg) \bigg] \\ &= \ \mathbb{E}\bigg[g\bigg(\int_{[0,t]} \psi_0(s)\, d^- \boldsymbol{x}(s) + \int_t^T \psi_0(s)\, d\boldsymbol{W}_s, \ldots \bigg) \bigg], \end{aligned}$$

where the second equality follows from the fact that the forward integral coincides with the Itô integral when the integrator is the Brownian motion (or, more generally, a continuous semimartingale), see for instance Proposition 6, Section 3.3, in [83].

In order to prove that $v \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$, we consider the following lifting $\hat{v}\colon \boldsymbol{\Lambda} \times \mathbb{R}^d \to \mathbb{R}$ of v:

$$\hat{v}(t,\boldsymbol{x},\mathbf{y}) \ = \ \mathbb{E}\bigg[g\bigg(\psi_0(t)(\mathbf{y} - \boldsymbol{x}(t)) + \int_{[0,t]} \psi_0(s)\, d^- \boldsymbol{x}(s) + \int_t^T \psi_0(s)\, d\boldsymbol{W}_s, \ldots \bigg) \bigg],$$

for all $(t,\boldsymbol{x},\mathbf{y}) \in \boldsymbol{\Lambda}\times\mathbb{R}^d$. Notice that

$$\hat v(t,\boldsymbol{x},\mathbf{y}) \ = \ \hat V\bigg(t\,,\,\psi_0(t)(\mathbf{y}-\boldsymbol{x}(t))+\int_{[0,t]}\psi_0(s)\,d^-\boldsymbol{x}(s)\,,\,\ldots\,,\,\psi_n(t)(\mathbf{y}-\boldsymbol{x}(t))+\int_{[0,t]}\psi_n(s)\,d^-\boldsymbol{x}(s)\bigg),$$

where $\hat V\colon[0,T]\times\mathbb{R}^{d(n+1)}\to\mathbb{R}$ is given by

$$\hat V(t,\mathbf{z}) \ = \ \mathbb{E}\bigg[g\bigg(\mathbf{z}_0+\int_t^T\psi_0(s)\,d\boldsymbol{W}_s\,,\,\ldots\,,\,\mathbf{z}_n+\int_t^T\psi_n(s)\,d\boldsymbol{W}_s\bigg)\bigg],$$

for all $t\in[0,T]$ and $\mathbf{z}\in\mathbb{R}^{d(n+1)}$, with $\mathbf{z}=(\mathbf{z}_0,\ldots,\mathbf{z}_n)$ and $\mathbf{z}_0,\ldots,\mathbf{z}_n\in\mathbb{R}^d$. Let $\sigma\colon[0,T]\to\mathbb{R}^{d(n+1)\times d}$ be given by

$$\sigma(t) \ = \ \begin{bmatrix}\psi_0(t)\,I\\ \psi_1(t)\,I\\ \vdots\\ \psi_n(t)\,I\end{bmatrix} \tag{C.1}$$

where I denotes the $d\times d$ identity matrix. It is well-known (see, for instance, Theorem 5.6.1 in [43]) that $\hat V\in C^{1,2}([0,T]\times\mathbb{R}^{d(n+1)})$ and satisfies (here $\partial_t\hat V(t,\mathbf{z})$ denotes the standard time derivative, while $\partial_{\mathbf{z}\mathbf{z}}\hat V(t,\mathbf{z})$ is the standard Hessian matrix of spatial derivatives)

$$\begin{cases}\partial_t\hat V(t,\mathbf{z})+\dfrac{1}{2}\mathrm{tr}\big[\sigma(t)\sigma^{\intercal}(t)\partial_{\mathbf{z}\mathbf{z}}\hat V(t,\mathbf{z})\big] \ = \ 0, & (t,\mathbf{z})\in[0,T)\times\mathbb{R}^{d(n+1)},\\ \hat V(T,\mathbf{z}) \ = \ g(\mathbf{z}), & \mathbf{z}\in\mathbb{R}^{d(n+1)}.\end{cases} \tag{C.2}$$

Let us find the expression of the pathwise derivatives of $\hat v$ in terms of $\hat V$. Concerning the horizontal derivative at $(t,\boldsymbol{x})\in\boldsymbol{\Lambda}$, $t<T$, we have

$$\begin{aligned}\partial_t^H\hat v(t,\boldsymbol{x}) \ &= \ \lim_{\delta\to0^+}\frac{\hat v(t+\delta,\boldsymbol{x}(\cdot\wedge t),\boldsymbol{x}(t))-\hat v(t,\boldsymbol{x},\boldsymbol{x}(t))}{\delta}\\ &= \ \lim_{\delta\to0^+}\frac{1}{\delta}\bigg\{\hat V\bigg(t+\delta,\psi_0(t+\delta)(\boldsymbol{x}(t)-\boldsymbol{x}((t+\delta)\wedge t))+\int_{[0,t+\delta]}\psi_0(s)\,d^-\boldsymbol{x}(s\wedge t),\ldots\bigg)\\ &\qquad -\hat V\bigg(t,\psi_0(t)(\boldsymbol{x}(t)-\boldsymbol{x}(t))+\int_{[0,t]}\psi_0(s)\,d^-\boldsymbol{x}(s),\ldots\bigg)\bigg\}\\ &= \ \lim_{\delta\to0^+}\frac{1}{\delta}\bigg\{\hat V\bigg(t+\delta,\int_{[0,t]}\psi_0(s)\,d^-\boldsymbol{x}(s),\ldots\bigg)-\hat V\bigg(t,\int_{[0,t]}\psi_0(s)\,d^-\boldsymbol{x}(s),\ldots\bigg)\bigg\}\\ &= \ \partial_t\hat V\bigg(t,\int_{[0,t]}\psi_0(s)\,d^-\boldsymbol{x}(s),\ldots\bigg).\end{aligned}$$

Concerning the first-order vertical derivatives at $(t,\boldsymbol{x},\mathbf{y})\in\boldsymbol{\Lambda}\times\mathbb{R}^d$, for every $i=1,\ldots,d$, we have

$$\begin{aligned}\partial_{x_i}^V\hat v(t,\boldsymbol{x},\mathbf{y}) \ &= \ \lim_{h\to0}\frac{\hat v(t,\boldsymbol{x},\mathbf{y}+h\mathbf{e}_i)-\hat v(t,\boldsymbol{x},\mathbf{y})}{h}\\ &= \ \lim_{h\to0}\frac{1}{h}\bigg\{\hat V\bigg(t\,,\,\psi_0(t)(\mathbf{y}+h\mathbf{e}_i-\boldsymbol{x}(t))+\int_{[0,t]}\psi_0(s)\,d^-\boldsymbol{x}(s)\,,\,\ldots\bigg)\\ &\qquad -\hat V\bigg(t\,,\,\psi_0(t)(\mathbf{y}-\boldsymbol{x}(t))+\int_{[0,t]}\psi_0(s)\,d^-\boldsymbol{x}(s)\,,\,\ldots\bigg)\bigg\}\\ &= \ \bigg\langle\sigma_i(t),\partial_{\mathbf{z}}\hat V\bigg(t\,,\,\psi_0(t)(\mathbf{y}-\boldsymbol{x}(t))+\int_{[0,t]}\psi_0(s)\,d^-\boldsymbol{x}(s)\,,\,\ldots\bigg)\bigg\rangle,\end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $\mathbb{R}^{d(n+1)}$, $\sigma_i(t)$ is the i-th column of the matrix $\sigma(t)$ in (C.1), and $\partial_{\mathbf{z}} \hat{V}$ is the standard gradient of spatial derivatives of $\hat{V}$.

Concerning the second-order vertical derivatives, it holds that

$$\partial^V_{\boldsymbol{x}\boldsymbol{x}} \hat{v}(t, \boldsymbol{x}, \mathbf{y}) \;=\; \sigma^{\intercal}(t) \partial_{\mathbf{z}\mathbf{z}} \hat{V}\bigg(t\,,\, \psi_0(t)(\mathbf{y} - \boldsymbol{x}(t)) + \int_{[0,t]} \psi_0(s)\, d^- \boldsymbol{x}(s)\,,\, \ldots\bigg) \sigma(t).$$

Since $\hat{V} \in C^{1,2}([0,T] \times \mathbb{R}^{d(n+1)})$, we deduce that $\hat{v} \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda} \times \mathbb{R}^d)$, so that $v \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$. Finally, since $\hat{V}$ is a classical (smooth) solution of equation (C.2), using the relations between the pathwise derivatives of $\hat{v}$ (and hence of v) and the derivatives of $\hat{V}$, we deduce that v is a classical (smooth) solution of equation (4.2). □

C.3. **Cylindrical approximation.**

Lemma C.2. *Suppose that $\xi \colon C([0,T]; \mathbb{R}^d) \to \mathbb{R}$ is continuous and satisfies the polynomial growth condition*

$$|\xi(\boldsymbol{x})| \;\le\; M(1 + \|\boldsymbol{x}\|_\infty^p), \qquad \textit{for all } \boldsymbol{x} \in C([0,T]; \mathbb{R}^d),$$

for some positive constants M and p. Then, there exists a sequence $\{\xi_n\}_n$, with $\xi_n \colon C([0,T]; \mathbb{R}^d) \to \mathbb{R}$, such that the following holds.

I) *$\{\xi_n\}_n$ converges pointwise to ξ as $n \to +\infty$.*
II) *If ξ is bounded then ξ_n is bounded uniformly with respect to n.*
III) *For every n, ξ_n is given by*

$$\xi_n(\boldsymbol{x}) \;=\; g_n\bigg(\int_{[0,T]} \psi_0(t)\, d^- \boldsymbol{x}(t), \ldots, \int_{[0,T]} \psi_n(t)\, d^- \boldsymbol{x}(t)\bigg),$$

where

i) *for every n, $g_n \colon \mathbb{R}^{d(n+1)} \to \mathbb{R}$ is of class $C^\infty(\mathbb{R}^{d(n+1)})$, with partial derivatives of every order satisfying a polynomial growth condition; moreover, g_n satisfies*

$$|g_n(\mathbf{z})| \;\le\; M'\big(1 + |\mathbf{z}|^{p'}\big), \qquad \textit{for all } \mathbf{z} \in \mathbb{R}^{d(n+1)},$$

for some positive constants M' and p', not depending on n;

ii) *the functions $\psi_\ell \colon [0,T] \to \mathbb{R}$, $\ell \ge 0$, satisfy:*
   a) *ψ_ℓ is of class $C^\infty([0,T])$;*
   b) *ψ_ℓ is uniformly bounded with respect to ℓ;*
   c) *the first derivative of ψ_ℓ is bounded in $L^1([0,T])$, uniformly with respect to ℓ.*

Proof. **Step I.** Let $\{e_\ell\}_{\ell \ge 0}$ be the following orthonormal basis of $L^2([0,T]; \mathbb{R})$:

$$e_0(t) \;=\; \frac{1}{\sqrt{T}}, \qquad e_{2\ell-1}(t) \;=\; \sqrt{\frac{2}{T}} \sin\bigg(2\ell\pi \frac{t}{T}\bigg), \qquad e_{2\ell}(t) \;=\; \sqrt{\frac{2}{T}} \cos\bigg(2\ell\pi \frac{t}{T}\bigg),$$

for all $\ell \ge 1$ and $t \in [0,T]$. Consider the linear operator $\Lambda \colon C([0,T]; \mathbb{R}^d) \to C([0,T]; \mathbb{R}^d)$ given by

$$(\Lambda \boldsymbol{x})(t) \;=\; \boldsymbol{x}(T) \frac{t}{T}, \qquad \text{for all } t \in [0,T],\ \boldsymbol{x} \in C([0,T]; \mathbb{R}^d).$$

Observe that $(\boldsymbol{x} - \Lambda \boldsymbol{x})(0) = (\boldsymbol{x} - \Lambda \boldsymbol{x})(T) = 0$. Now, for every $n \ge 0$, define

$$s_n(\boldsymbol{x} - \Lambda \boldsymbol{x}) \;=\; \sum_{\ell=0}^n (\boldsymbol{x} - \Lambda \boldsymbol{x})_\ell\, e_\ell,$$

$$\sigma_n(\boldsymbol{x} - \Lambda \boldsymbol{x}) \;=\; \frac{s_0(\boldsymbol{x} - \Lambda \boldsymbol{x}) + \cdots + s_n(\boldsymbol{x} - \Lambda \boldsymbol{x})}{n+1} \;=\; \sum_{\ell=0}^n \frac{n+1-\ell}{n+1} (\boldsymbol{x} - \Lambda \boldsymbol{x})_\ell\, e_\ell,$$

where

$$(\boldsymbol{x}-\Lambda\boldsymbol{x})_\ell \;=\; \int_0^T(\boldsymbol{x}(t)-(\Lambda\boldsymbol{x})(t))e_\ell(t)dt \;=\; \int_0^T\boldsymbol{x}(t)e_\ell(t)dt-\boldsymbol{x}(T)\mathcal{E}_\ell(T)+x(T)\frac{1}{T}\int_0^T\mathcal{E}_\ell(t)dt,$$

with $\mathcal{E}_\ell$ being a primitive of e_ℓ. In particular, we take

$$\mathcal{E}_0(t) \;=\; \frac{t}{\sqrt{T}}-\frac{\sqrt{T}}{2}, \qquad \mathcal{E}_{2\ell-1}(t) \;=\; -\sqrt{\frac{T}{2}}\frac{1}{\ell\pi}\cos\Big(2\ell\pi\frac{t}{T}\Big), \qquad \mathcal{E}_{2\ell}(t) \;=\; \sqrt{\frac{T}{2}}\frac{1}{\ell\pi}\sin\Big(2\ell\pi\frac{t}{T}\Big),$$

for all $\ell\geq 1$ and $t\in[0,T]$. Then

$$(\boldsymbol{x}-\Lambda\boldsymbol{x})_\ell \;=\; \int_0^T\boldsymbol{x}(t)e_\ell(t)dt-\boldsymbol{x}(T)\mathcal{E}_\ell(T) \;=\; -\int_{[0,T]}\mathcal{E}_\ell(t)\,d^-\boldsymbol{x}(t),$$

for all $\ell\geq 0$. By Fejér's theorem (see for instance Theorem III.3.4 in [91]), we have

$$\|\sigma_n(\boldsymbol{x}-\Lambda\boldsymbol{x})-(\boldsymbol{x}-\Lambda\boldsymbol{x})\|_\infty \overset{n\to\infty}{\longrightarrow} 0 \qquad\text{and}\qquad \|\sigma_n(\boldsymbol{x}-\Lambda\boldsymbol{x})\|_\infty \;\leq\; \|\boldsymbol{x}-\Lambda\boldsymbol{x}\|_\infty,$$

for all $\boldsymbol{x}\in C([0,T];\mathbb{R}^d)$. Consider the linear operator $T_n\colon C([0,T];\mathbb{R}^d)\to C([0,T];\mathbb{R}^d)$ given by ($e_{-1}(t):=\frac{t}{T}$, for all $t\in[0,T]$)

$$\begin{aligned}T_n\boldsymbol{x} \;=\; \Lambda\boldsymbol{x}+\sigma_n(\boldsymbol{x}-\Lambda\boldsymbol{x})-\big(\sigma_n(\boldsymbol{x}-\Lambda\boldsymbol{x})\big)(0) \;&=\; \boldsymbol{x}(T)e_{-1}+\sum_{\ell=0}^n\frac{n+1-\ell}{n+1}(\boldsymbol{x}-\Lambda\boldsymbol{x})_\ell\,(e_\ell-e_\ell(0))\\ &=\; \boldsymbol{x}(T)e_{-1}+\sum_{\ell=1}^n\frac{n+1-\ell}{n+1}(\boldsymbol{x}-\Lambda\boldsymbol{x})_\ell\,(e_\ell-e_\ell(0)),\end{aligned}$$

for all $n\geq 0$, where the latter equality follows from the fact that e_0 is constant. Then, for any $\boldsymbol{x}\in C([0,T];\mathbb{R}^d)$, $\|T_n\boldsymbol{x}-\boldsymbol{x}\|_\infty\to 0$, as n tends to infinity. Furthermore, there exists a positive constant C, independent of n, such that

$$\|T_n\boldsymbol{x}\|_\infty \;\leq\; C\|\boldsymbol{x}\|_\infty, \qquad \text{for all } \boldsymbol{x}\in C([0,T];\mathbb{R}^d),\ n\geq 0. \tag{C.3}$$

Then, we define

$$\tilde{\xi}_n(\boldsymbol{x}) \;:=\; \xi(T_n\boldsymbol{x}), \qquad \text{for all } \boldsymbol{x}\in C([0,T];\mathbb{R}^d),\ n\geq 0.$$

Notice that

$$\begin{aligned}\tilde{\xi}_n(\boldsymbol{x}) \;&=\; \xi\bigg(\boldsymbol{x}(T)e_{-1}+\sum_{\ell=1}^n\frac{n+1-\ell}{n+1}(\boldsymbol{x}-\Lambda\boldsymbol{x})_\ell\,(e_\ell-e_\ell(0))\bigg)\\ &=\; \xi\bigg(\boldsymbol{x}(T)e_{-1}-\sum_{\ell=1}^n\frac{n+1-\ell}{n+1}\bigg(\int_{[0,T]}\mathcal{E}_\ell(t)\,d^-\boldsymbol{x}(t)\bigg)\,(e_\ell-e_\ell(0))\bigg)\\ &=\; \tilde{g}_n\bigg(\boldsymbol{x}(T)\,,\int_{[0,T]}\mathcal{E}_1(t)\,d^-\boldsymbol{x}(t)\,,\,\ldots\,,\int_{[0,T]}\mathcal{E}_n(t)\,d^-\boldsymbol{x}(t)\bigg)\\ &=\; \tilde{g}_n\bigg(\int_{[0,T]}1\,d^-\boldsymbol{x}(t)\,,\int_{[0,T]}\mathcal{E}_1(t)\,d^-\boldsymbol{x}(t)\,,\,\ldots\,,\int_{[0,T]}\mathcal{E}_n(t)\,d^-\boldsymbol{x}(t)\bigg),\end{aligned}$$

where the last inequality follows from the identity $\boldsymbol{x}(T)=\int_{[0,T]}1\,d^-\boldsymbol{x}(t)$, while $\tilde{g}_n\colon\mathbb{R}^{d(n+1)}\to\mathbb{R}$ is given by

$$\tilde{g}_n(\mathbf{z}) \;:=\; \xi\bigg(\mathbf{z}_0e_{-1}-\sum_{\ell=1}^n\frac{n+1-\ell}{n+1}\,\mathbf{z}_\ell\,(e_\ell-e_\ell(0))\bigg), \tag{C.4}$$

for all $\mathbf{z}\in\mathbb{R}^{d(n+1)}$, with $\mathbf{z}=(\mathbf{z}_0,\ldots,\mathbf{z}_n)$ and $\mathbf{z}_0,\ldots,\mathbf{z}_n\in\mathbb{R}^d$. From now on we denote

$$\psi_0(t) \;=\; 1, \qquad\qquad \psi_\ell(t) \;=\; \mathcal{E}_\ell(t), \qquad \ell\geq 1. \tag{C.5}$$

Step II. We begin introducing the double sequence $\{g_{n,k}\}_{n\geq 0,k\geq 1}$, with $g_{n,k}\colon \mathbb{R}^{d(n+1)}\to\mathbb{R}$ given by

$$g_{n,k}(\mathbf{z}) \;=\; \int_{\mathbb{R}^{d(n+1)}} \tilde g_n(\mathbf{z}-\mathbf{w})\,\zeta_{n,k}(\mathbf{w})\,d\mathbf{w},$$

where $\zeta_{n,k}(\mathbf{z}) = k^{d(n+1)}\zeta_n(k\,\mathbf{z})$, for all $\mathbf{z}\in\mathbb{R}^{d(n+1)}$, with

$$\zeta_n(\mathbf{z}) \;=\; c_n\prod_{\ell=0}^n \exp\bigg(\frac{1}{\mathbf{z}_\ell^2-2^{-2\ell}}\bigg)1_{\{|\mathbf{z}_\ell|<2^{-\ell}\}}, \qquad \text{for all } \mathbf{z}=(\mathbf{z}_0,\ldots,\mathbf{z}_n)\in\mathbb{R}^{d(n+1)}.$$

The constant $c_n>0$ is such that $\int_{\mathbb{R}^{d(n+1)}}\zeta_n(\mathbf{z})\,d\mathbf{z}=1$. We also introduce the double sequence $\{\xi_{n,k}\}_{n\geq0,k\geq1}$, with $\xi_{n,k}\colon C([0,T];\mathbb{R}^d)\to\mathbb{R}$ given by

$$\xi_{n,k}(\boldsymbol{x}) \;=\; g_{n,k}\bigg(\int_{[0,T]}\psi_0(t)\,d^-\boldsymbol{x}(t),\ldots,\int_{[0,T]}\psi_n(t)\,d^-\boldsymbol{x}(t)\bigg), \qquad \boldsymbol{x}\in C([0,T];\mathbb{R}^d),$$

with $\psi_0,\ldots,\psi_n$ as in (C.5). Our aim is to apply Lemma D.1 in [18]. To this end, we need to prove the following items:

a) $\xi_{n,k}$ is continuous;
b) for every $\boldsymbol{x}\in \boldsymbol{C}([0,T];\mathbb{R}^d)$, $|\xi_{n,k}(\boldsymbol{x})-\tilde\xi_n(\boldsymbol{x})|\to0$ as $k\to+\infty$;
c) $\{\xi_{n,k}\}_{n\geq0,k\geq1}$ is equicontinuous on compact sets.

Suppose for a moment that items a)-b)-c) hold true. Then, by Lemma D.1 in [18] we deduce the existence of a subsequence $\{\xi_{n,k_n}\}_n$ converging pointwise to ξ. Hence, we set

$$\xi_n \;:=\; \xi_{n,k_n}, \qquad\qquad g_n \;:=\; g_{n,k_n}.$$

It is then easy to see that $\{\xi_n\}_n$ is the claimed sequence. It remains to prove a)-b)-c). As items a) and b) can be easily proved, we only report the proof of item c).

Step III. Let us prove item c). We begin noting that $g_{n,k}$ can be rewritten as follows:

$$g_{n,k}(\mathbf{z}) \;=\; \int_{\boldsymbol{E}_n}\tilde g_n(\mathbf{z}-\mathbf{w})\,\zeta_{n,k}(\mathbf{w})\,d\mathbf{w},$$

where

$$\boldsymbol{E}_n \;:=\; \big\{\mathbf{z}=(\mathbf{z}_0,\ldots,\mathbf{z}_n)\in\mathbb{R}^{d(n+1)}\colon |\mathbf{z}_\ell|\leq 2^{-\ell},\,\ell=0,\ldots,n\big\}.$$

Then, the claim follows if we prove that for any compact set $\boldsymbol{K}\subset C([0,T];\mathbb{R}^d)$ there exists a continuity modulus $\rho_{\boldsymbol{K}}$ such that, for every n,

$$\text{(C.6)}\quad \bigg|\tilde g_n\bigg(\int_{[0,T]}\psi_0(t)\,d^-\boldsymbol{x}(t)+\mathbf{z}_0,\ldots,\int_{[0,T]}\psi_n(t)\,d^-\boldsymbol{x}(t)+\mathbf{z}_n\bigg) - \tilde g_n\bigg(\int_{[0,T]}\psi_0(t)\,d^-\boldsymbol{x}'(t)+\mathbf{z}_0,\ldots,\int_{[0,T]}\psi_n(t)\,d^-\boldsymbol{x}'(t)+\mathbf{z}_n\bigg)\bigg| \;\leq\; \rho_{\boldsymbol{K}}(\|\boldsymbol{x}-\boldsymbol{x}'\|_\infty),$$

for all $\boldsymbol{x},\boldsymbol{x}'\in\boldsymbol{K}$ and $\mathbf{z}=(\mathbf{z}_0,\ldots,\mathbf{z}_n)\in\boldsymbol{E}_n$. Let us prove (C.6). Given a compact set $\boldsymbol{K}\subset C([0,T];\mathbb{R}^d)$, we denote

$$\boldsymbol{\mathcal{K}} \;:=\; \bigg\{\boldsymbol{x}\in C([0,T];\mathbb{R}^d)\colon \boldsymbol{x} \;=\; T_n\bar{\boldsymbol{x}}+\mathbf{z}_0e_{-1}-\sum_{\ell=1}^n\frac{n+1-\ell}{n+1}\mathbf{z}_\ell(e_\ell-e_\ell(0)), \quad \text{for some } \bar{\boldsymbol{x}}\in\boldsymbol{K},\,\mathbf{z}\in\boldsymbol{E}_n,\,n\geq0\bigg\}.$$

Recalling (C.4), we see that if $\boldsymbol{\mathcal{K}}$ is relatively compact then (C.6) follows from the uniform continuity of ξ on compact sets (which in turn follows from the continuity of ξ). In order to prove that $\mathcal{K}$ is relatively compact, we observe that $\mathcal{K} \subset \mathcal{K}_1 + \mathcal{K}_2 := \{\boldsymbol{x} \in C([0,T];\mathbb{R}^d) \colon \boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2,\, \boldsymbol{x}_1 \in \mathcal{K}_1,\, \boldsymbol{x}_2 \in \mathcal{K}_2\}$, where

$$\mathcal{K}_1 := \big\{\boldsymbol{x} \in C([0,T];\mathbb{R}^d) \colon \boldsymbol{x} \ = \ T_n\bar{\boldsymbol{x}}, \text{ for some } \bar{\boldsymbol{x}} \in \boldsymbol{K},\, n \geq 0\big\},$$

$$\mathcal{K}_2 := \bigg\{\boldsymbol{x} \in C([0,T];\mathbb{R}^d) \colon \boldsymbol{x} \ = \ \mathbf{z}_0 e_{-1} - \sum_{\ell=1}^n \frac{n+1-\ell}{n+1}\mathbf{z}_\ell(e_\ell - e_\ell(0)), \text{ for some } \mathbf{z} \in \boldsymbol{E}_n,\, n \geq 0\bigg\}.$$

If we prove that $\mathcal{K}_1$ and $\mathcal{K}_2$ are relatively compact, it follows that $\mathcal{K}$ is also relatively compact.

- $\mathcal{K}_1$ *is relatively compact.* Let $\{\boldsymbol{x}_h\}_h$ be a sequence in $\mathcal{K}_1$. Let us prove that, up to a subsequence, $\{\boldsymbol{x}_h\}_h$ converges. For each h, there exists $\bar{\boldsymbol{x}}_h \in \boldsymbol{K}$ and $n_h \geq 0$ such that $\boldsymbol{x}_h = T_{n_h}\bar{\boldsymbol{x}}_h$. Suppose that, up to a subsequence, n_h goes to infinity (the proof is simpler when n_h is bounded). Since $\{\bar{\boldsymbol{x}}_h\}_h \subset \boldsymbol{K}$, there exists $\bar{\boldsymbol{x}} \in \boldsymbol{K}$ such that $\{\bar{\boldsymbol{x}}_h\}_h$ converges, up to a subsequence, to $\bar{\boldsymbol{x}}$. Then

$$\|\boldsymbol{x}_h - \bar{\boldsymbol{x}}\|_\infty \ = \ \|T_{n_h}\bar{\boldsymbol{x}}_h - \bar{\boldsymbol{x}}\|_\infty \ \leq \ \|T_{n_h}\bar{\boldsymbol{x}}_h - T_{n_h}\bar{\boldsymbol{x}}\|_\infty + \|T_{n_h}\bar{\boldsymbol{x}} - \bar{\boldsymbol{x}}\|_\infty.$$

  By (C.3) and $\|T_{n_h}\bar{\boldsymbol{x}} - \bar{\boldsymbol{x}}\|_\infty \to 0$, the claim follows.
- $\mathcal{K}_2$ *is relatively compact.* Let $\{\boldsymbol{x}_h\}_h$ be a sequence in $\mathcal{K}_2$. In order to prove that, up to a subsequence, $\{\boldsymbol{x}_h\}_h$ is convergent, we begin noting that

$$\boldsymbol{x}_h \ = \ \mathbf{z}_{0,h}e_{-1} - \sum_{\ell=1}^{n_h} \frac{n_h+1-\ell}{n_h+1}\mathbf{z}_{\ell,h}(e_\ell - e_\ell(0)),$$

  for some $n_h \geq 0$ and $\mathbf{z}_h = (\mathbf{z}_{0,h},\ldots,\mathbf{z}_{n_h,h}) \in \boldsymbol{E}_{n_h}$. Suppose that, up to a subsequence, n_h goes to infinity, otherwise the proof is simpler. Notice that, each sequence $\{\mathbf{z}_{\ell,h}\}_h$ converges, up to a subsequence, to some $\mathbf{z}_\ell$, with $|\mathbf{z}_\ell| \leq 2^{-\ell}$. By Cantor's diagonal argument, there exists a subsequence of $\{\boldsymbol{x}_h\}_h$, which we still denote $\{\boldsymbol{x}_h\}_h$, such that every $\{\mathbf{z}_{\ell,h}\}_h$ converges to $\mathbf{z}_\ell$. We construct this subsequence in such a way that, for every h, $|\mathbf{z}_{0,h} - \mathbf{z}_0| + \cdots + |\mathbf{z}_{n_h,h} - \mathbf{z}_{n_h}| \leq 1/h$. It follows that $\|\boldsymbol{x}_h - \boldsymbol{x}\|_\infty \to 0$, where $\boldsymbol{x} = \mathbf{z}_0 e_{-1} - \sum_{\ell=1}^\infty \mathbf{z}_\ell(e_\ell - e_\ell(0))$.

□

APPENDIX D.
Extended Borwein-Preiss variational principle on $\boldsymbol{\Lambda}$

The proof of the comparison Theorem (Theorem 4.2) is based on Corollary D.1 below, which in turn relies on a generalization of the so-called Borwein-Preiss smooth variant ([6]) of Ekeland's variational principle ([31]), corresponding to Theorem 2.5.2 in [7]. We now state such a generalization in Lemma D.1 for the case of real-valued (rather than $\mathbb{R}\cup\{+\infty\}$-valued as in [7]) maps on $\boldsymbol{\Lambda}$.

Lemma D.1. *Let $G \colon \boldsymbol{\Lambda} \to \mathbb{R}$ be an upper semicontinuous map, bounded from above. Suppose that $\Psi \colon \boldsymbol{\Lambda}\times\boldsymbol{\Lambda} \to [0,+\infty)$ is a gauge-type function (according to Definition 3.1) and $\{\delta_n\}_{n\geq 0}$ is a sequence of strictly positive real numbers. For every $\varepsilon > 0$, let $(t_0,\boldsymbol{x}_0) \in \boldsymbol{\Lambda}$ such that*

$$\sup G - \varepsilon \ \leq \ G(t_0,\boldsymbol{x}_0).$$

Then, there exists a sequence $\{(t_n,\boldsymbol{x}_n)\}_{n\geq 1} \subset \boldsymbol{\Lambda}$ which converges to some $(\bar{t},\bar{\boldsymbol{x}}) \in \boldsymbol{\Lambda}$ satisfying the following.

i) $\Psi((\bar{t},\bar{\boldsymbol{x}}),(t_n,\boldsymbol{x}_n)) \leq \frac{\varepsilon}{2^n\delta_0}$, *for every $n \geq 0$.*
ii) $G(t_0,\boldsymbol{x}_0) \leq G(\bar{t},\bar{\boldsymbol{x}}) - \sum_{n=0}^{+\infty}\delta_n\,\Psi((\bar{t},\bar{\boldsymbol{x}}),(t_n,\boldsymbol{x}_n))$.

iii) *For every* $(t, \boldsymbol{x}) \neq (\bar{t}, \bar{\boldsymbol{x}})$,

$$G(t, \boldsymbol{x}) - \sum_{n=0}^{+\infty} \delta_n \Psi\big((t, \boldsymbol{x}), (t_n, \boldsymbol{x}_n)\big) \;<\; G(\bar{t}, \bar{\boldsymbol{x}}) - \sum_{n=0}^{+\infty} \delta_n \Psi\big((\bar{t}, \bar{\boldsymbol{x}}), (t_n, \boldsymbol{x}_n)\big).$$

Proof. Lemma D.1 follows trivially from Theorem 2.5.2 in [7], the only difference being that this latter result is stated on complete *metric* spaces, while here $\boldsymbol{\Lambda}$ is a complete *pseudometric* space. □

We now apply Lemma D.1 to the smooth gauge-type function ρ_∞ with bounded derivatives defined by (3.20), taking $\delta_0 := \delta > 0$ and $\delta_n := \delta/2^n$, for every $n \geq 1$.

Corollary D.1. *Let* $\delta > 0$ *and* $G\colon \boldsymbol{\Lambda} \to \mathbb{R}$ *be an upper semicontinuous map, bounded from above. For every* $\varepsilon > 0$*, let* $(t_0, \boldsymbol{x}_0) \in \boldsymbol{\Lambda}$ *satisfy*

$$\sup G - \varepsilon \;\leq\; G(t_0, \boldsymbol{x}_0).$$

Then, there exists a sequence $\{(t_n, \boldsymbol{x}_n)\}_{n\geq 1} \subset \boldsymbol{\Lambda}$ *which converges to some* $(\bar{t}, \bar{\boldsymbol{x}}) \in \boldsymbol{\Lambda}$ *such that:*

i) $\rho_\infty((\bar{t}, \bar{\boldsymbol{x}}), (t_n, \boldsymbol{x}_n)) \leq \frac{\varepsilon}{2^n \delta}$*, for every* $n \geq 0$.

ii) $G(t_0, \boldsymbol{x}_0) \leq G(\bar{t}, \bar{\boldsymbol{x}}) - \delta \varphi_\varepsilon(t, \boldsymbol{x})$*, where the map* $\varphi_\varepsilon\colon \boldsymbol{\Lambda} \to [0, +\infty)$ *is defined as*

$$\varphi_\varepsilon(t, \boldsymbol{x}) \;:=\; \sum_{n=0}^{+\infty} \frac{1}{2^n} \rho_\infty\big((t, \boldsymbol{x}), (t_n, \boldsymbol{x}_n)\big), \qquad \forall\, (t, \boldsymbol{x}) \in \boldsymbol{\Lambda}.$$

iii) *For every* $(t, \boldsymbol{x}) \neq (\bar{t}, \bar{\boldsymbol{x}})$,

$$G(t, \boldsymbol{x}) - \delta\, \varphi_\varepsilon(t, \boldsymbol{x}) \;<\; G(\bar{t}, \bar{\boldsymbol{x}}) - \delta\, \varphi_\varepsilon(\bar{t}, \bar{\boldsymbol{x}}).$$

Finally, the map φ_ε *satisfies the following properties.*

1) $\varphi_\varepsilon \in \boldsymbol{C}^{1,2}(\boldsymbol{\Lambda})$ *and is bounded.*
2) $\big|\partial_t^H \varphi_\varepsilon(t, \boldsymbol{x})\big| \leq 2\left(2T + \sqrt{\frac{2}{\pi \mathrm{e}}}\right)$*, for every* $(t, \boldsymbol{x}) \in [0, T) \times C([0, T]; \mathbb{R}^d)$.
3) *For every* $i, j = 1, \dots, d$*,* $\partial_{x_i}^V \varphi_\varepsilon$ *is bounded by the constant* 2 *and* $\partial_{x_i x_j}^V \varphi_\varepsilon$ *is bounded by the constant* $2\left(\sqrt{\frac{2}{\pi}} + 2\right)$.

Proof. Items i)-ii)-iii) follow directly from Lemma D.1, while items 1)-2)-3) follow easily from items 1)-2)-3) of Lemma 3.3. □

Acknowledgments

The work of the second named author was supported by a public grant as part of the *Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH,* in a joint call with Gaspard Monge Program for optimization, operations research and their interactions with data sciences.

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] M. Bambi, G. Fabbri, and F. Gozzi. Optimal policy and consumption smoothing effects in the time-to-build AK model. *Econom. Theory*, 50(3):635–669, 2012.

[3] A. Barrasso and F. Russo. Decoupled mild solutions of path-dependent PDEs and Integro PDEs represented by BSDEs driven by cadlag martingales. *Potential Analysis. To appear*, 2018. Preprint hal-01774823.

[4] E. Bayraktar and C. Keller. Path-dependent Hamilton-Jacobi equations in infinite dimensions. *J. Funct. Anal.*, 275(8):2096–2161, 2018.

[5] A. Bensoussan, M. H. M. Chau, Y. Lai, and S. C. P. Yam. Linear-quadratic mean field Stackelberg games with state and control delays. *SIAM J. Control Optim.*, 55(4):2748–2781, 2017.

[6] 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.

[7] 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.

[8] R. Buckdahn, C. Keller, J. Ma, and J. Zhang. Fully nonlinear stochastic and rough PDEs: Classical and viscosity solutions. *Preprint* arXiv:1501.06978.

[9] R. Carmona, J.-P. Fouque, S. M. Mousavi, and L.-H. Sun. Systemic risk and stochastic games with delay. *J. Optim. Theory Appl.*, 179(2):366–399, 2018.

[10] A. Chojnowska-Michalik. Representation theorem for general stochastic delay equations. *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, 26(7):635–642, 1978.

[11] 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.

[12] R. Cont and D.-A. Fournié. Functional Itô calculus and stochastic integral representation of martingales. *Ann. Probab.*, 41(1):109–133, 2013.

[13] 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.

[14] A. Cosso, C. Di Girolami, and F. Russo. Calculus via regularizations in Banach spaces and Kolmogorov-type path-dependent equations. In *Probability on algebraic and geometric structures*, volume 668 of *Contemp. Math.*, pages 43–65. Amer. Math. Soc., Providence, RI, 2016.

[15] 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.

[16] A. Cosso and F. Russo. Functional and Banach space stochastic calculi: path-dependent Kolmogorov equations associated with the frame of a Brownian motion. In *Stochastics of environmental and financial economics – Centre of Advanced Study, Oslo, Norway, 2014–2015*, volume 138 of *Springer Proc. Math. Stat.*, pages 27–80. Springer, Cham, 2016.

[17] 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.

[18] A. Cosso and F. Russo. Strong-viscosity solutions: classicalg and path-dependent PDEs. *Osaka J. Math.*, 56(2):323–373, 2019.

[19] 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.

[20] 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.

[21] M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277(1):1–42, 1983.

[22] M. G. Crandall and P.-L. Lions. Hamilton-Jacobi equations in infinite dimensions. I. Uniqueness of viscosity solutions. *J. Funct. Anal.*, 62(3):379–396, 1985.

[23] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996.

[24] C. Di Girolami, G. Fabbri, and F. Russo. The covariation for Banach space valued processes and applications. *Metrika*, 77(1):51–104, 2014.

[25] C. Di Girolami and F. Russo. Infinite dimensional stochastic calculus via regularization and applications. *Preprint* HAL-INRIA, inria-00473947 version 1, (Unpublished), 2010.

[26] C. Di Girolami and F. Russo. Clark-Ocone type formula for non-semimartingales with finite quadratic variation. *C. R. Math. Acad. Sci. Paris*, 349(3-4):209–214, 2011.

[27] C. Di Girolami and F. Russo. Generalized covariation and extended Fukushima decomposition for Banach space-valued processes. Applications to windows of Dirichlet processes. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.*, 15(2):1250007, 50, 2012.

[28] C. Di Girolami and F. Russo. Generalized covariation for Banach space valued processes, Itô formula and applications. *Osaka J. Math.*, 51(3):729–783, 2014.

[29] C. Di Girolami and F. Russo. About classical solutions of the path-dependent heat equation. *Random Operators Stochastic Equations (ROSE)*, 2019, to appear.

[30] B. Dupire. Functional Itô calculus. *Portfolio Research Paper*, Bloomberg, 2009.

[31] I. Ekeland. On the variational principle. *J. Math. Anal. Appl.*, 47:324–353, 1974.

[32] I. Ekren, C. Keller, N. Touzi, and J. Zhang. On viscosity solutions of path dependent PDEs. *Ann. Probab.*, 42(1):204–236, 2014.

[33] 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.

[34] 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.

[35] I. Elsanosi, B. Øksendal, and A. Sulem. Some solvable stochastic control problems with delay. *Stochastics Stochastics Rep.*, 71(1-2):69–89, 2000.

[36] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[37] G. Fabbri, F. Gozzi, and A. Święch. *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.

[38] S. Federico. A stochastic control problem with delay arising in a pension fund model. *Finance Stoch.*, 15(3):421–459, 2011.

[39] S. Federico, B. Goldys, and F. Gozzi. HJB equations for the optimal control of differential equations with delays and state constraints, I: regularity of viscosity solutions. *SIAM J. Control Optim.*, 48(8):4910–4937, 2010.

[40] S. Federico, B. Goldys, and F. Gozzi. HJB equations for the optimal control of differential equations with delays and state constraints, II: Verification and optimal feedbacks. *SIAM J. Control Optim.*, 49(6):2378–2414, 2011.

[41] F. Flandoli and G. Zanco. An infinite-dimensional approach to path-dependent Kolmogorov equations. *Ann. Probab.*, 44(4):2643–2693, 2016.

[42] D.-A. Fournié. *Functional Itô calculus and applications*. Ph.D. Thesis - Columbia University, 2010.

[43] A. Friedman. *Stochastic differential equations and applications. Vol. 1.* Academic Press, New York, 1975. Probability and Mathematical Statistics, Vol. 28.

[44] M. Fuhrman, F. Masiero, and G. Tessitore. Stochastic equations with delay: optimal control via BSDEs and regular solutions of Hamilton-Jacobi-Bellman equations. *SIAM J. Control Optim.*, 48(7):4624–4651, 2010.

[45] L. Göllmann, D. Kern, and H. Maurer. Optimal control problems with delays in state and control variables subject to mixed control-state constraints. *Optimal Control Appl. Methods*, 30(4):341–365, 2009.

[46] L. Göllmann and H. Maurer. Optimal control problems with time delays: two case studies in biomedicine. *Math. Biosci. Eng.*, 15(5):1137–1154, 2018.

[47] F. Gozzi and C. Marinelli. Stochastic optimal control of delay equations arising in advertising models. In *Stochastic partial differential equations and applications—VII*, volume 245 of *Lect. Notes Pure Appl. Math.*, pages 133–148. Chapman & Hall/CRC, Boca Raton, FL, 2006.

[48] F. Gozzi, C. Marinelli, and S. Savin. On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects. *J. Optim. Theory Appl.*, 142(2):291–321, 2009.

[49] A. Ichikawa. Quadratic control of evolution equations with delays in control. *SIAM J. Control Optim.*, 20(5):645–668, 1982.

[50] K. Itô and M. Nisio. On stationary solutions of a stochastic differential equation. *J. Math. Kyoto Univ.*, 4:1–75, 1964.

[51] A. F. Ivanov, Y. I. Kazmerchuk, and A. V. Swishchuk. Theory, stochastic stability and applications of stochastic delay differential equations: a survey of results. *Differential Equations Dynam. Systems*, 11(1-2):55–115, 2003.

[52] R. Jensen, M. Kocan, and A. Święch. Good and viscosity solutions of fully nonlinear elliptic equations. *Proc. Amer. Math. Soc.*, 130(2):533–542, 2002.

[53] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[54] A. V. Kim. *Functional differential equations*, volume 479 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 1999. Application of i-smooth calculus.

[55] A. V. Kim. *i-smooth analysis*. Scrivener Publishing, Salem, MA; John Wiley & Sons, Inc., Hoboken, NJ, 2015. Theory and applications.

[56] A. V. Kim and A. V. Ivanov. *Systems with delays*. John Wiley & Sons, Inc., Hoboken, NJ; Scrivener Publishing, Salem, MA, 2015. Analysis, control, and computations.

[57] B. Larssen and N. H. Risebro. When are HJB-equations in stochastic control of delay systems finite dimensional? *Stochastic Anal. Appl.*, 21(3):643–671, 2003.

[58] D. Leão, A. Ohashi, and A. B. Simas. A weak version of path-dependent functional Itô calculus. *Ann. Probab.*, 46(6):3399–3441, 2018.

[59] P.-L. Lions. On the Hamilton-Jacobi-Bellman equations. *Acta Appl. Math.*, 1(1):17–41, 1983.

[60] P.-L. Lions. Optimal control of diffusion processes and Hamilton-Jacobi-Bellman equations. I. The dynamic programming principle and applications. *Comm. Partial Differential Equations*, 8(10):1101–1174, 1983.

[61] 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.

[62] P.-L. Lions. Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. I. The case of bounded stochastic evolutions. *Acta Math.*, 161(3-4):243–278, 1988.

[63] P.-L. Lions. Viscosity solutions of fully nonlinear second order equations and optimal stochastic control in infinite dimensions. II. Optimal control of Zakai's equation. In *Stochastic partial differential equations and applications, II (Trento, 1988)*, volume 1390 of *Lecture Notes in Math.*, pages 147–170. Springer, Berlin, 1989.

[64] P.-L. Lions. Viscosity solutions of fully nonlinear second-order equations and optimal stochastic control in infinite dimensions. III. Uniqueness of viscosity solutions for general second-order equations. *J. Funct. Anal.*, 86(1):1–18, 1989.

[65] N. Y. Lukoyanov. A minimax solution of functional equations of Hamilton-Jacobi type for hereditary systems. *Differ. Equ.*, 37(2):246–256, 2001.

[66] N. Y. Lukoyanov. Functional Hamilton-Jacobi type equations in ci-derivatives for systems with distributed delays. *Nonlinear Funct. Anal. Appl.*, 8(3):365–397, 2003.

[67] N. Y. Lukoyanov. Functional Hamilton-Jacobi type equations with ci-derivatives in control problems with hereditary information. *Nonlinear Funct. Anal. Appl.*, 8(4):535–555, 2003.

[68] N. Y. Lukoyanov. Viscosity solution of nonanticipating equations of Hamilton-Jacobi type. *Differ. Equ.*, 43(12):1715–1723, 2007.

[69] S. E. A. Mohammed. *Stochastic functional differential equations*, volume 99 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[70] S. E. A. Mohammed. Stochastic differential systems with memory: theory, examples and applications. In *Stochastic analysis and related topics, VI (Geilo, 1996)*, volume 42 of *Progr. Probab.*, pages 1–77. Birkhäuser Boston, Boston, MA, 1998.

[71] H. Oberhauser. The functional Itô formula under the family of continuous semimartingale measures. *Stoch. Dyn.*, 16(4):1650010, 26, 2016.

[72] B. Øksendal, A. Sulem, and T. Zhang. Optimal control of stochastic delay equations and time-advanced backward stochastic differential equations. *Adv. in Appl. Probab.*, 43(2):572–596, 2011.

[73] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[74] S. Peng. Note on viscosity solution of path-dependent PDE and G-martingales. *Preprint* arXiv: 1106.1144v2, 2012.

[75] 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.

[76] S. Peng and F. Wang. BSDE, path-dependent PDE and nonlinear Feynman-Kac formula. *Sci. China Math.*, 59(1):19–36, 2016.

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

[78] Z. Ren. Viscosity solutions of fully nonlinear elliptic path dependent partial differential equations. *Ann. Appl. Probab.*, 26(6):3381–3414, 2016.

[79] Z. Ren, N. Touzi, and J. Zhang. Comparison of viscosity solutions of semi-linear path-dependent PDEs. *Preprint* arXiv:1410.7281.

[80] 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.

[81] F. Rodrigues, C. J. Silva, D. F. M. Torres, and H. Maurer. Optimal control of a delayed HIV model. *Discrete Contin. Dyn. Syst. Ser. B*, 23(1):443–458, 2018.

[82] F. Russo and P. Vallois. The generalized covariation process and Itô formula. *Stochastic Process. Appl.*, 59(1):81–104, 1995.

[83] F. Russo and P. Vallois. Elements of stochastic calculus via regularization. In *Séminaire de Probabilités XL*, volume 1899 of *Lecture Notes in Math.*, pages 147–185. Springer, Berlin, 2007.

[84] M. Scheutzow. Qualitative behaviour of stochastic delay equations with a bounded memory. *Stochastics*, 12(1):41–80, 1984.

[85] H. M. Soner. On the Hamilton-Jacobi-Bellman equations in Banach spaces. *J. Optim. Theory Appl.*, 57(3):429–437, 1988.

[86] A. I. Subbotin. Generalization of the fundamental equation of the theory of differential games. *Dokl. Akad. Nauk SSSR*, 254(2):293–297, 1980.

[87] A. I. Subbotin and N. N. Subbotina. Necessary and sufficient conditions for the piecewise smooth value of a differential game. *Dokl. Akad. Nauk SSSR*, 243(4):862–865, 1978.

[88] A. Święch. "Unbounded" second order partial differential equations in infinite-dimensional Hilbert spaces. *Comm. Partial Differential Equations*, 19(11-12):1999–2036, 1994.

[89] 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.

[90] R. B. Vinter and R. H. Kwong. The infinite time quadratic control problem for linear systems with state and control delays: an evolution equation approach. *SIAM J. Control Optim.*, 19(1):139–153, 1981.

[91] A. Zygmund. *Trigonometric series. Vol. I, II.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2002.

ANDREA COSSO, UNIVERSITY OF BOLOGNA, DEPARTMENT OF MATHEMATICS, PIAZZA DI PORTA SAN DONATO 5, 40126 BOLOGNA, ITALY
Email address: andrea.cosso@unibo.it

FRANCESCO RUSSO, ENSTA PARIS, INSTITUT POLYTECHNIQUE DE PARIS, UNITÉ DE MATHÉMATIQUES APPLIQUÉES, 828, BOULEVARD DES MARÉCHAUX, F-91120 PALAISEAU, FRANCE
Email address: francesco.russo@ensta-paris.fr