# Nonlinear continuous semimartingales[*]

David Criens[†]    Lars Niemann[‡]

### Abstract

In this paper we study a family of nonlinear (conditional) expectations that can be understood as a continuous semimartingale with uncertain local characteristics. Here, the differential characteristics are prescribed by a set-valued function that depends on time and path in a non-Markovian way. We provide a dynamic programming principle for the nonlinear expectation and we link the corresponding value function to a variational form of a nonlinear path-dependent partial differential equation. In particular, we establish conditions that allow us to identify the value function as the unique viscosity solution. Furthermore, we prove that the nonlinear expectation solves a nonlinear martingale problem, which confirms our interpretation as a nonlinear semimartingale.

## 1 Introduction

In this paper we study a family of nonlinear (conditional) expectations which we call *nonlinear continuous semimartingale* and which we consider as a continuous semimartingale with uncertain local characteristics. This line of research started with the seminal work of Peng [35, 36] on the $G$-Brownian motion. In recent years, there have been several extensions to construct larger classes of nonlinear (Markov) processes, see [12, 18, 32, 33]. At this point we highlight the articles [11, 31, 34] which establish general abstract measure theoretic concepts to construct nonlinear expectations. Next to this approach, nonlinear Markov processes have also been constructed via nonlinear

[†]Albert-Ludwigs University of Freiburg, Germany.
   E-mail: david.criens@stochastik.uni-freiburg.de
[‡]Albert-Ludwigs University of Freiburg, Germany.
   E-mail: lars.niemann@stochastik.uni-freiburg.de

semigroups, see [7, 17, 26, 27] and, in particular, [17, Chapter 4] for a comparison of the methods.

This paper investigates nonlinear processes with non-Markovian dynamics. We define a nonlinear expectation via

$$\mathcal{E}_t(\psi)(\omega) := \sup_{P \in \mathcal{C}(t,\omega)} E^P[\psi], \quad (t,\omega) \in \mathbb{R}_+ \times C(\mathbb{R}_+; \mathbb{R}^d),$$

where $\psi \colon C(\mathbb{R}_+; \mathbb{R}^d) \to \mathbb{R}$ is an upper semianalytic function and $\mathcal{C}(t,\omega)$ is a set of probability measures on the Wiener space which give point mass to the path $\omega$ till time $t$ and afterwards coincide with the law of a semimartingale with absolutely continuous characteristics. In this paper we parameterize drift and quadratic variation by a compact parameter space $F$ and two functions $b \colon F \times \mathbb{R}_+ \times C(\mathbb{R}_+; \mathbb{R}^d) \to \mathbb{R}^d$ and $a \colon F \times \mathbb{R}_+ \times C(\mathbb{R}_+; \mathbb{R}^d) \to \mathbb{S}_+^d$ such that

$$\mathcal{C}(t,\omega) := \Big\{ P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}(t) \colon P(X = \omega \text{ on } [0,t]) = 1,$$
$$(\lambda \otimes P)\text{-a.e. } (dB_{\cdot+t}^P/d\lambda, dC_{\cdot+t}^P/d\lambda) \in \Theta(\cdot + t, X) \Big\},$$

where

$$\Theta(t,\omega) := \big\{ (b(f,t,\omega), a(f,t,\omega)) \colon f \in F \big\},$$

and $\mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}(t)$ denotes the set of semimartingale laws after $t$ with absolutely continuous characteristics. This framework includes nonlinear Lévy processes as introduced (with jumps) in [32] and the class of nonlinear affine processes as studied in [12]. Furthermore, our setting can also be used to model path-dependent dynamics such as stochastic delay equations under parameter uncertainty.

For this nonlinear expectation we prove the dynamic programming principle (DPP), i.e., we prove the tower property

$$\mathcal{E}_\sigma(\psi) = \mathcal{E}_\sigma(\mathcal{E}_\tau(\psi))$$

for all finite stopping times $\tau \geq \sigma$. To prove the DPP we use an abstract theorem from [11]. The work lies in the verification of its prerequisites. To check them we extend certain results from [31] on the measurability of the semimartingale property and the behavior of the characteristics to a dynamic framework.

In a second step we identify two properties of $\mathcal{E}$ which confirm our interpretation as a nonlinear continuous semimartingale. First, we relate the value function

$$v(t,\omega) := \mathcal{E}_t(\psi)(\omega) = \sup_{P \in \mathcal{C}(t,\omega)} E^P[\psi] \tag{1.1}$$

to a path-dependent Kolmogorov type partial differential equation and, in the second part, we show that $\mathcal{E}$ solves a type of nonlinear martingale problem. Let us discuss our contributions in more detail.

Under mere continuity and linear growth conditions on the drift $b$ and the quadratic variation $a$, and under the hypothesis that the set $\{(b(f,t,\omega), a(f,t,\omega)) \colon f \in F\}$ is convex for every $(t,\omega) \in \mathbb{R}_+ \times C(\mathbb{R}_+, \mathbb{R}^d)$, we show that the value function $v$ is a weak sense viscosity solution, i.e., a path-dependent Crandall–Lions type viscosity solution without regularity properties, to the following path-dependent partial differential equation (PPDE):

$$\begin{cases} \dot\partial v(t,\omega) + G(t,\omega,v) = 0, & \text{for } (t,\omega) \in [0,T) \times C(\mathbb{R}_+; \mathbb{R}^d), \\ v(T,\omega) = \psi(\omega), & \text{for } \omega \in C(\mathbb{R}_+; \mathbb{R}^d), \end{cases} \tag{1.2}$$

where

$$G(t,\omega,\phi) := \sup\left\{\langle\nabla\phi(t,\omega), b(f,t,\omega)\rangle + \tfrac{1}{2}\operatorname{tr}\left[\nabla^2\phi(t,\omega)a(f,t,\omega)\right] : f \in F\right\}. \qquad (1.3)$$

The proof for the viscosity property is split into two parts, i.e., we prove the sub- and the supersolution property. The ideas of proof are based on applications of Berge's maximum theorem, Skorokhod's existence theorem for stochastic differential equations and Lebesgue's differentiation theorem. In contrast to the proofs from [12, 32] for the viscosity subsolution property in Lévy and continuous affine frameworks respectively, we do not work with explicit moment estimates. This allows us to extend the class of test functions in our framework to $C^{1,2}$ in comparison to the class $C^{2,3}$ as used in [12, 32].

Further, we investigate when the value function is not only a *weak sense* viscosity solution but has in addition some regularity properties. By virtue of the last term in (1.1) and Berge's maximum theorem, it is a natural idea to deduce regularity properties of $v$ from corresponding properties of the set-valued map $(t,\omega) \mapsto \mathcal{C}(t,\omega)$. To the best of our knowledge, the idea to deduce regularity properties of value functions from related properties of set-valued maps traces back to the seminal paper [10] that investigates a controlled diffusion framework. Due to the appearance of $\mathfrak{P}^{\mathrm{ac}}_{\mathrm{sem}}(t)$ and $(dB^P_{\cdot+t}/d\lambda, dC^P_{\cdot+t}/d\lambda)$ in the definition of the set $\mathcal{C}(t,\omega)$, regularity properties of $(t,\omega) \mapsto \mathcal{C}(t,\omega)$ seem at first glance to be difficult to verify. To get a more convenient condition, we show that

$$v(t,\omega) = \sup_{P\in\mathcal{C}(t,\omega)} E^P\big[\psi\big] = \sup_{P\in\mathcal{R}(t,\omega)} E^P\big[\psi(\omega\,\widetilde{\otimes}_t X)\big],$$

where

$$\mathcal{R}(t,\omega) := \Big\{P \in \mathfrak{P}^{\mathrm{ac}}_{\mathrm{sem}} : P \circ X_0^{-1} = \delta_{\omega(t)},$$
$$(\lambda \otimes P)\text{-a.e. } (dB^P/d\lambda, dC^P/d\lambda) \in \Theta(\cdot + t, \omega\,\widetilde{\otimes}_t X)\Big\},$$

and

$$\omega\,\widetilde{\otimes}_t \omega' := \omega\mathbb{1}_{[0,t)} + (\omega(t) + \omega'(\cdot - t) - \omega'(0))\mathbb{1}_{[t,\infty)}.$$

This reformulation of the value function $v$ explains that it suffices to investigate the correspondence $(t,\omega) \mapsto \mathcal{R}(t,\omega)$ and it connects the two (seemingly closely related) approaches from [11] and [34] for the construction of nonlinear expectations. We show that $(t,\omega) \mapsto \mathcal{R}(t,\omega)$ is upper hemicontinuous and compact-valued, which establishes upper semicontinuity of $v$. This requires a profound analysis of the limiting behaviour of semimartingale characteristics and hinges on the continuity of the correspondence $(t,\omega) \mapsto \Theta(t,\omega)$. We also present a counterexample which explains that mere continuity and linear growth of $b$ and $a$ are insufficient for lower hemicontinuity of $(t,\omega) \mapsto \mathcal{R}(t,\omega)$.

Provided $b$ and $a$ are additionally locally Lipschitz continuous, we prove lower hemicontinuity of $(t,\omega) \mapsto \mathcal{R}(t,\omega)$. To this end, we combine arguments based on an implicit function theorem, a strong existence property for stochastic differential equations with random locally Lipschitz coefficients and Gronwall's lemma. To the best of our knowledge, this is the first result regarding lower hemicontinuity in a path-dependent setting related to nonlinear stochastic processes.

Under a uniform Lipschitz continuity condition on the coefficients $b$ and $a$ and the terminal function $\psi$, we also show that the value function $v$ has a certain Lipschitz property w.r.t. the function

$$\mathsf{d}((t,\omega),(s,\alpha)) := \Big(1 + \sup_{r\in[0,t]}\|\omega(r)\| + \sup_{r\in[0,s]}\|\alpha(r)\|\Big)|t-s|^{1/2}$$
$$+ \sup_{r\in[0,T]}\|\omega(r\wedge t) - \alpha(r\wedge s)\|,$$

for $(t,\omega),(s,\alpha) \in [0,T] \times C([0,T];\mathbb{R}^d)$. This observation allows us to invoke a novel

uniqueness result from [43] that identifies $v$ as the unique viscosity solution to the PPDE (1.2) that is bounded and Lipschitz continuous w.r.t. d.

Next, we discuss the martingale problem related to $\mathcal{E}$. In case the coefficients $b$ and $a$ are compactly parameterized independently of each other, and under linear growth and mere continuity assumptions, for suitable test functions $\phi$, we show that the process

$$\phi(X_t) - \int_0^t G(s, X, \phi)ds, \quad t \in \mathbb{R}_+,$$

is a local $\mathcal{E}$-martingale. This seems to be a novel connection between nonlinear martingale problems and nonlinear processes in a path-dependent setting. Our proof is based on applications of a measurable maximum theorem and Skorokhod's existence theorem.

This paper is structured as follows: in Section 2 we introduce our setting. In Section 3 we state the DPP and in Section 4 we discuss the relation of the correspondence $\mathcal{R}$ to the PPDE (1.2). In Section 5 the nonlinear martingale problem is stated. The proofs for our main results are given in the remaining sections. More precisely, the DPP is proved in Section 6, the viscosity property is proved in Section 7, regularity of $\mathcal{R}$ and $v$ are established in Section 8 and the martingale problem is proved in Section 9. The Lévy case with constant uncertainty set $\Theta$ allows for a simplified proof of the continuity of $\mathcal{R}$. We detail the argument in Appendix A.

## 2 The setting

Let $d \in \mathbb{N}$ be a fixed dimension and define $\Omega$ to be the space of continuous functions $\mathbb{R}_+ \to \mathbb{R}^d$ endowed with the local uniform topology. The Euclidean scalar product and the corresponding Euclidean norm are denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$. We write $X$ for the canonical process on $\Omega$, i.e., $X_t(\omega) = \omega(t)$ for $\omega \in \Omega$ and $t \in \mathbb{R}_+$. It is well-known that $\mathcal{F} := \mathcal{B}(\Omega) = \sigma(X_t, t \geq 0)$. We define $\mathbf{F} := (\mathcal{F}_t)_{t \geq 0}$ as the canonical filtration generated by $X$, i.e., $\mathcal{F}_t := \sigma(X_s, s \leq t)$ for $t \in \mathbb{R}_+$. Notice that we do not make the filtration $\mathbf{F}$ right-continuous. The set of probability measures on $(\Omega, \mathcal{F})$ is denoted by $\mathfrak{P}(\Omega)$ and endowed with the usual topology of convergence in distribution. Let $F$ be a metrizable space and let $b \colon F \times \mathbb{R}_+ \times \Omega \to \mathbb{R}^d$ and $a \colon F \times \mathbb{R}_+ \times \Omega \to \mathbb{S}_+^d$ be Borel functions such that $(t, \omega) \mapsto b(f, t, \omega)$ and $(t, \omega) \mapsto a(f, t, \omega)$ are predictable for every $f \in F$. Here, $\mathbb{S}_+^d$ denotes the space of all real-valued symmetric positive semidefinite $d \times d$ matrices. We define the correspondence, i.e., the set-valued mapping, $\Theta \colon \mathbb{R}_+ \times \Omega \twoheadrightarrow \mathbb{R}^d \times \mathbb{S}_+^d$ by

$$\Theta(t, \omega) := \big\{ (b(f, t, \omega), a(f, t, \omega)) \colon f \in F \big\}.$$

**Standing Assumption 2.1.** $\Theta$ *has a measurable graph, i.e., the graph*

$$\operatorname{gr} \Theta = \big\{ (t, \omega, b, a) \in \mathbb{R}_+ \times \Omega \times \mathbb{R}^d \times \mathbb{S}_+^d \colon (b, a) \in \Theta(t, \omega) \big\}$$

*is Borel.*

In Lemma 2.8 below we will see that this standing assumption holds once $F$ is compact and $b$ and $a$ are continuous in the $F$ variable.

We call an $\mathbb{R}^d$-valued continuous process $Y = (Y_t)_{t \geq 0}$ a (continuous) *semimartingale after a time* $t^* \in \mathbb{R}_+$ if the process $Y_{\cdot + t^*} = (Y_{t+t^*})_{t \geq 0}$ is a $d$-dimensional semimartingale for its natural right-continuous filtration. Notice that it comes without loss of generality that we consider the right-continuous version of the filtration (see [31, Proposition 2.2]). The law of a semimartingale after $t^*$ is said to be a *semimartingale law after* $t^*$ and the set of them is denoted by $\mathfrak{P}_{\text{sem}}(t^*)$. Notice also that $P \in \mathfrak{P}_{\text{sem}}(t^*)$ if and only if the coordinate process is a semimartingale after $t^*$, see Lemma 6.4 below. For $P \in \mathfrak{P}_{\text{sem}}(t^*)$ we denote the semimartingale characteristics of the shifted coordinate process $X_{\cdot + t^*}$ by $(B_{\cdot + t^*}^P, C_{\cdot + t^*}^P)$. Moreover, we set

$$\mathfrak{P}_{\text{sem}}^{\text{ac}}(t^*) := \big\{ P \in \mathfrak{P}_{\text{sem}}(t^*) \colon P\text{-a.s. } (B_{\cdot + t^*}^P, C_{\cdot + t^*}^P) \ll \lambda \big\}, \quad \mathfrak{P}_{\text{sem}}^{\text{ac}} := \mathfrak{P}_{\text{sem}}^{\text{ac}}(0),$$

where $\lambda$ denotes the Lebesgue measure. For $\omega, \omega' \in \Omega$ and $t \in \mathbb{R}_+$, we define the concatenation

$$\omega \otimes_t \omega' := \omega \mathbb{1}_{[0,t)} + (\omega(t) + \omega' - \omega'(t)) \mathbb{1}_{[t,\infty)}.$$

Finally, for $(t, \omega) \in \mathbb{R}_+ \times \Omega$, we define $\mathcal{C}(t, \omega) \subset \mathfrak{P}(\Omega)$ by

$$\mathcal{C}(t, \omega) := \Big\{ P \in \mathfrak{P}^{\mathrm{ac}}_{\mathrm{sem}}(t) \colon P(X^t = \omega^t) = 1,$$

$$(\lambda \otimes P)\text{-a.e. } (dB^P_{\cdot+t}/d\lambda, dC^P_{\cdot+t}/d\lambda) \in \Theta(\cdot + t, \omega \otimes_t X) \Big\},$$

where we use the standard notation $X^t := X_{\cdot \wedge t}$.

To lighten our notation, let us further define, for two stopping times $S$ and $T$, the stochastic interval

$$[\![S, T[\![ := \{(t, \omega) \in \mathbb{R}_+ \times \Omega \colon S(\omega) \le t < T(\omega)\}.$$

The stochastic intervals $]\!]S, T[\![, [\![S, T]\!], ]\!]S, T]\!]$ are defined accordingly. In particular, the equality $[\![0, \infty[\![ = \mathbb{R}_+ \times \Omega$ holds.

**Standing Assumption 2.2.** $\mathcal{C}(t, \omega) \ne \varnothing$ for all $(t, \omega) \in [\![0, \infty[\![$.

In Lemma 2.10 below we will see that this standing assumption holds under continuity and linear growth conditions on $b$ and $a$.

**Remark 2.3.**

(i) The sets $(\mathcal{C}(t, \omega))_{(t,\omega) \in [\![0,\infty[\![}$ are adapted in the following sense: for each $t \in \mathbb{R}_+$, the set $\mathcal{C}(t, \omega)$ depends only on the path of $\omega$ up to time $t$.

(ii) Notice that

$$\mathcal{C}(t, \omega) = \Big\{ P \in \mathfrak{P}^{\mathrm{ac}}_{\mathrm{sem}}(t) \colon P(X^t = \omega^t) = 1,$$

$$(\lambda \otimes P)\text{-a.e. } (dB^P_{\cdot+t}/d\lambda, dC^P_{\cdot+t}/d\lambda) \in \Theta(\cdot + t, X) \Big\}.$$

We defined $\mathcal{C}(t, \omega)$ with the seemingly more complicated ingredient $\Theta(\cdot + t, \omega \otimes_t X)$, instead of $\Theta(\cdot + t, X)$, to prevent confusions about measurability. Namely, both $(dB^P_{\cdot+t}/d\lambda, dC^P_{\cdot+t}/d\lambda)$ and $\Theta(\cdot+t, \omega \otimes_t X)$ are measurable with respect to $\sigma(X_s, s \ge t)$, while $\Theta(\cdot + t, X)$ might also depend on $X_s, s < t$.

In the following we state and discuss some conditions needed to formulate our main results.

**Condition 2.4** (Linear Growth). *For every $T > 0$, there exists a constant $C = C_T > 0$ such that*

$$\|b(f, t, \omega)\|^2 + \mathrm{tr}\left[a(f, t, \omega)\right] \le C \Big(1 + \sup_{s \in [0,t]} \|\omega(s)\|^2\Big)$$

*for all $f \in F$ and $(t, \omega) \in [\![0, T]\!]$.*

**Condition 2.5** (Continuity in Control). *For each $(t, \omega) \in [\![0, \infty[\![$, the maps $f \mapsto b(f, t, \omega)$ and $f \mapsto a(f, t, \omega)$ are continuous.*

**Condition 2.6** (Joint Continuity). *The functions $(f, t, \omega) \mapsto b(f, t, \omega)$ and $(f, t, \omega) \mapsto a(f, t, \omega)$ are continuous.*

**Condition 2.7** (Convexity). *For every $(t, \omega) \in [\![0, \infty[\![$, the set $\{(b(f, t, \omega), a(f, t, \omega)) \colon f \in F\} \subset \mathbb{R}^d \times \mathbb{S}^d_+$ is convex.*

Before we present our main results, let us shortly show that our standing assumptions hold under some of the above conditions. We start with Standing Assumption 2.1.

**Lemma 2.8.** *If $F$ is a compact metrizable space and Condition 2.5 holds, then the correspondence $\Theta$ has measurable graph, i.e., Standing Assumption 2.1 holds.*

The previous lemma is a direct consequence of the following general observation.

**Lemma 2.9.** *Let $(\Sigma, \mathcal{G})$ be a measurable space, let $F$ be a compact metrizable space, let $E$ be a separable metrizable space, and finally let $g: F \times \Sigma \to E$ be a Carathéodory function, i.e., $g$ is continuous in the first and measurable in the second variable. Then, the correspondence $\varphi$ defined by $\varphi(\sigma) := \{g(f, \sigma): f \in F\}$ has a measurable graph.*

*Proof.* First, we show that $\varphi$ is weakly measurable, that is, for every open set $U \subset E$ the lower inverse $\varphi^l(U) := \{\sigma \in \Sigma: \varphi(\sigma) \cap U \neq \varnothing\}$ is measurable. As $F$ is compact, there exists a countable dense subset $\{f_n: n \in \mathbb{N}\} \subset F$. Since $g$ is a Carathéodory function, the functions $\{g_n: n \in \mathbb{N}\}$, where $g_n(\sigma) := g(f_n, \sigma)$, form a Castaing representation of $g(F, \cdot) \subset E$, i.e., a countable family of measurable functions such that

$$\mathrm{cl}(\{g_n(\sigma): n \in \mathbb{N}\}) = g(F, \sigma),$$

for every $\sigma \in \Sigma$. Thus, for each open subset $U \subset E$,

$$\{\sigma \in \Sigma: \varphi(\sigma) \cap U \neq \varnothing\} = \{\sigma \in \Sigma: g(F, \sigma) \cap U \neq \varnothing\} = \bigcup_{n \in \mathbb{N}} \{g_n \in U\} \in \mathcal{G}.$$

Moreover, as $\varphi$ is compact-valued, and in particular closed-valued, it follows from [1, Theorem 18.6] that $\varphi$ has measurable graph. $\square$

Moreover, also our second standing assumption follows from some conditions above.

**Lemma 2.10.** *If the Conditions 2.4 and 2.6 hold, then $\mathcal{C}(t, \omega) \neq \varnothing$ for all $(t, \omega) \in [\![0, \infty[\![$, i.e., Standing Assumption 2.2 holds.*

*Proof.* This follows from Lemmata 7.1, 7.2 and 7.11 below. $\square$

## 3 The dynamic programming principle

We now define a nonlinear expectation and prove the dynamic programming principle (DPP), i.e., the tower property. Suppose that $\psi: \Omega \to [-\infty, \infty]$ is an upper semianalytic function, i.e., $\{\omega \in \Omega: \psi(\omega) > c\}$ is analytic for every $c \in \mathbb{R}$, and define the so-called *value function* by

$$v(t, \omega) := \sup_{P \in \mathcal{C}(t, \omega)} E^P[\psi], \quad (t, \omega) \in [\![0, \infty[\![.$$

For every finite stopping time $\tau$, we set $v(\tau, X) := v(\tau(X), X)$.

**Theorem 3.1** (Dynamic Programming Principle)**.** *The value function $v$ is upper semianalytic. Moreover, for every pair $(t, \omega) \in [\![0, \infty[\![$ and every stopping time $\tau$ with $t \leq \tau < \infty$, we have*

$$v(t, \omega) = \sup_{P \in \mathcal{C}(t, \omega)} E^P[v(\tau, X)]. \tag{3.1}$$

The value function can be interpreted as a nonlinear expectation $\mathcal{E}$ given by

$$\mathcal{E}_t(\psi)(\omega) := v(t, \omega), \qquad (t, \omega) \in [\![0, \infty[\![.$$

The DPP in Theorem 3.1 provides the tower rule for $\mathcal{E}$. Namely, (3.1) means that

$$\mathcal{E}_t(\psi)(\omega) = \mathcal{E}_t(\mathcal{E}_\tau(\psi))(\omega), \quad t \leq \tau \leq \infty. \tag{3.2}$$

By its pathwise structure, the equality (3.2) also implies that

$$\mathcal{E}_\sigma(\psi) = \mathcal{E}_\sigma(\mathcal{E}_\tau(\psi)),$$

for all finite stopping times $\tau \geq \sigma$. To prove Theorem 3.1 we use a general theorem from [11]. The work lies in the verification of the prerequisites, which are (i) a measurable graph property, (ii) a stability property for conditioning, and (iii) a stability property for pasting. The proof is given in Section 6 below.

**Example 3.2.** In the following we mention some examples of stochastic models that are covered by our framework. We stress that it includes many previously studied frameworks but also some new ones which are of interest for future investigations.

(i) The case where $\Theta(t, \omega) \equiv \Theta$ is independent of time $t$ and path $\omega$ corresponds to the generalized $G$-Brownian motion as introduced in [37], cf. also [32] for a nonlinear Lévy setting *with jumps*.

(ii) The situation where $\Theta(t, \omega) \equiv \Theta(\omega(t))$ depends on $(t, \omega)$ only through the value $\omega(t)$ corresponds to a Markovian setting that has, for instance, been studied in [5, 17].

(iii) The example where, for $(f, t, \omega) \in F \times [\![0, \infty[\![$ and $r > 0$,

$$b(f, t, \omega) := b_0(f)\omega(t) + \int_{(t-r)\vee 0}^t b_1(f, s)\omega(s)ds, \qquad a(f, t, \omega) := a_0(f),$$

with Borel functions

$$b_0 \colon F \to \mathbb{R}, \quad b_1 \colon F \times \mathbb{R}_+ \to \mathbb{R}, \quad a_0 \colon F \to \mathbb{R}_+,$$

corresponds to a class of linear stochastic delay differential equations with parameter uncertainty (that is captured by the $F$-dependence of $b$ and $a$). Control problems for such a path-dependent setting were studied in [14]. From a modeling perspective, natural assumptions such as sign-constraints and ellipticity can be incorporated through the functions $b_0, b_1$ and $a_0$, respectively. More references involving control problems for delay equations can be found in [3]. Our framework seems to be the first that captures stochastic delay equations from the perspective of nonlinear stochastic processes.

Given the DPP, we proceed studying more properties of $v$ and $\mathcal{E}$. In the following section we identify $v$ as a viscosity solution to a certain nonlinear path-dependent partial differential equation (PPDE) and, in Section 5, we identify $\mathcal{E}$ as a solution to a nonlinear martingale problem.

## 4 The nonlinear Kolmogorov equation

In the following we discuss the relation of the value function to a path-dependent Kolmogorov type partial differential equation. This section is structured as follows. In Section 4.1 we define the class of test functions for the concept of Crandall–Lions type viscosity solutions in our path-dependent setting. The Kolmogorov type equation is introduced in Section 4.2. Finally, we present our main results in Section 4.3.

### 4.1 The class of test functions

In the first part of this section we introduce the set of test functions for the concept of Crandall–Lions type viscosity solutions in our path-dependent setting. The following definitions are adapted from [3, 4]. Let $T > 0$ and $t_0 \in [0, T)$. Further, let $D(\mathbb{R}_+; \mathbb{R}^d)$ be

the space of càdlàg functions from $\mathbb{R}_+$ into $\mathbb{R}^d$. We define $\Lambda(t_0) := [t_0, T] \times D(\mathbb{R}_+; \mathbb{R}^d)$ and, on $[0, T] \times D(\mathbb{R}_+; \mathbb{R}^d)$, we further define the pseudometric $d$ as

$$d((t, \omega), (s, \omega')) := |t - s| + \sup_{r \in [0,T]} \|\omega(r \wedge t) - \omega'(r \wedge s)\|. \tag{4.1}$$

We denote the restriction of $d$ to $\Lambda(t_0)$ again by $d$. For a map $F \colon \Lambda(t_0) \to \mathbb{R}$ we say that $F$ admits a *horizontal derivative* at $(t, \omega) \in \Lambda(t_0)$ with $t < T$ if

$$\dot{\partial}F(t, \omega) := \lim_{h \searrow 0} \frac{F(t + h, \omega(\cdot \wedge t)) - F(t, \omega(\cdot \wedge t))}{h}$$

exists. At $t = T$, the horizontal derivative is defined as

$$\dot{\partial}F(T, \omega) := \lim_{h \nearrow T} \dot{\partial}F(h, \omega).$$

Further, we say that $F$ admits a *vertical derivative* at $(t, \omega) \in \Lambda(t_0)$ if

$$\partial_i F(t, \omega) := \lim_{h \to 0} \frac{F(t, \omega + h e_i \mathbb{1}_{[t,T]}) - F(t, \omega)}{h}, \quad i = 1, 2, \ldots, d,$$

exist, where $e_1, \ldots, e_d$ are the standard unit vectors in $\mathbb{R}^d$. Accordingly, the second vertical derivatives $\partial_{ij}^2 F(t, \omega), i, j = 1, \ldots, d$, at $(t, \omega) \in \Lambda(t_0)$ are defined as

$$\partial_{ij}^2 F(t, \omega) := \partial_i(\partial_j F)(t, \omega).$$

We write $\nabla F := (\partial_1 F, \ldots, \partial_d F)$ for the *vertical gradient* and $\nabla^2 F := (\partial_{ij}^2 F)_{i,j=1,\ldots,d}$ for the *vertical Hessian matrix*.

Next, we denote by $C^{1,2}(\Lambda(t_0); \mathbb{R})$ the set of functions $F \colon \Lambda(t_0) \to \mathbb{R}$, continuous with respect to $d$, such that

$$\dot{\partial}F, \nabla F, \nabla^2 F$$

exist everywhere on $\Lambda(t_0)$ and are continuous with respect to $d$.

The set $C^{1,2}(\llbracket t_0, T \rrbracket; \mathbb{R})$ consists of functions $F \colon \llbracket t_0, T \rrbracket \to \mathbb{R}$ such that there exists a function $\hat{F} \in C^{1,2}(\Lambda(t_0); \mathbb{R})$ with

$$F(t, \omega) = \hat{F}(t, \omega), \quad (t, \omega) \in \llbracket t_0, T \rrbracket.$$

In this case, we define, for $(t, \omega) \in \llbracket t_0, T \rrbracket$,

$$\dot{\partial}F(t, \omega) := \dot{\partial}\hat{F}(t, \omega), \quad \nabla F(t, \omega) := \nabla \hat{F}(t, \omega), \quad \nabla^2 F(t, \omega) := \nabla^2 \hat{F}(t, \omega).$$

By [3, Lemma 2.1], the derivatives $\dot{\partial}F, \nabla F, \nabla^2 F$ are well-defined for $F \in C^{1,2}(\llbracket t_0, T \rrbracket; \mathbb{R})$.

Finally, the set $C_{pol}^{1,2}(\llbracket t_0, T \rrbracket; \mathbb{R})$ consists of all $F \in C^{1,2}(\llbracket t_0, T \rrbracket; \mathbb{R})$ such that there exist constants $C, q \geq 0$ with

$$|\dot{\partial}F(t, \omega)| + \|\nabla F(t, \omega)\| + \operatorname{tr}\left[\nabla^2 F(t, \omega)\right] \leq C\left(1 + \sup_{r \in [t_0, T]} \|\omega(r \wedge t)\|^q\right)$$

for all $(t, \omega) \in \llbracket t_0, T \rrbracket$.

**Remark 4.1.** Every map $F \colon [0, T] \times D(\mathbb{R}_+; \mathbb{R}^d)$ that is continuous with respect to $d$ is *non-anticipative*, i.e., $F(t, \omega) = F(t, \omega(\cdot \wedge t))$ for all $(t, \omega) \in [0, T] \times D(\mathbb{R}_+; \mathbb{R}^d)$.

## 4.2 The nonlinear Kolmogorov equation

For $(t, \omega, \phi) \in [\![0, T]\!] \times C^{1,2}([\![0,T]\!]; \mathbb{R})$, we set

$$G(t, \omega, \phi) := \sup \left\{ \langle \nabla \phi(t, \omega), b(f, t, \omega) \rangle + \tfrac{1}{2} \operatorname{tr} \left[ \nabla^2 \phi(t, \omega) a(f, t, \omega) \right] : f \in F \right\}.$$

One of our goals is the identification of the value function $v$ as a so-called *viscosity solution* to the nonlinear PPDE

$$\begin{cases} \dot{\partial} v(t, \omega) + G(t, \omega, v) = 0, & \text{for } (t, \omega) \in [\![0, T[\![, \\ v(T, \omega) = \psi(\omega), & \text{for } \omega \in \Omega, \end{cases} \tag{4.2}$$

where $\psi \colon \Omega \to \mathbb{R}$ is a bounded continuous function such that $\psi(\omega) = \psi(\omega(\cdot \wedge T))$. For notational convenience, we fix the terminal function $\psi$ from now on.

In contrast to the classical case, where the solution runs over time and space, we consider viscosity solutions which run over time and path. Let us provide a precise definition of a viscosity solution in our setting.

**Definition 4.2** (Viscosity Solution over Time and Path). *A function $u \colon [\![0, T]\!] \to \mathbb{R}$ is said to be a* weak sense viscosity subsolution *to (4.2) if the following two properties hold:*

  (a) $u(T, \cdot) \leq \psi$;
  (b) *for any $(t, \omega) \in [\![0, T[\![$ and $\phi \in C_{pol}^{1,2}([\![t, T]\!]; \mathbb{R})$ satisfying*

$$0 = (u - \phi)(t, \omega) = \sup\{(u - \phi)(s, \omega') \colon (s, \omega') \in [\![t, T]\!]\},$$

  *we have $\dot{\partial} \phi(t, \omega) + G(t, \omega, \phi) \geq 0$.*

*Moreover, a function $u \colon [\![0, T]\!] \to \mathbb{R}$ is said to be a* weak sense viscosity supersolution *to (4.2) if the following two properties hold:*

  (a) $u(T, \cdot) \geq \psi$;
  (b) *for any $(t, \omega) \in [\![0, T[\![$ and $\phi \in C_{pol}^{1,2}([\![t, T]\!]; \mathbb{R})$ satisfying*

$$0 = (u - \phi)(t, \omega) = \inf\{(u - \phi)(s, \omega') \colon (s, \omega') \in [\![t, T]\!]\},$$

  *we have $\dot{\partial} \phi(t, \omega) + G(t, \omega, \phi) \leq 0$.*

*Further, $u$ is called* weak sense viscosity solution *if it is a weak sense viscosity sub- and supersolution. Finally, a continuous weak sense viscosity solution is called* viscosity solution.

## 4.3 Main results

Before we present our main results, we need a last bit of notation. For $\omega, \omega' \in \Omega$ and $t \in \mathbb{R}_+$, we define the concatenation

$$\omega \stackrel{\sim}{\otimes}_t \omega' := \omega \mathbb{1}_{[0,t)} + (\omega(t) + \omega'(\cdot - t) - \omega'(0)) \mathbb{1}_{[t,\infty)},$$

and the set

$$\mathcal{R}(t, \omega) := \Big\{ P \in \mathfrak{P}_{\text{sem}}^{\text{ac}} \colon P \circ X_0^{-1} = \delta_{\omega(t)},$$
$$(\lambda \otimes P)\text{-a.e. } (dB^P/d\lambda, dC^P/d\lambda) \in \Theta(\cdot + t, \omega \stackrel{\sim}{\otimes}_t X) \Big\}. \tag{4.3}$$

We are in the position to present the main results of this section. Let us start with the viscosity solution part.

**Theorem 4.3.** *Assume that $F$ is a compact metrizable space and that the Conditions 2.4, 2.6 and 2.7 hold. Then, the value function $v$ is a weak sense viscosity solution to (4.2). If the correspondence $(t,\omega) \mapsto \mathcal{R}(t,\omega)$ is continuous (lower, upper hemicontinuous),[1] then $v$ is continuous (lower, upper semicontinuous).*

For a first result regarding upper semicontinuity of $v$ in a Markovian framework beyond the Lévy case (albeit under uniform boundedness and global Lipschitz assumptions) we refer to [17, Lemma 4.42]. The thesis [17] contains no explicit conditions for lower semicontinuity that appears to be related to martingale problems with possibly non-regular coefficients, which are difficult to study, see also [17, Remark 4.43], [26, Remark 3.4] and [27, Remark 5.4] for comments in this direction.

As shown in Theorem 4.3, (lower, upper) hemicontinuity of $(t,\omega) \mapsto \mathcal{R}(t,\omega)$ provides (lower, upper) semicontinuity of the value function $(t,\omega) \mapsto v(t,\omega)$. In the framework of nonlinear Lévy processes from [32], reduced to our path-continuous setting, continuity of $\mathcal{R}$ is rather straightforward to verify. That is, in case $\Theta \subset \mathbb{R}^d \times \mathbb{S}_+^d$ is a convex and compact set, the correspondence

$$\mathcal{R}(t,\omega) = \left\{ P \in \mathfrak{P}_{\text{sem}}^{\text{ac}} \colon P \circ X_0^{-1} = \delta_{\omega(t)}, \ (\lambda \otimes P)\text{-a.e. } (dB^P/d\lambda, dC^P/d\lambda) \in \Theta \right\},$$

is continuous. We present the details in Appendix A.

In the following, we present more general conditions for upper and lower hemicontinuity of the correspondence $(t,\omega) \mapsto \mathcal{R}(t,\omega)$ which lead to explicit conditions for the continuity of the value function and thereby identify it as a viscosity solution to the nonlinear PPDE (4.2).

**Theorem 4.4** (Upper hemicontinuity of $\mathcal{R}$)**.** *Assume that $F$ is a compact metrizable space and that the Conditions 2.4, 2.6 and 2.7 hold. Then, the correspondence $(t,\omega) \mapsto \mathcal{R}(t,\omega)$ is upper hemicontinuous with compact values.*

The following example shows that the conditions from Theorem 4.4 are not sufficient for lower hemicontinuity of $(t,\omega) \mapsto \mathcal{R}(t,\omega)$.

**Example 4.5.** Suppose that $d = 1, a \equiv 0$ and that $b(f,t,\omega) \equiv b^\circ(\omega(t))$ for all $(f,t,\omega) \in F \times [\![0,\infty[\![$, where $b^\circ \colon \mathbb{R} \to \mathbb{R}$ is any bounded continuous function such that $b^\circ(x) = \text{sgn}(x)\sqrt{|x|}$ for $|x| \leq 1$ and which is continuously differentiable off $(-1,1)$. According to [41, Exercise 12.4.2], whenever $\omega(t) \neq 0$, the set $\mathcal{R}(t,\omega)$ is a singleton $\{P_{\omega(t)}\}$ and $\mathbb{R}\backslash\{0\} \ni x \mapsto P_x$ has a weak limit through (strictly) positive or (strictly) negative values (that converge to zero) but these limits are different. To see that $\mathcal{R}$ is not lower hemicontinuous, take $(t^n, \omega^n) := (1/n, \text{id}) \in [\![0,\infty[\![$, with $n \in \mathbb{N}$, and $P := \lim_{x \nearrow 0} P_x \in \mathcal{R}(0, \text{id})$. Then, as $\lim_{n\to\infty} P_{\omega^n(1/n)} = \lim_{n\to\infty} P_{1/n} \neq P$, we get from [1, Theorem 17.21] that $(t,\omega) \mapsto \mathcal{R}(t,\omega)$ is not lower hemicontinuous.

Next, we show that the correspondence $(t,\omega) \mapsto \mathcal{R}(t,\omega)$ is lower hemicontinuous in case we impose an additional local Lipschitz condition. For a matrix $A$, we denote its transposed by $A^*$ and we denote by $\|A\|_o$ the operator norm of $A$. Let $\mathscr{P}$ be the predictable $\sigma$-field on $[\![0,\infty[\![$.

**Condition 4.6** (Local Lipschitz continuity in the last variable uniformly in the first two)**.** *There exists a dimension $r \in \mathbb{N}$ and a $\mathcal{B}(F) \otimes \mathscr{P}$-measurable function $\sigma \colon F \times [\![0,\infty[\![ \to \mathbb{R}^{d \times r}$ such that $a = \sigma\sigma^*$. Furthermore, for every $T, M > 0$, there exists a constant $C = C_{T,M} > 0$ such that*

$$\|b(f,t,\omega) - b(f,t,\alpha)\| + \|\sigma(f,t,\omega) - \sigma(f,t,\alpha)\|_o \leq C \sup_{s \in [0,t]} \|\omega(s) - \alpha(s)\|$$

*for all $\omega, \alpha \in \Omega \colon \sup_{s \in [0,t]} \|\omega(s)\| \vee \|\alpha(s)\| \leq M$ and $(f,t) \in F \times [0,T]$.*

---

[1]see [1, Definition 17.2].

The following theorem is seemingly the first result regarding lower hemicontinuity and, together with Theorem 4.3, the first result on lower semicontinuity of the value function in a path-dependent framework related to nonlinear stochastic processes.

**Theorem 4.7** (Lower hemicontinuity of $\mathcal{R}$). *Assume that $F$ is a compact metrizable space and that the Conditions 2.4, 2.6, 2.7 and 4.6 hold. Then, the correspondence $(t, \omega) \mapsto \mathcal{R}(t, \omega)$ is lower hemicontinuous.*

**Example 4.8.** An interesting situation is the case where $d = 1$ and

$$\Theta(t, \omega) = [\underline{b}_t(\omega), \overline{b}_t(\omega)] \times [\underline{a}_t(\omega), \overline{a}_t(\omega)]$$

with predictable functions

$$\underline{b}, \overline{b} \colon [\![0, \infty[\![ \to \mathbb{R}, \quad \underline{a}, \overline{a} \colon [\![0, \infty[\![ \to \mathbb{R}_+$$

such that

$$\underline{b}_t(\omega) \leq \overline{b}_t(\omega), \quad \underline{a}_t(\omega) \leq \overline{a}_t(\omega), \quad (t, \omega) \in [\![0, \infty[\![.$$

This situation is included in our setting. For instance, take $F := [0, 1] \times [0, 1]$ and, for $((f_1, f_2), t, \omega) \in F \times [\![0, \infty[\![$,

$$b((f_1, f_2), t, \omega) := \underline{b}_t(\omega) + f_1 \cdot (\overline{b}_t(\omega) - \underline{b}_t(\omega)),$$
$$a((f_1, f_2), t, \omega) := \underline{a}_t(\omega) + f_2 \cdot (\overline{a}_t(\omega) - \underline{a}_t(\omega)).$$

In the following, we consider these choices of $b$ and $a$. Evidently, the convexity assumption given by Condition 2.7 is satisfied. Further, the linear growth Condition 2.4 holds once the functions $\underline{b}, \overline{b}, \underline{a}$ and $\overline{a}$ satisfy itself linear growth conditions, i.e., in case for every $T > 0$ there exists a constant $C = C_T > 0$ such that

$$|\underline{b}_t(\omega)|^2 + |\overline{b}_t(\omega)|^2 + |\underline{a}_t(\omega)| + |\overline{a}_t(\omega)| \leq C\Big(1 + \sup_{s \in [0, t]} |\omega(s)|^2\Big)$$

for all $(t, \omega) \in [\![0, T]\!]$. Similarly, the continuity Condition 2.6 is implied by (joint) continuity of $\underline{b}, \overline{b}, \underline{a}$ and $\overline{a}$. Under these conditions, i.e., continuity and linear growth, Theorem 4.4 implies that the correspondence $(t, \omega) \mapsto \mathcal{R}(t, \omega)$ is upper hemicontinuous with compact values, while Theorem 4.3 implies that the value function $v$ is upper semicontinuous. Moreover, if $b$ and $a$ satisfy the local Lipschitz condition given by Condition 4.6, then Theorem 4.7 shows that the correspondence $(t, \omega) \mapsto \mathcal{R}(t, \omega)$ is lower hemicontinuous. In particular, the value function $v$ is then continuous. For instance, the local Lipschitz conditions hold in case for every $T, M > 0$ there exists a constant $C = C_{T,M} > 0$ such that $\underline{a}_t(\omega) \geq 1/C$ and

$$|\underline{b}_t(\omega) - \underline{b}_t(\alpha)| + |\overline{b}_t(\omega) - \overline{b}_t(\alpha)| \leq C \sup_{s \in [0, t]} |\omega(s) - \alpha(s)|,$$
$$|\underline{a}_t(\omega) - \underline{a}_t(\alpha)| + |\overline{a}_t(\omega) - \overline{a}_t(\alpha)| \leq C \sup_{s \in [0, t]} |\omega(s) - \alpha(s)|,$$

for all $\omega, \alpha \in \Omega \colon \sup_{s \in [0, t]} |\omega(s)| \vee |\alpha(s)| \leq M$ and $t \in [0, T]$.

**Remark 4.9.** Let us also comment on a result from the paper [33] that deals with a regularity property of a value function $V \colon [\![0, T]\!] \to \mathbb{R}$ that is closely related to ours. Let $\mathbf{D} \colon [\![0, T]\!] \twoheadrightarrow \mathbb{S}_+^d$ be a progressively measurable closed-valued correspondence and define, for $\delta > 0$ and $D \subseteq \mathbb{S}_+^d$, $\mathrm{Int}^\delta D := \{x \in D : B_\delta(x) \subseteq D\}$, where $B_\delta(x)$ denotes the ball around $x$ with radius $\delta$. Further, for $(t, \omega) \in [\![0, T]\!]$, let $\mathcal{M}(t, \omega)$ be the set of all $P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}(t)$ such that $P(X^t = \omega^t) = 1$, $(\lambda \otimes P)$-a.e. $b_{\cdot+t}^P = 0$, i.e., that are local martingale measures

after time $t$, and for that there exists a $\delta = \delta(t, \omega, P) > 0$ such that $(\lambda \otimes P)$-a.e. $a^P_{\cdot + t} \in \mathbb{S}^d_{++}$ and $a^P_{\cdot + t} \in \mathrm{Int}^\delta \mathbf{D}(\cdot + t, X)$. The value function $V$ from [33] is given by

$$V(t, \omega) = \sup_{P \in \mathcal{M}(t, \omega)} E^P[\xi],$$

where $\xi \colon [\![0, T]\!] \to \mathbb{R}$ is a bounded uniformly continuous input function. Our setting covers the (in a certain sense) limiting case where $\delta(t, \omega, P) = 0$, cf. [33, Remark 3.7]. Corollary 3.6 from [33] shows that, for fixed $t \in [0, T]$, the map $\omega \mapsto V(t, \omega)$ is lower semicontinuous in case $\mathbf{D}$ satisfies the uniform continuity property as defined in [33, Definition 3.2]. To compare this regularity result to our main theorems, we notice that the special case $\mathbf{D}(t, \omega) = [\underline{a}_t(\omega), \overline{a}_t(\omega)]$ satisfies the hypothesis from [33, Corollary 3.6] when the functions $\underline{a}$ and $\overline{a}$ are uniformly continuous in the sense that for all $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\omega, \alpha \in \Omega, \ \sup_{s \in [0, T]} |\omega(s) - \alpha(s)| < \delta \implies \sup_{s \in [0, T]} |a_s(\omega) - a_s(\alpha)| < \varepsilon, \quad a = \underline{a}, \overline{a}.$$

In our setting, mere continuity of $\underline{a}$ and $\overline{a}$ imply upper semicontinuity of our value function, see Example 4.8, while we require some local Lipschitz continuity for lower semicontinuity.

In the following corollary we summarize our main observations concerning the correspondence $\mathcal{R}$ and the value function $v$.

**Corollary 4.10** (Continuity of $\mathcal{R}$ and the value function)**.** *Assume that $F$ is a compact metrizable space and that the Conditions 2.4, 2.6, 2.7 and 4.6 hold. Then, the correspondence $(t, \omega) \mapsto \mathcal{R}(t, \omega)$ is continuous and the value function $[\![0, T]\!] \ni (t, \omega) \mapsto v(t, \omega)$ is a viscosity solution to (4.2).*

It is a natural question whether the value function is the unique viscosity solution to the nonlinear PPDE (4.2). Uniqueness results for certain Hamilton–Jacobi–Bellman PPDEs were recently proved in [4, 43]. In the following we apply a theorem from [43] and show that, under global Lipschitz conditions, our value function is the unique viscosity solution among all solutions that satisfy a certain Lipschitz property that we explain now. We define $\mathsf{d} = \mathsf{d}_T \colon [\![0, T]\!] \times [\![0, T]\!] \to \mathbb{R}_+$ by

$$\mathsf{d}((t, \omega), (s, \alpha)) := \left(1 + \sup_{r \in [0, t]} \|\omega(r)\| + \sup_{r \in [0, s]} \|\alpha(r)\|\right) |t - s|^{1/2} \tag{4.4}$$
$$+ \sup_{r \in [0, T]} \|\omega(r \wedge t) - \alpha(r \wedge s)\|.$$

A function $f \colon [\![0, T]\!] \to \mathbb{R}$ is said to be $\mathsf{d}$-*Lipschitz continuous*, if there exists a constant $L = L_T > 0$ such that

$$|f(t, \omega) - f(s, \alpha)| \leq L \mathsf{d}((t, \omega), (s, \alpha))$$

for all $(t, \omega), (s, \alpha) \in [\![0, T]\!]$. We remark that $\mathsf{d}$-Lipschitz continuity entails continuity on $[\![0, T]\!]$.

**Condition 4.11** (Global Lipschitz continuity)**.** *There exists a dimension $r \in \mathbb{N}$ and a $\mathcal{B}(F) \otimes \mathscr{P}$-measurable function $\sigma \colon F \times [\![0, \infty[\![ \to \mathbb{R}^{d \times r}$ such that $a = \sigma \sigma^*$. Moreover, there exists a constant $C = C_T > 0$ such that*

$$\|b(f, t, \omega) - b(f, t, \alpha)\| + \|\sigma(f, t, \omega) - \sigma(f, t, \alpha)\|_o \leq C \sup_{s \in [0, t]} \|\omega(s) - \alpha(s)\|$$

*for all $\omega, \alpha \in \Omega$ and $(f, t) \in F \times [0, T]$.*

**Theorem 4.12.** *Assume that $F$ is a compact metrizable space and that the Conditions 2.4, 2.5, 2.7 and 4.11 hold. Furthermore, suppose that the terminal function $\psi$ is bounded and Lipschitz continuous, i.e., there exists a constant $L > 0$ such that*

$$|\psi(\omega) - \psi(\alpha)| \leq L \sup_{s \in [0,T]} \|\omega(s) - \alpha(s)\|$$

*for all $\omega, \alpha \in \Omega$. Then, the value function is $d$-Lipschitz continuous.*

Estimates similar to Theorem 4.12 are standard in stochastic optimal control, see, e.g., [43, Theorem 2.11] where a version of Theorem 4.12 for a path-dependent control framework is given.

When it comes to mere continuity, Theorem 4.12 requires stronger regularity assumption in the path variable than Corollary 4.10. In particular, also the input function $\psi$ has to be Lipschitz continuous in Theorem 4.12. Contrary to Corollary 4.10, Theorem 4.12 shows that the value function is Lipschitz continuous in space for fixed, but arbitrary, times with a uniform (in time) Lipschitz constant. We think this observation is of independent interest.

We are in the position to present a uniqueness result for the value function.

**Theorem 4.13.** *Assume that $F$ is a compact metrizable space and that the Conditions 2.4, 2.6, 2.7 and 4.11 hold. Furthermore, suppose that the terminal function $\psi$ is bounded and Lipschitz continuous, i.e., there exists a constant $L > 0$ such that*

$$|\psi(\omega) - \psi(\alpha)| \leq L \sup_{s \in [0,T]} \|\omega(s) - \alpha(s)\|$$

*for all $\omega, \alpha \in \Omega$. Then, the value function $v$ is the unique viscosity solution that is bounded and $d$-Lipschitz continuous.*

*Proof.* Thanks to the Theorems 4.3 and 4.12, the value function $v$ is a viscosity solution to (4.2) that is $d$-Lipschitz continuous. Hence, the uniqueness statement follows from [43, Theorem 6.2]. $\square$

**Remark 4.14.** Under more assumptions on the coefficients, our value function can be identified as the unique viscosity solution in a class of uniformly continuous functions. This follows from a uniqueness result that was established in the recent paper [4].

**Remark 4.15.**

(i) In the Markovian case where $b(f, t, \omega)$ and $a(f, t, \omega)$ depend on $(t, \omega)$ only through the value $\omega(t)$, assumptions in the spirit of those from Theorem 4.13 imply that the (point-dependent) value function is unique among all bounded viscosity solutions, see [5] for a precise result in this direction.

(ii) Let us again comment on the situation from Example 4.8, i.e., the case where $d = 1$ and

$$\Theta(t, \omega) = [\underline{b}_t(\omega), \overline{b}_t(\omega)] \times [\underline{a}_t(\omega), \overline{a}_t(\omega)]$$

with predictable functions

$$\underline{b}, \overline{b} \colon [\![0, \infty[\![ \to \mathbb{R}, \quad \underline{a}, \overline{a} \colon [\![0, \infty[\![ \to \mathbb{R}_+$$

such that

$$\underline{b}_t(\omega) \leq \overline{b}_t(\omega), \quad \underline{a}_t(\omega) \leq \overline{a}_t(\omega), \quad (t, \omega) \in [\![0, \infty[\![.$$

As explained in Example 4.8, this situation can be included in our framework with $F := [0, 1] \times [0, 1]$ and

$$b((f_1, f_2), t, \omega) := \underline{b}_t(\omega) + f_1 \cdot (\overline{b}_t(\omega) - \underline{b}_t(\omega)),$$

$$a((f_1, f_2), t, \omega) := \underline{a}_t(\omega) + f_2 \cdot (\overline{a}_t(\omega) - \underline{a}_t(\omega)),$$

for $(f_1, f_2) \in F$ and $(t, \omega) \in [\![0, \infty[\![$. In this setting, Condition 4.11 holds in case there exists a constant $C > 0$ such that

$$|\underline{b}(t, \omega)|, |\overline{b}(t, \omega)| \leq C, \qquad \frac{1}{C} \leq \underline{a}(t, \omega), \overline{a}(t, \omega) \leq C,$$

and

$$|\underline{b}(t, \omega) - \underline{b}(t, \alpha)| + |\overline{b}(t, \omega) - \overline{b}(t, \alpha)| \leq C \sup_{s \in [0,t]} |\omega(s) - \alpha(s)|,$$

$$|\underline{a}(t, \omega) - \underline{a}(t, \alpha)| + |\overline{a}(t, \omega) - \overline{a}(t, \alpha)| \leq C \sup_{s \in [0,t]} |\omega(s) - \alpha(s)|,$$

for all $t \in [0, T]$ and $\omega, \alpha \in \Omega$.

## 5 The martingale problem

As a final main result, we show that $\mathcal{E}$ solves a type of nonlinear martingale problem. This result supports our interpretation of $\mathcal{E}$ as a nonlinear continuous semimartingale. We restrict our attention to the one-dimensional situation, i.e., we presume that $d = 1$. Let $\mathbb{M}_{icx}$ be the set of all $\phi \in C^2(\mathbb{R}; \mathbb{R})$ such that $\phi', \phi'' \geq 0$. Furthermore, for $n > 0$, we set

$$\rho_n := \inf\{t \geq 0 \colon |X_t| \geq n\}.$$

**Definition 5.1.** *We say that $\mathcal{E}$ solves the martingale problem associated to $G$ if for all $n \in \mathbb{N}$ and $\phi \in \mathbb{M}_{icx}$ the process*

$$\phi(X_{t \wedge \rho_n}) - \int_0^{t \wedge \rho_n} G(r, X, \phi) dr, \quad t \in \mathbb{R}_+,$$

*is an $\mathcal{E}$-martingale, i.e.,*

$$\mathcal{E}_s\left(\phi(X_{t \wedge \rho_n}) - \phi(X_{s \wedge \rho_n}) - \int_{s \wedge \rho_n}^{t \wedge \rho_n} G(r, X, \phi) dr\right)(\omega) = 0, \quad s \leq t, \ \omega \in \Omega.$$

Of course, the acronym $icx$ stands for increasing and convex. Although this set of test functions looks non-standard at first sight, it is perfectly fine in the linear setting, because it is well-known to be measure-determining for the set of Borel probability measures on $\mathbb{R}$ with finite first moment (cf. [30, Theorems 5.2, 5.3] and [8, Theorem 2.1]).

**Theorem 5.2.** *Suppose that $F = F_0 \times F_1$ for two compact metrizable spaces $F_0$ and $F_1$, that $b(f, t, \omega) = b(f_0, t, \omega)$ and $a(f, t, \omega) = a(f_1, t, \omega)$ for all $(f, t, \omega) = (f_0, f_1, t, \omega) \in F \times [\![0, \infty[\![$, and that the Conditions 2.4 and 2.6 hold. Then, $\mathcal{E}$ solves the martingale problem associated to $G$.*

In [15] a related result for generalized $G$-Brownian motion was proved. Let us discuss the relation of Theorem 5.2 and the approach from [15] in more detail. In [15] it is shown that a generalized $G$-Brownian motion (see [37]) solves a nonlinear martingale problem that is defined not only via test functions but also via a class of nonlinear test generators. This approach uses the power of $G$-Itô calculus in a crucial manner. In our framework we cannot rely on such a stochastic calculus. Instead, we only identify a suitable class of test functions, namely $\mathbb{M}_{icx}$, for that we can prove a nonlinear martingale property. In that sense our treatment of the martingale problem is certainly less complete than those for the generalized $G$-Brownian motion from [15]. We think that Theorem 5.2 is a good indicator for our interpretation of $\mathcal{E}$ as a nonlinear continuous semimartingale.

## 6 Proof of the dynamic programming principle: Theorem 3.1

The proof is based on an application of [11, Theorem 2.1], which provides three abstract conditions on the correspondence $\mathcal{C}$ which imply the DPP. The work lies in the verification of these conditions. For reader's convenience, let us restate them. For a probability measure $P$ on $(\Omega, \mathcal{F})$, a kernel $\Omega \ni \omega \mapsto Q_\omega \in \mathfrak{P}(\Omega)$, and a finite stopping time $\tau$, we define the pasting measure

$$(P \otimes_\tau Q)(A) := \iint \mathbb{1}_A(\omega \otimes_{\tau(\omega)} \omega') Q_\omega(d\omega') P(d\omega), \quad A \in \mathcal{F}.$$

We are in the position to formulate the assumptions from [11, Theorem 2.1].

(i) *Measurable graph condition:* The set

$$\left\{ (t, \omega, P) \in [\![0, \infty[\![ \times \mathfrak{P}(\Omega) \colon P \in \mathcal{C}(t, \omega) \right\}$$

is analytic.

(ii) *Stability under conditioning:* For any $t \in \mathbb{R}_+$, any stopping time $\tau$ with $t \leq \tau < \infty$, and any $P \in \mathcal{C}(t, \alpha)$ there exists a family $\{P(\cdot|\mathcal{F}_\tau)(\omega) \colon \omega \in \Omega\}$ of regular $P$-conditional probabilities given $\mathcal{F}_\tau$ such that $P$-a.s. $P(\cdot|\mathcal{F}_\tau) \in \mathcal{C}(\tau, X)$.

(iii) *Stability under pasting:* For any $\alpha \in \Omega$, $t \in \mathbb{R}_+$, any stopping time $\tau$ with $t \leq \tau < \infty$, any $P \in \mathcal{C}(t, \alpha)$ and any $\mathcal{F}_\tau$-measurable map $\Omega \ni \omega \mapsto Q_\omega \in \mathfrak{P}(\Omega)$ the following implication holds:

$$P\text{-a.s. } Q \in \mathcal{C}(\tau, X) \quad \Longrightarrow \quad P \otimes_\tau Q \in \mathcal{C}(t, \alpha).$$

**Remark 6.1.** To verify (ii) above, it is necessary to introduce the sets $\mathfrak{P}_{\text{sem}}^{\text{ac}}(t)$ for $t > 0$. Notice that for a finite stopping time $\tau$ it is not necessarily true that $P \in \mathfrak{P}_{\text{sem}}^{\text{ac}}$ implies $\{P(\cdot|\mathcal{F}_\tau)(\omega) \colon \omega \in \Omega\} \subset \mathfrak{P}_{\text{sem}}^{\text{ac}}$. Indeed, we only have the semimartingale property of $X$ after $\tau(\omega)$ under $P(\cdot|\mathcal{F}_\tau)(\omega)$. This forces us to study the correspondence $t \mapsto \mathfrak{P}_{\text{sem}}^{\text{ac}}(t)$. To see the issue, let $P$ be the Wiener measure, i.e., the law of a one-dimensional Brownian motion. Then, for any $t > 0$ and $P$-a.a. $\omega \in \Omega$, $P(\cdot|\mathcal{F}_t)(\omega) \notin \mathfrak{P}_{\text{sem}}^{\text{ac}}$. This follows simply from the fact that $P$-a.a. $\omega \in \Omega$ are locally of infinite variation, which implies that the stopped process $X_{\cdot \wedge t}$, which coincides $P(\cdot|\mathcal{F}_t)(\omega)$-a.s. with $\omega(\cdot \wedge t)$, is no $P(\cdot|\mathcal{F}_t)(\omega)$-semimartingale for $P$-a.a. $\omega \in \Omega$ (see [23, Proposition I.4.28]).

In the following three sections we check these properties. In the fourth (and last) section, we finalize the proof of Theorem 3.1.

Before we start our program let us shortly comment on our strategy of proof and relate it to existing literature. Dynamic programming principles for nonlinear expectations related to nonlinear Lévy processes, and more general Markovian semimartingales, have been proved in [17, 32]. In contrast to our approach, which has also been used in [11] for controlled martingale problems, the proofs in [17, 32] are based on an abstract result from [34]. In general, the methodologies are different in the sense that the uncertainty sets of measures in [17, 32] consist of laws of semimartingales that start at zero, while time and path dependence persists in both the set $\Theta$ and the test function $\psi$. In our setting, however, the dependence on $(t, \omega)$ only enters via the set $\mathcal{C}(t, \omega)$. This construction seems to us closer to those of a classical linear conditional expectation.

Compared to [17, 32], the main difference in our setting is the time-dependence $t \mapsto \mathfrak{P}_{\text{sem}}^{\text{ac}}(t)$ for which we have to prove new measurability and stability properties. While this differs from [17, 32] on a technical level, we closely follow, however, their general strategy of proof.

## 6.1 Measurable graph condition

The proof of the measurable graph condition is split into several parts. For $t \in \mathbb{R}_+$, we define the usual shift $\theta_t \colon \Omega \to \Omega$ by $\theta_t(\omega) = \omega(\cdot + t)$ for all $\omega \in \Omega$. The next two lemmata give some rather elementary observations.

**Lemma 6.2.** *The map $(t, P) \mapsto P \circ \theta_t^{-1} =: P_t$ is continuous.*

*Proof.* This follows from the continuity of $(t, \omega) \mapsto \theta_t(\omega)$ and [2, Theorem 8.10.61]. $\square$

**Lemma 6.3.** *Let $Y = (Y_t)_{t \geq 0}$ be an $\mathbb{R}^d$-valued continuous process and set $\mathcal{F}_t^Y := \sigma(Y_s, s \leq t)$ for $t \in \mathbb{R}_+$. Then, $Y^{-1}(\mathcal{F}_{s+}) = \mathcal{F}_{s+}^Y$ for all $s \in \mathbb{R}_+$.*

*Proof.* First of all, it is clear that $Y^{-1}(\mathcal{F}_s) = \mathcal{F}_s^Y$ for all $s \in \mathbb{R}_+$. Since $\mathcal{F}_{s+} \subset \mathcal{F}_{s+\varepsilon}$ for all $\varepsilon > 0$, we obtain $Y^{-1}(\mathcal{F}_{s+}) \subset \bigcap_{\varepsilon > 0} Y^{-1}(\mathcal{F}_{s+\varepsilon}) = \mathcal{F}_{s+}^Y$. Conversely, take $A \in \mathcal{F}_{s+}^Y$. By definition, $A \in Y^{-1}(\mathcal{F}_{s+1/n})$ for every $n \in \mathbb{N}$. Hence, for each $n \in \mathbb{N}$ there exists a set $B_n \in \mathcal{F}_{s+1/n}$ such that $A = Y^{-1}(B_n)$. Define $G := \bigcap_{n \in \mathbb{N}} \bigcup_{m \geq n} B_m$ and notice that $G \in \mathcal{F}_{s+}$. Finally, as $Y^{-1}(G) = \bigcap_{n \in \mathbb{N}} \bigcup_{m \geq n} Y^{-1}(B_m) = A$, we conclude that $A \in Y^{-1}(\mathcal{F}_{s+})$. The proof is complete. $\square$

The following lemma is a restatement of [21, Lemma 2.9 a)] for a path-continuous setting.

**Lemma 6.4.** *Let $\mathbf{B}^* = (\Omega^*, \mathcal{F}^*, (\mathcal{F}_t^*)_{t \geq 0}, P^*)$ and $\mathbf{B}' = (\Omega', \mathcal{F}', (\mathcal{F}_t')_{t \geq 0}, P')$ be two filtered probability spaces with right-continuous filtrations and the property that there is a map $\phi \colon \Omega' \to \Omega^*$ such that $\phi^{-1}(\mathcal{F}^*) \subset \mathcal{F}', P^* = P' \circ \phi^{-1}$ and $\phi^{-1}(\mathcal{F}_t^*) = \mathcal{F}_t'$ for all $t \in \mathbb{R}_+$. Then, $X^*$ is a $d$-dimensional continuous semimartingale (local martingale) on $\mathbf{B}^*$ if and only if $X' = X^* \circ \phi$ is a $d$-dimensional continuous semimartingale (local martingale) on $\mathbf{B}'$. Moreover, $(B^*, C^*)$ are the characteristics of $X^*$ if and only if $(B^* \circ \phi, C^* \circ \phi)$ are the characteristics of $X' = X^* \circ \phi$.*

**Lemma 6.5.** *We have*

$$\big\{(t, \omega, P) \colon P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}(t), (\lambda \otimes P)\text{-}a.e. \ (dB_{\cdot+t}^P/d\lambda, dC_{\cdot+t}^P/d\lambda) \in \Theta(\cdot + t, \omega \otimes_t X)\big\}$$
$$= \big\{(t, \omega, P) \colon P_t \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}, (\lambda \otimes P_t)\text{-}a.e. \ (dB^{P_t}/d\lambda, dC^{P_t}/d\lambda) \in \Theta(\cdot + t, \omega \widetilde{\otimes}_t X)\big\}.$$

*Proof.* Since $(\omega \widetilde{\otimes}_t X) \circ \theta_t = \omega \otimes_t X$, the claim follows from Lemmata 6.3 and 6.4. $\square$

Finally, we are in the position to prove the measurable graph condition.

**Lemma 6.6.** *The set $\{(t, \omega, P) \in [\![0, \infty[\![ \times \mathfrak{P}(\Omega) \colon P \in \mathcal{C}(t, \omega)\}$ is Borel.*

*Proof.* Notice that $\{(t, \omega, P) \in [\![0, \infty[\![ \times \mathfrak{P}(\Omega) \colon P(X^t = \omega^t) = 1\}$ is Borel by [2, Theorem 8.10.61]. Further, notice that

$$\big\{(t, \omega, P) \colon P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}, (\lambda \otimes P)\text{-}a.e. \ (dB^P/d\lambda, dC^P/d\lambda) \in \Theta(\cdot + t, \omega \widetilde{\otimes}_t X)\big\}$$
$$= \big\{(t, \omega, P) \colon P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}, (\lambda \otimes P)((\cdot + t, \omega \widetilde{\otimes}_t X, dB^P/d\lambda, dC^P/d\lambda) \notin \mathrm{gr} \, \Theta) = 0\big\}.$$

By virtue of [31, Theorem 2.6] and [2, Theorem 8.10.61], this set is Borel. Consequently, the lemma follows from Lemmata 6.2 and 6.5. $\square$

## 6.2 Stability under conditioning

Next, we check stability under conditioning. Throughout this section, we fix $(t^*, \omega^*) \in [\![0, \infty[\![$, a stopping time $\tau$ with $t^* \leq \tau < \infty$, and a probability measure $P$ on $(\Omega, \mathcal{F})$ such that $P(X = \omega^* \text{ on } [0, t^*]) = 1$. We denote by $P(\cdot | \mathcal{F}_\tau)$ a version of the regular conditional $P$-probability given $\mathcal{F}_\tau$. The following three observations are simple but useful.

**Lemma 6.7.** *There exists a $P$-null set $N \in \mathcal{F}_\tau$ such that $P(A|\mathcal{F}_\tau)(\omega) = \mathbb{1}_A(\omega)$ for all $A \in \mathcal{F}_\tau$ and $\omega \notin N$.*

*Proof.* Notice that $\mathcal{F}_\tau = \sigma(X_{t \wedge \tau}, t \in \mathbb{Q}_+)$ is countably generated (see [41, Lemma 1.3.3]). Now, the claim is classical ([40, Theorem 9.2.1]). $\square$

**Lemma 6.8.** *If $\omega \mapsto Q_\omega$ is a kernel from $\mathcal{F}$ into $\mathcal{F}$ such that $Q_\omega(\omega = X$ on $[0, \tau(\omega)]) = 1$ for $P$-a.a. $\omega \in \Omega$, then $Q_\omega(\omega \otimes_{\tau(\omega)} X = X) = 1$ for $P$-a.a. $\omega \in \Omega$. In particular, there exists a $P$-null set $N \in \mathcal{F}_\tau$ such that $P(\omega \otimes_{\tau(\omega)} X = X|\mathcal{F}_\tau)(\omega) = 1$ for all $\omega \notin N$.*

*Proof.* The first claim is obvious and the second follows from the first and Lemma 6.7. $\square$

**Lemma 6.9.** *Take a predictable process $H = (H_t)_{t \geq 0}, t \in \mathbb{R}_+$ and $\omega \in \Omega$. Then, $H_{\cdot + t}(\omega \otimes_t X)$ is predictable for the (right-continuous) filtration generated by $X_{\cdot + t}$.*

*Proof.* This follows from [20, Proposition 10.35 (d)]. $\square$

The next lemma is a partial restatement of [41, Theorem 1.2.10].

**Lemma 6.10.** *If $M - M^{t^*}$ is a $P$-martingale, then there exists a $P$-null set $N$ such that, for all $\omega \notin N$, $M - M^{\tau(\omega)}$ is a $P(\cdot|\mathcal{F}_\tau)(\omega)$-martingale.*

The following lemma should be compared to [32, Theorem 3.1].

**Lemma 6.11.** *Suppose that $P \in \mathfrak{P}_{\mathrm{sem}}(t^*)$ and denote the $P$-semimartingale characteristics of $X_{\cdot + t^*}$ by $(B_{\cdot + t^*}, C_{\cdot + t^*})$. Then, there exists a $P$-null set $N$ such that, for all $\omega \notin N$, $X_{\cdot + \tau(\omega)}$ is a $P(\cdot|\mathcal{F}_\tau)(\omega)$-semimartingale (for its right-continuous natural filtration) and the corresponding $P(\cdot|\mathcal{F}_\tau)(\omega)$-characteristics are given by*

$$(B_{\cdot + \tau(\omega)} - B_{\tau(\omega)})(\omega \otimes_{\tau(\omega)} X), \quad (C_{\cdot + \tau(\omega)} - C_{\tau(\omega)})(\omega \otimes_{\tau(\omega)} X).$$

*Proof.* For notational convenience, we prove the claim only for $t^* = 0$. Define $M := X - B - X_0$ and

$$\tau_n := \inf\{t \geq 0 \colon \|M_t\| \geq n\}, \quad n > 0.$$

The process $M^{\tau_n}$ is clearly bounded. Thus, by definition of the first characteristic, the process $M^{\tau_n}$ is a $P$-martingale. By Lemma 6.10, there exists a $P$-null set $N$, such that, for all $\omega \notin N$, $M^{\tau_n} - M^{\tau_n \wedge \tau(\omega)}$ is a $P(\cdot|\mathcal{F}_\tau)(\omega)$-martingale. As

$$E^P\left[P\left(\liminf_{n \to \infty} \tau_n < \infty | \mathcal{F}_\tau\right)\right] = 0,$$

possibly making $N$ a bit larger, we can w.l.o.g. assume that $P(\tau_n \to \infty|\mathcal{F}_\tau)(\omega) = 1$ for all $\omega \notin N$. Now, using Lemmata 6.7 and 6.8, possibly making $N$ again a bit larger, we get $P(\cdot|\mathcal{F}_\tau)(\omega)$-a.s.

$$X^{\tau_n} - X^{\tau_n \wedge \tau(\omega)} = (M^{\tau_n} - M^{\tau_n \wedge \tau(\omega)})(\omega \otimes_{\tau(\omega)} X) + (B^{\tau_n} - B^{\tau_n \wedge \tau(\omega)}(\omega))(\omega \otimes_{\tau(\omega)} X).$$

By the tower rule, $(M^{\tau_n}_{\cdot + \tau(\omega)} - M_{\tau_n \wedge \tau(\omega)})(\omega \otimes_{\tau(\omega)} X)$ is a $P(\cdot|\mathcal{F}_\tau)(\omega)$-martingale for the filtration generated by $X_{\cdot + \tau(\omega)}$. Hence, for all $\omega \notin N$, by virtue of Lemma 6.9, we deduce that $X_{\cdot + \tau(\omega)}$ is a $P(\cdot|\mathcal{F}_\tau)(\omega)$-semimartingale (for its right-continuous canonical filtration), and further we get the formula for the first characteristic. For the second characteristic, notice that, for all $\omega \notin N$, $P(\cdot|\mathcal{F}_\tau)(\omega)$-a.s.

$$[X - X^{\tau(\omega)}, X - X^{\tau(\omega)}] = [X, X] - 2[X, X^{\tau(\omega)}] + [X^{\tau(\omega)}, X^{\tau(\omega)}]$$
$$= C - C^{\tau(\omega)}$$
$$= (C - C^{\tau(\omega)}(\omega))(\omega \otimes_{\tau(\omega)} X).$$

This observation yields the formula for the second characteristic. $\square$

We are in the position to deduce stability under conditioning. The following corollary should be compared to [32, Corollary 3.2].

**Corollary 6.12.** *If $P \in \mathcal{C}(t^*, \omega^*)$, then $P$-a.s. $P(\cdot|\mathcal{F}_\tau) \in \mathcal{C}(\tau, X)$.*

*Proof.* Let $N$ be as in Lemma 6.7 and take $\omega \notin N$. Then, $P(\tau = \tau(\omega)|\mathcal{F}_\tau)(\omega) = 1$ and

$$P(X^{\tau(\omega)} = \omega^{\tau(\omega)}|\mathcal{F}_\tau)(\omega) = P(X^\tau = \omega^{\tau(\omega)}|\mathcal{F}_\tau)(\omega) = \mathbb{1}\{X^\tau = \omega^{\tau(\omega)}\}(\omega) = 1.$$

Thanks to Lemma 6.11, for $P$-a.a. $\omega \in \Omega$, $P(\cdot|\mathcal{F}_\tau)(\omega) \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}(\tau(\omega))$ and the Lebesgue densities of the characteristics of $X_{\cdot+\tau(\omega)}$ are given by $(b_{\cdot+\tau(\omega)}, a_{\cdot+\tau(\omega)})(\omega \otimes_{\tau(\omega)} X)$, where $(b_{\cdot+t^*}, a_{\cdot+t^*})$ are the Lebesgue densities of the characteristics of $X_{\cdot+t^*}$ under $P$. Thus, we get from Lemma 6.8, Fubini's theorem, Lemma 6.7, the tower rule, and $P \in \mathcal{C}(t^*, \omega^*)$ that

$$\iint_{\tau(\omega)}^\infty E^P\big[\mathbb{1}\{(b_t, a_t)(\omega \otimes_{\tau(\omega)} X) \notin \Theta(t, \omega \otimes_{\tau(\omega)} X)\}|\mathcal{F}_\tau\big](\omega) dt P(d\omega)$$

$$= \iint_{\tau(\omega)}^\infty E^P\big[\mathbb{1}\{(b_t, a_t) \notin \Theta(t, X)\}|\mathcal{F}_\tau\big](\omega) dt P(d\omega)$$

$$= \int E^P\Big[\int_{\tau(\omega)}^\infty \mathbb{1}\{(b_t, a_t) \notin \Theta(t, X)\} dt\big|\mathcal{F}_\tau\Big](\omega) P(d\omega)$$

$$= \int E^P\Big[\int_\tau^\infty \mathbb{1}\{(b_t, a_t) \notin \Theta(t, X)\} dt\big|\mathcal{F}_\tau\Big](\omega) P(d\omega)$$

$$= E^P\Big[\int_\tau^\infty \mathbb{1}\{(b_t, a_t) \notin \Theta(t, X)\} dt\Big] = 0.$$

This observation completes the proof of $P$-a.s. $P(\cdot|\mathcal{F}_\tau) \in \mathcal{C}(\tau, X)$. $\qquad\square$

### 6.3 Stability under pasting

In this section we check stability under pasting. The proof is split into several steps. Throughout this section, we fix $(t^*, \omega^*) \in [\![0, \infty[\![$, a probability measure $P$ on $(\Omega, \mathcal{F})$ such that $P(X = \omega^*$ on $[0, t^*]) = 1$, a stopping time $\tau$ with $t^* \leq \tau < \infty$, and an $\mathcal{F}_\tau$-measurable map $\Omega \ni \omega \mapsto Q_\omega \in \mathfrak{P}(\Omega)$ such that $Q_\omega(\omega = X$ on $[0, \tau(\omega)]) = 1$ for $P$-a.a. $\omega \in \Omega$. For simplicity, we further set $\overline{P} := P \otimes_\tau Q$. The following corollary is an immediate consequence of Lemma 6.8.

**Corollary 6.13.** $\overline{P} = E^P\big[Q(\cdot)\big]$.

**Lemma 6.14.** $P = \overline{P}$ on $\mathcal{F}_\tau$ and $\overline{P}$-a.s. $\overline{P}(\cdot|\mathcal{F}_\tau) = Q$.

*Proof.* Take $A \in \mathcal{F}_\tau$. By hypothesis, $Q_\omega(\omega \otimes_{\tau(\omega)} X = \omega$ on $[0, \tau(\omega)]) = 1$ for $P$-a.a. $\omega \in \Omega$. Hence, Galmarino's test yields that

$$\overline{P}(A) = \int Q_\omega(\omega \in A) P(d\omega) = P(A).$$

This is the first claim. For the second claim, take $G \in \mathcal{F}$. Then, again by Galmarino's test, Lemma 6.8 and the first part,

$$\overline{P}(A \cap G) = \int Q_\omega(\omega \in A, \omega \otimes_{\tau(\omega)} X \in G) P(d\omega) = \int_A Q_\omega(G) P(d\omega) = E^{\overline{P}}\big[\mathbb{1}_A Q(G)\big].$$

This yields $\overline{P}$-a.s. $\overline{P}(G|\mathcal{F}_\tau) = Q(G)$. Since $\mathcal{F}_\tau$ is countably generated, a monotone class argument completes the proof. $\qquad\square$

The following two lemmata should be compared to [32, Proposition 4.2].

**Lemma 6.15.** *Let $P \in \mathfrak{P}_{\text{sem}}(t^*)$. If $P$-a.s. $Q \in \mathfrak{P}_{\text{sem}}(\tau)$, then $\overline{P} \in \mathfrak{P}_{\text{sem}}(t^*)$.*

*Proof.* Let $T > t^*$, and take a sequence $(H^n)_{n \in \mathbb{Z}_+}$ of simple predictable processes on $[\![t^*, T]\!]$ such that $H^n \to H^0$ uniformly in time and $\omega$. Then, by Lemma 6.14, we get, for every $i = 1, 2, \ldots, d$,

$$E^{\overline{P}}\left[\left|\int_{t^*}^T (H_s^n - H_s)dX_s^{(i)}\right| \wedge 1\right]$$
$$\leq E^P\left[\left|\int_{t^*}^{\tau \wedge T} (H_s^n - H_s)dX_s^{(i)}\right| \wedge 1\right]$$
$$+ \int E^{Q_\omega}\left[\left|\int_{\tau(\omega) \wedge T}^T (H_s^n - H_s)(\omega \otimes_{\tau(\omega)} X)dX_s^{(i)}\right| \wedge 1\right] P(d\omega).$$

The first term converges to zero since $P \in \mathfrak{P}_{\text{sem}}(t^*)$ and thanks to the Bichteler–Dellacherie (BD) Theorem ([39, Theorem III.43]). The second term converges to zero by dominated convergence, the assumption that $P$-a.s. $Q \in \mathfrak{P}_{\text{sem}}(\tau)$ and, by virtue of Lemma 6.9, again the BD Theorem. Consequently, invoking the BD Theorem a third time yields that $\overline{P} \in \mathfrak{P}_{\text{sem}}(t^*)$. $\square$

**Lemma 6.16.** *Let $P \in \mathfrak{P}_{\text{sem}}^{\text{ac}}(t^*)$. If $P$-a.s. $Q \in \mathfrak{P}_{\text{sem}}^{\text{ac}}(\tau)$, then $\overline{P} \in \mathfrak{P}_{\text{sem}}^{\text{ac}}(t^*)$.*

*Proof.* By Lemma 6.15, we already know that $\overline{P} \in \mathfrak{P}_{\text{sem}}(t^*)$. Denote by $B_{\cdot + t^*}$ the first characteristic of $X_{\cdot + t^*}$ under $\overline{P}$, and let

$$B_{\cdot + t^*} = \int_{t^*}^{\cdot} \phi_s ds + \psi$$

be the Lebesgue decomposition of (the paths of) $B$. Since $P \in \mathfrak{P}_{\text{sem}}^{\text{ac}}(t^*)$ and $\overline{P} = P$ on $\mathcal{F}_\tau$ by Lemma 6.14, simply by the definition of the first characteristic, we get that $B \ll \lambda$ on $[\![0, \tau]\!]$. Hence, by virtue of Corollary 6.13, it suffices to show that

$$D := \left\{ B_{\cdot + \tau} - B_\tau \neq \int_\tau^{\cdot + \tau} \phi_s ds \right\}$$

is a $Q_\omega$-null set for $\overline{P}$-a.a. $\omega \in \Omega$. Due to Lemmata 6.11 and 6.14, for $\overline{P}$-a.a. $\omega \in \Omega$, we have $Q_\omega \in \mathfrak{P}_{\text{sem}}(\tau(\omega))$ and the $Q_\omega$-characteristics of the shifted process $X_{\cdot + \tau(\omega)}$ are given by $(B_{\cdot + \tau(\omega)} - B_{\tau(\omega)}, C_{\cdot + \tau(\omega)} - C_{\tau(\omega)})(\omega \otimes_{\tau(\omega)} X)$. As $P$-a.s. $Q \in \mathfrak{P}_{\text{sem}}^{\text{ac}}(\tau)$, Lemma 6.14 and the uniqueness of the Lebesgue decomposition yield that for $\overline{P}$-a.a. $\omega \in \Omega$

$$Q_\omega\left(B_{\cdot + \tau(\omega)} - B_{\tau(\omega)} \neq \int_{\tau(\omega)}^{\cdot + \tau(\omega)} \phi_s ds\right) = 0.$$

As $Q_\omega(\tau = \tau(\omega)) = 1$ for $\overline{P}$-a.a. $\omega \in \Omega$, $\overline{P}$-a.s. $Q(D) = 0$. One may proceed similarly for the other characteristic. $\square$

**Lemma 6.17.** *Let $P \in \mathcal{C}(t^*, \omega^*)$. If $Q_\omega \in \mathcal{C}(\tau(\omega), \omega)$ for $P$-a.e. $\omega \in \Omega$, then $\overline{P} \in \mathcal{C}(t^*, \omega^*)$.*

*Proof.* Lemma 6.16 implies $\overline{P} \in \mathfrak{P}_{\text{sem}}^{\text{ac}}(t^*)$. Clearly, $\{X = \omega^* \text{ on } [0, t^*]\} \in \mathcal{F}_\tau$ as $\tau \geq t^*$. Thus, Lemma 6.14 yields $\overline{P}(X = \omega^* \text{ on } [0, t^*]) = P(X = \omega^* \text{ on } [0, t^*]) = 1$. Denote by $(b_{\cdot + t^*}, a_{\cdot + t^*})$ the Lebesgue densities of the characteristics of $X_{\cdot + t^*}$ under $\overline{P}$. Since $\overline{P} = P$ on $\mathcal{F}_\tau$, it suffices to show that

$$R := \left\{ (t, \omega) \in [\![\tau, \infty[\![ : (b_t(\omega), a_t(\omega)) \notin \Theta(t, \omega) \right\}$$

is a $(\lambda \otimes \overline{P})$-null set. By virtue of Lemmata 6.11 and 6.14, the assumption that $Q_\omega \in \mathcal{C}(\tau(\omega), \omega)$ for $P$-a.a. $\omega \in \Omega$ yields that $P$-a.s.

$$(\lambda \otimes Q)\big((t, \omega') \in [\![\tau, \infty[\![ \,: (b_t(\omega'), a_t(\omega')) \notin \Theta(t, \omega')\big) = 0.$$

Finally, as $Q_\omega(\tau = \tau(\omega)) = 1$ for $P$-a.a. $\omega \in \Omega$, Corollary 6.13 and Fubini's theorem yield that

$$
\begin{aligned}
(\lambda \otimes \overline{P})(R) &= E^{\overline{P}}\Big[ \int_\tau^\infty \mathbb{1}\{(b_t, a_t) \notin \Theta(t, X)\} dt \Big] \\
&= E^P\Big[ E^Q\Big[ \int_\tau^\infty \mathbb{1}\{(b_t, a_t) \notin \Theta(t, X)\} dt \Big]\Big] \\
&= E^P\Big[ \int_\tau^\infty Q\big((b_t, a_t) \notin \Theta(t, X)\big) dt \Big] = 0.
\end{aligned}
$$

This completes the proof. $\qquad\square$

### 6.4 Proof of Theorem 3.1

Lemma 6.6, Corollary 6.12 and Lemma 6.17 show that the prerequisites of [11, Theorem 2.1] are fulfilled and this theorem then directly implies the DPP. $\qquad\square$

## 7 A nonlinear Kolmogorov equation: proof of Theorem 4.3

In this section we prove that the value function is a weak sense viscosity solution to the PPDE (4.2) and we discuss some of its regularity properties. By its very definition, regularity of the value function is closely linked to the regularity of the correspondence $(t, \omega) \mapsto \mathcal{C}(t, \omega)$. Due to the appearance of $\mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}(t)$ and $(dB_{\cdot+t}^P/d\lambda, dC_{\cdot+t}^P/d\lambda)$ in the set $\mathcal{C}(t, \omega)$, certain regularity properties of $(t, \omega) \mapsto \mathcal{C}(t, \omega)$ seem at first glance to be difficult to verify. To get a more convenient condition, in Section 7.1, we show that

$$v(t, \omega) = \sup_{P \in \mathcal{C}(t, \omega)} E^P\big[\psi\big] = \sup_{P \in \mathcal{R}(t, \omega)} E^P\big[\psi(\omega \,\widetilde{\otimes}_t\, X)\big].$$

This reformulation of the value function $v$ explains that it suffices to investigate the correspondence $(t, \omega) \mapsto \mathcal{R}(t, \omega)$ from (4.3). Thereby, we shift the main $(t, \omega)$ dependence to the explicitly given correspondence $\Theta$, as the remaining parts from $\mathcal{R}(t, \omega)$ only depend on $\mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}} = \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}(0)$ and $(dB^P/d\lambda, dC^P/d\lambda)$. In the Sections 7.2 and 7.3, we show that $v$ is a viscosity sub- and supersolution to (4.2). The ideas of proof are based on applications of Berge's maximum theorem, Skorokhod's existence theorem for stochastic differential equations and Lebesgue's differentiation theorem. In contrast to the proofs from [12, 32] for the viscosity subsolution property in Lévy and continuous affine frameworks respectively, we do not work with explicit moment estimates. This allows us to extend the class of test functions to $C^{1,2}$ in comparison to the class $C^{2,3}$ as used in [12, 32]. Further, this extension also enables us to apply the uniqueness result from [43] to deduce Theorem 4.13. Given the above mentioned results, the proof of Theorem 4.3 is finalized in Section 7.4.

### 7.1 Some preparations

To execute the program outlined above, we require more notation. First, for $t \in \mathbb{R}_+$, we define another shift operator $\gamma_t\colon \Omega \to \Omega$ by $\gamma_t(\omega) := \omega((\cdot - t)^+)$ for all $\omega \in \Omega$. For $P \in \mathfrak{P}(\Omega)$, we denote $P^t := P \circ \gamma_t^{-1}$. Moreover, for $(t, \omega) \in [\![0, \infty[\![$, we write $\xi_{t,\omega}(\omega') := \omega \otimes_t \omega'$ and define

$$
\begin{aligned}
\mathcal{Q}(t, \omega) := \Big\{ P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}(t)\colon\ &P \circ X_t^{-1} = \delta_{\omega(t)}, \\
&(\lambda \otimes P)\text{-a.e. } (dB_{\cdot+t}^P/d\lambda, dC_{\cdot+t}^P/d\lambda) \in \Theta(\cdot + t, \omega \otimes_t X) \Big\}.
\end{aligned}
$$

Further, recall the definition of $\mathcal{R}$ as given in (4.3).

**Lemma 7.1.** *The equality $\mathcal{R}(t,\omega) = \{P_t \colon P \in \mathcal{Q}(t,\omega)\}$ holds for every $(t,\omega) \in [\![0,\infty[\![$.*

*Proof.* Lemma 6.5 shows the inclusion $\{P_t \colon P \in \mathcal{Q}(t,\omega)\} \subset \mathcal{R}(t,\omega)$, as $P_t \circ X_0^{-1} = P \circ X_t^{-1} = \delta_{\omega(t)}$ for every $P \in \mathcal{Q}(t,\omega)$. Conversely, given $P \in \mathcal{R}(t,\omega)$, the measure $P^t$ is contained in $\mathcal{Q}(t,\omega)$. To see this, note first that $P^t \circ X_t^{-1} = P \circ X_0^{-1} = \delta_{\omega(t)}$. Second, $P^t \in \mathfrak{P}_{\text{sem}}^{\text{ac}}(t)$ since $(P^t)_t = P \in \mathfrak{P}_{\text{sem}}^{\text{ac}}$ as $\theta_t \circ \gamma_t = \text{id}$. Finally, Lemma 6.4 together with $(\omega \otimes_t X) \circ \gamma_t = \omega \,\widetilde{\otimes}_t\, X$ implies that $(\lambda \otimes P^t)$-a.e. $(dB_{\cdot+t}^{P^t}/d\lambda, dC_{\cdot+t}^{P^t}/d\lambda) \in \Theta(\cdot+t, \omega \otimes_t X)$. $\qquad\square$

**Lemma 7.2.** *The equality $\mathcal{C}(t,\omega) = \{P \circ \xi_{t,\omega}^{-1} \colon P \in \mathcal{Q}(t,\omega)\}$ holds for every $(t,\omega) \in [\![0,\infty[\![$.*

*Proof.* For any measure $P \in \mathfrak{P}(\Omega)$, and every $(t,\omega) \in [\![0,\infty[\![$, we have $P(X = \omega \text{ on } [0,t]) = 1$ if and only if $P = P \circ \xi_{t,\omega}^{-1}$. Note that the canonical process $X$ is a semimartingale after time $t$ if and only if the process $\omega \otimes_t X$ is a semimartingale after time $t$. Thus, Lemma 6.4 implies, together with the identity $\xi_{t,\omega} \circ \xi_{t,\omega} = \xi_{t,\omega}$, that $P \in \mathcal{Q}(t,\omega)$ if and only if $P \circ \xi_{t,\omega}^{-1} \in \mathcal{Q}(t,\omega)$. This completes the proof. $\qquad\square$

Summarizing the above, we may conclude the following corollary, which provides, for our framework, the connection between the approaches in [11] and [34] for the construction of nonlinear expectations on path spaces.

**Corollary 7.3.** *For every upper semianalytic function $\psi \colon \Omega \to \mathbb{R}$, the equality*

$$\sup_{P \in \mathcal{C}(t,\omega)} E^P[\psi] = \sup_{P \in \mathcal{R}(t,\omega)} E^P[\psi(\omega \,\widetilde{\otimes}_t\, X)]$$

*holds for every $(t,\omega) \in [\![0,\infty[\![$.*

*Proof.* Using Lemma 7.2 for the first, the identity $\xi_{t,\omega} = (\omega \,\widetilde{\otimes}_t\, X) \circ \theta_t$ for the second, and Lemma 7.1 for the final equality, we obtain

$$\begin{aligned}
\sup_{P \in \mathcal{C}(t,\omega)} E^P[\psi] &= \sup_{P \in \mathcal{Q}(t,\omega)} E^P[\psi \circ \xi_{t,\omega}] \\
&= \sup_{P \in \mathcal{Q}(t,\omega)} E^{P_t}[\psi(\omega \,\widetilde{\otimes}_t\, X)], \\
&= \sup_{P \in \mathcal{R}(t,\omega)} E^P[\psi(\omega \,\widetilde{\otimes}_t\, X)].
\end{aligned}$$

The proof is complete. $\qquad\square$

For $N > 0$, we call a set $G \subset \Omega$ *$N$-bounded* if

$$\sup\big\{\|\omega(t)\| \colon \omega \in G, t \in [0,N]\big\} < \infty.$$

By the Arzelà–Ascoli theorem, any relatively compact set $G \subset \Omega$ is $N$-bounded for every $N > 0$.

**Lemma 7.4.** *Suppose that Condition 2.4 holds. For any $N > 0$, any $N$-bounded set $G \subset \Omega$ and any bounded set $K \subset \mathbb{R}^d$, the set*

$$\mathcal{R}^* := \bigcup_{\substack{t \in [0,N] \\ \omega \in G}} \Big\{ P \in \mathfrak{P}_{\text{sem}}^{\text{ac}} \colon P \circ X_0^{-1} \in \{\delta_x \colon x \in K\},$$

$$(\lambda \otimes P)\text{-}a.e.\ (dB^P/d\lambda, dC^P/d\lambda) \in \Theta(\cdot + t, \omega \,\widetilde{\otimes}_t\, X) \Big\}$$

*is relatively compact in $\mathfrak{P}(\Omega)$. Moreover, for every $p \geq 1$,*

$$\sup_{P \in \mathcal{R}^*} E^P\Big[\sup_{s \in [0,T]} \|X_s\|^{2p}\Big] < \infty. \tag{7.1}$$

*Proof.* Thanks to Prohorov's theorem, we need to show that $\mathcal{R}^*$ is tight, which we do by an application of Kolmogorov's criterion ([25, Theorem 21.42]), i.e., we show the following two conditions:

(a) the family $\{P \circ X_0^{-1} \colon P \in \mathcal{R}^*\}$ is tight;

(b) for each $T > 0$ there are numbers $C, \alpha, \beta > 0$ such that, for all $s, t \in [0, T]$, we have

$$\sup_{P \in \mathcal{R}^*} E^P\big[\|X_s - X_t\|^\alpha\big] \leq C|s - t|^{\beta+1}.$$

Part (a) follows easily from the fact that $\sup_{P \in \mathcal{R}^*} E^P[\|X_0\|] < \infty$, which uses the boundedness of $K$. We now show (b). Fix $T > 0, p \geq 1$ and set

$$T_n := \inf\{t \geq 0 \colon \|X_t\| \geq n\}, \quad n \in \mathbb{N}.$$

For a moment, take $P \in \mathcal{R}^*$ and let $(b^P, a^P)$ be the Lebesgue densities of the semimartingale characteristics of $X$ under $P$. Using the Burkholder–Davis–Gundy inequality, Hölder's inequality and the linear growth assumption, i.e., Condition 2.4, for all $t \in [0, T]$, we obtain that

$$E^P\Big[\sup_{s \in [0, t \wedge T_n]} \|X_s\|^{2p}\Big]$$

$$\leq C\Big(1 + T^{2p-1} \int_0^t E^P\big[\|b_{s \wedge T_n}^P\|^{2p}\big]ds + T^{p-1} \int_0^t E^P\big[\operatorname{tr}\big[a_{s \wedge T_n}^P\big]^p\big]ds\Big)$$

$$\leq C\Big(1 + \int_0^t E^P\Big[\sup_{r \in [0, s \wedge T_n]} \|X_r\|^{2p}\Big]ds\Big).$$

Here, the constant might depend on $d, p, T, N, G$ and $K$ but it is independent of $n, t$ and $P$. Finally, using Gronwall's lemma and Fatou's lemma, we conclude that (7.1) holds. It remains to finish the proof for (b). Take $0 \leq s \leq t \leq T$. Using again the Burkholder–Davis–Gundy inequality, the linear growth assumption and (7.1), we get, for every $P \in \mathcal{R}^*$, that

$$E^P\big[\|X_t - X_s\|^4\big] \leq C\Big((t-s)^3 \int_s^t E^P\big[\|b_r^P\|^4\big]dr + (t-s) \int_s^t E^P\big[\operatorname{tr}\big[a_r^P\big]^2\big]dr\Big)$$

$$\leq C\big((t-s)^4 + (t-s)^2\big)\Big(1 + E^P\Big[\sup_{r \in [0, T]} \|X_r\|^4\Big]\Big)$$

$$\leq C\big((t-s)^4 + (t-s)^2\big),$$

where the constant $C > 0$ is independent of $s, t$ and $P$. Consequently, we conclude that (b) holds with $\alpha = 4$ and $\beta = 1$. The proof is complete. $\qquad\square$

We collect more technical observations. Recall that $D(\mathbb{R}_+; \mathbb{R}^d)$ denotes the space of càdlàg functions from $\mathbb{R}_+$ into $\mathbb{R}^d$. In the following we endow $D(\mathbb{R}_+; \mathbb{R}^d)$ with the Skorokhod $J_1$ topology, see [23, Section VI.1] or [16, Section XV.1] for details on this topology. For $\omega, \alpha \in \Omega$ and $t \in \mathbb{R}_+$, we define the concatenation

$$\omega \widehat{\otimes}_t \alpha := \omega \mathbb{1}_{[0,t)} + \alpha \mathbb{1}_{[t,\infty)}.$$

Notice that $\omega \widehat{\otimes}_t \alpha \in D(\mathbb{R}_+; \mathbb{R}^d)$. Furthermore, recall that we endow $\Omega$ with the local uniform topology.

**Lemma 7.5.** *The map* $\Omega \times \Omega \times (0, \infty) \ni (\omega, \alpha, t) \mapsto \omega \widehat{\otimes}_t \alpha \in D(\mathbb{R}_+; \mathbb{R}^d)$ *is continuous.*

*Proof.* Take two sequences $(\omega^n)_{n\in\mathbb{N}}, (\alpha^n)_{n\in\mathbb{N}} \subset \Omega$ and two functions $\omega, \alpha \in \Omega$ such that $\omega^n \to \omega$ and $\alpha^n \to \alpha$. Furthermore, take a sequence $(t^n)_{n\in\mathbb{N}} \subset (0,\infty)$ and a time $t > 0$ such that $t^n \to t$. We have to prove that $\omega^n \widehat{\otimes}_{t^n} \alpha^n \to \omega \widehat{\otimes}_t \alpha$ in the Skorokhod $J_1$ topology. For $n \in \mathbb{N}$, define

$$\lambda^n(s) := \tfrac{t^n s}{t}\mathbb{1}_{\{s<t\}} + (s + t^n - t)\mathbb{1}_{\{s\geq t\}}, \quad s \in \mathbb{R}_+.$$

Evidently, $\lambda^n$ is a strictly increasing continuous function such that $\lambda^n(0) = 0$ and $\lambda^n(s) \to \infty$ as $s \to \infty$. Furthermore, for every $N > 0$, we have

$$\sup_{s\in[0,N]} |\lambda^n(s) - s| \leq N|\tfrac{t^n}{t} - 1| + |t^n - t| \to 0$$

as $n \to \infty$. Notice that

$$(\omega^n \widehat{\otimes}_{t^n} \alpha^n)(\lambda_n(s)) = \omega^n(\tfrac{t^n s}{t})\mathbb{1}_{\{s<t\}} + \alpha^n(s + t^n - t)\mathbb{1}_{\{s\geq t\}}, \quad s \in \mathbb{R}_+.$$

Thus, for every $N > 0$ and some $T = T_N > 0$ large enough, we have

$$\sup_{s\in[0,N]} \|(\omega^n \widehat{\otimes}_{t^n} \alpha^n)(\lambda^n(s)) - (\omega \widehat{\otimes}_t \alpha)(s)\|$$

$$\leq \sup\{\|\omega^n(s) - \omega(s)\|\colon s \in [0,N]\} + \sup\{\|\alpha^n(s) - \alpha(s)\|\colon s \in [0,N]\}$$
$$+ \sup\{\|\omega^n(z) - \omega^n(v)\|\colon z, v \in [0,T], |z - v| \leq N|\tfrac{t^n}{t} - 1|\}$$
$$+ \sup\{\|\alpha^n(z) - \alpha^n(v)\|\colon z, v \in [0,T], |z - v| \leq |t^n - t|\}.$$

Thanks to the Arzelà–Ascoli theorem, we conclude that all terms on the r.h.s. converge to zero as $n \to \infty$, which implies that

$$\sup_{s\in[0,N]} \|(\omega^n \widehat{\otimes}_{t^n} \alpha^n)(\lambda^n(s)) - (\omega \widehat{\otimes}_t \alpha)(s)\| \to 0$$

as $n \to \infty$. Consequently, by virtue of [16, Theorem 15.10], $\omega^n \widehat{\otimes}_{t^n} \alpha^n \to \omega \widehat{\otimes}_t \alpha$ in the Skorokhod $J_1$ topology. The proof is complete. □

In the following lemma we take care of the case $t = 0$.

**Lemma 7.6.** *Let* $(\omega^n)_{n\in\mathbb{Z}_+}, (\alpha^n)_{n\in\mathbb{Z}_+} \subset \Omega$ *and* $(t^n)_{n\in\mathbb{N}} \subset \mathbb{R}_+$ *such that* $\omega^n \to \omega^0, \alpha^n \to \alpha^0$ *and* $t^n \to 0$. *Furthermore, suppose that* $\omega^0(0) = \alpha^0(0)$. *Then,*

$$\omega^n \widehat{\otimes}_{t^n} \alpha^n \to \alpha^0$$

*locally uniformly.*

*Proof.* For every $T > 0$, we estimate

$$\sup_{s\in[0,T]} \|(\omega^n \widehat{\otimes}_{t^n} \alpha^n)(s) - \alpha^0(s)\|$$

$$\leq \sup_{s\in[0,T]} \|\omega^n(s) - \alpha^0(s)\|\mathbb{1}_{\{s<t^n\}} + \sup_{s\in[0,T]} \|\alpha^n(s) - \alpha^0(s)\|\mathbb{1}_{\{s\geq t^n\}}$$
$$\leq \sup_{s\in[0,T]} \|\omega^n(s) - \omega^n(0) + \omega^n(0) - \omega^0(0) + \alpha^0(0) - \alpha^0(s)\|\mathbb{1}_{\{s<t^n\}}$$
$$+ \sup_{s\in[0,T]} \|\alpha^n(s) - \alpha^0(s)\|$$
$$\leq \sup\{\|\omega^n(s) - \omega^n(r)\|\colon s, r \in [0,T], |s - r| \leq t^n\} + \|\omega^n(0) - \omega^0(0)\|$$
$$+ \sup\{\|\alpha^0(s) - \alpha^0(r)\|\colon s, r \in [0,T], |s - r| \leq t^n\}$$
$$+ \sup_{s\in[0,T]} \|\alpha^n(s) - \alpha(s)\|.$$

By the Arzelà–Ascoli theorem, the r.h.s. tends to zero as $n \to \infty$. This completes the proof. □

**Corollary 7.7.** *The maps*

$$\Omega \times \Omega \times \mathbb{R}_+ \ni (\omega, \alpha, t) \mapsto \omega \otimes_t \alpha \in \Omega, \qquad \Omega \times \Omega \times \mathbb{R}_+ \ni (\omega, \alpha, t) \mapsto \omega \widetilde{\otimes}_t \alpha \in \Omega$$

*are continuous.*

*Proof.* We only prove that $(\omega, \alpha, t) \mapsto \omega \widetilde{\otimes}_t \alpha$ is continuous. The continuity of $(\omega, \alpha, t) \mapsto \omega \otimes_t \alpha$ follows the same way. Let $(\omega^n)_{n \in \mathbb{Z}_+}, (\alpha^n)_{n \in \mathbb{Z}_+} \subset \Omega$ and $(t^n)_{n \in \mathbb{Z}_+} \subset \mathbb{R}_+$ be sequences such that $\omega^n \to \omega^0, \alpha^n \to \alpha^0$ and $t^n \to t^0$. Furthermore, define

$$\gamma^n := \omega^n(t^n) + \alpha^n((\cdot - t^n)^+) - \alpha^n(0), \qquad n \in \mathbb{Z}_+.$$

For every $T > \sup_{n \in \mathbb{Z}_+} t^n$, we estimate

$$\sup_{s \in [0,T]} \|\gamma^n(s) - \gamma^0(s)\|$$

$$\leq \|\omega^n(t^n) - \omega^0(t^0)\| + \sup_{s \in [0,T]} \|\alpha^n((s - t^n)^+) - \alpha^0((s - t^0)^+)\| + \|\alpha^n(0) - \alpha^0(0)\|$$

$$\leq \sup\{\|\omega^n(s) - \omega^n(r)\| : s, r \in [0,T], |s - r| \leq |t^n - t^0|\} + \|\omega^n(t^0) - \omega^0(t^0)\|$$

$$+ \sup\{\|\alpha^n(s) - \alpha^n(r)\| : s, r \in [0,T], |s - r| \leq |t^n - t^0|\}$$

$$+ \sup_{s \in [0,T]} \|\alpha^n(s) - \alpha^0(s)\| + \|\alpha^n(0) - \alpha^0(0)\|.$$

By the Arzelà–Ascoli theorem, the r.h.s. converges to zero as $n \to \infty$. Hence, we conclude that $\gamma^n \to \gamma^0$. If $t^0 > 0$, then we can assume that $t^n > 0$ for all $n \in \mathbb{N}$ and Lemma 7.5 yields that

$$\omega^n \widetilde{\otimes}_{t^n} \alpha^n = \omega^n \widehat{\otimes}_{t^n} \gamma^n \to \omega^0 \widehat{\otimes}_{t^0} \gamma^0 = \omega^0 \widetilde{\otimes}_{t^0} \alpha^0 \tag{7.2}$$

in the Skorokhod $J_1$ topology. As $\omega^0 \widetilde{\otimes}_{t^0} \alpha^0 \in \Omega$, we get from [23, Proposition VI.1.17] that the convergence is even locally uniformly. Finally, in case $t^0 = 0$, Lemma 7.6 yields that (7.2) holds locally uniformly, because $\gamma^0(0) = \omega^0(0)$. We conclude the continuity of $(\omega, \alpha, t) \mapsto \omega \widetilde{\otimes}_t \alpha$. □

For $\alpha, \omega \in D(\mathbb{R}_+; \mathbb{R}^d)$ and $t \in \mathbb{R}_+$, we define

$$\omega \,\overline{\otimes}_t\, \alpha := \omega \mathbb{1}_{[0,t)} + \alpha(\cdot - t)\mathbb{1}_{[t,\infty)} \in D(\mathbb{R}_+; \mathbb{R}^d).$$

**Lemma 7.8.** *Let* $(t, \omega) \in [0, T] \times D(\mathbb{R}_+; \mathbb{R}^d)$ *and* $\phi \in C^{1,2}([t, T] \times D(\mathbb{R}_+; \mathbb{R}^d); \mathbb{R})$. *Then,*

$$((s, \alpha) \mapsto F(s, \alpha) := \phi(s + t, \omega \,\overline{\otimes}_t\, \alpha)) \in C^{1,2}([0, T - t] \times D(\mathbb{R}_+; \mathbb{R}^d); \mathbb{R}),$$

*with*

$$\dot\partial F(s, \alpha) = \dot\partial \phi(s + t, \omega \,\overline{\otimes}_t\, \alpha),$$
$$\nabla F(s, \alpha) = \nabla \phi(s + t, \omega \,\overline{\otimes}_t\, \alpha), \tag{7.3}$$
$$\nabla^2 F(s, \alpha) = \nabla^2 \phi(s + t, \omega \,\overline{\otimes}_t\, \alpha).$$

*for all* $(s, \alpha) \in [0, T - t] \times D(\mathbb{R}_+; \mathbb{R}^d)$.

*Proof.* For $\alpha \in D(\mathbb{R}_+; \mathbb{R}^d)$, $0 \leq s < T - t$ and $h > 0$ small enough, a short computation shows that

$$F(h + s, \alpha(\cdot \wedge s)) - F(s, \alpha(\cdot \wedge s))$$

$$= \phi(h + s + t, \omega \,\overline{\otimes}_t\, \alpha(\cdot \wedge s)) - \phi(s + t, \omega \,\overline{\otimes}_t\, \alpha(\cdot \wedge s))$$

$$= \phi(h + s + t, (\omega \overline{\otimes}_t \alpha)(\cdot \wedge (s + t))) - \phi(s + t, (\omega \overline{\otimes}_t \alpha)(\cdot \wedge (s + t))),$$

and we conclude that $\dot{\partial} F(s, \alpha) = \dot{\partial} \phi(s + t, \omega \overline{\otimes}_t \alpha)$. Furthermore, we get, for every $i = 1, \ldots, d$, that

$$F(s, \alpha + he_i \mathbb{1}_{[s, T-t]}) - F(s, \alpha) = \phi(s + t, \omega \overline{\otimes}_t (\alpha + he_i \mathbb{1}_{[s, T-t]})) - \phi(s + t, \omega \overline{\otimes}_t \alpha)$$
$$= \phi(s + t, (\omega \overline{\otimes}_t \alpha) + he_i \mathbb{1}_{[s+t, T]}) - \phi(s + t, \omega \overline{\otimes}_t \alpha),$$

which implies $\nabla F(s, \alpha) = \nabla \phi(s + t, \omega \overline{\otimes}_t \alpha)$. In the same way, we derive the formula for $\nabla^2 F$. Similar to (4.1), for every $(u, \zeta), (s, \alpha) \in [0, T] \times D(\mathbb{R}_+; \mathbb{R}^d)$, we define

$$d_T((u, \zeta), (s, \alpha)) := |u - s| + \sup_{r \in [0, T]} \|\zeta(r \wedge u) - \alpha(r \wedge s)\|.$$

For every $(u, \zeta), (s, \alpha) \in [0, T - t] \times D(\mathbb{R}_+; \mathbb{R}^d)$, we compute that

$$d_T((u + t, \omega \overline{\otimes}_t \zeta), (s + t, \omega \overline{\otimes}_t \alpha))$$
$$= |u - s| + \sup_{r \in [0, T]} \|\zeta((r - t) \wedge u) - \alpha((r - t) \wedge s)\| \mathbb{1}_{\{r \geq t\}}$$
$$\leq d_{T-t}((u, \zeta), (s, \alpha)).$$

Hence, the map

$$(s, \alpha) \mapsto (s + t, \omega \overline{\otimes}_t \alpha) \tag{7.4}$$

is (Lipschitz) continuous from $([0, T - t] \times D(\mathbb{R}_+; \mathbb{R}^d), d_{T-t})$ into $([0, T] \times D(\mathbb{R}_+; \mathbb{R}^d), d_T)$. We conclude that the claimed continuity properties of $F, \dot{\partial} F, \nabla F, \nabla^2 F$ follow from continuity of (7.4) together with the assumed continuity properties of $\phi, \dot{\partial} \phi, \nabla \phi, \nabla^2 \phi$. The proof is complete. $\qquad \square$

## 7.2 Subsolution property

In this section we prove that the value function is a weak sense viscosity subsolution to the nonlinear PPDE (4.2).

**Lemma 7.9.** *Assume that $F$ is a compact metrizable space and that the Conditions 2.4, 2.6 and 2.7 hold. The value function $v$ is a weak sense viscosity subsolution to (4.2).*

*Proof.* Notice that

$$v(T, \omega) = \sup_{P \in \mathcal{C}(T, \omega)} E^P[\psi] = \psi(\omega),$$

as $P(X = \omega \text{ on } [0, T]) = 1$ for all $P \in \mathcal{C}(T, \omega)$. Thus, we only have to prove the subsolution property. Take $(t, \omega) \in [\![0, T[\![$ and $\phi \in C^{1,2}_{pol}([\![t, T]\!]; \mathbb{R})$ satisfying

$$0 = (v - \phi)(t, \omega) = \sup\{(v - \phi)(s, \omega') \colon (s, \omega') \in [\![t, T]\!]\}.$$

Fix $0 < u < T - t$. The DPP (Theorem 3.1) and Corollary 7.3 show that

$$0 = \sup_{P \in \mathcal{C}(t, \omega)} E^P[v(u + t, X) - v(t, \omega)] \leq \sup_{P \in \mathcal{C}(t, \omega)} E^P[\phi(u + t, X) - \phi(t, \omega)]$$
$$= \sup_{P \in \mathcal{R}(t, \omega)} E^P[\phi(u + t, \omega \widetilde{\otimes}_t X) - \phi(t, \omega)]. \tag{7.5}$$

We fix $P \in \mathcal{R}(t, \omega)$ and denote the Lebesgue densities of the $P$-characteristics of $X$ by $(b^P, a^P)$. The pathwise Itô formula given by [3, Theorem 2.2] yields, together with Lemma 7.8, that

$$\phi(u + t, \omega \widetilde{\otimes}_t X) - \phi(t, X)$$
$$= \int_0^u \dot{\partial} \phi(s + t, \omega \widetilde{\otimes}_t X) ds + \frac{1}{2} \int_0^u \text{tr} \left[ \nabla^2 \phi(s + t, \omega \widetilde{\otimes}_t X) a_s^P \right] ds$$
$$+ \int_0^u \langle \nabla \phi(s + t, \omega \widetilde{\otimes}_t X), dX_s \rangle.$$

By virtue of the linear growth condition, the polynomial growth of $\nabla\phi$ and the moment bound from Lemma 7.4, the local martingale part of the stochastic integral above is a true martingale and we get

$$E^P\Big[\int_0^u \langle\nabla\phi(s+t,\omega \;\widetilde{\otimes}_t\; X),dX_s\rangle\Big] = E^P\Big[\int_0^u \langle\nabla\phi(s+t,\omega \;\widetilde{\otimes}_t\; X),b_s^P\rangle ds\Big],$$

which implies that

$$
\begin{aligned}
E^P\big[\phi(u+t,\omega \;&\widetilde{\otimes}_t\; X) - \phi(t,\omega)\big] \\
&= E^P\Big[\int_0^u \dot\partial\phi(s+t,\omega \;\widetilde{\otimes}_t\; X)ds \\
&\quad + \int_0^u \big(\langle\nabla\phi(s+t,\omega \;\widetilde{\otimes}_t\; X),b_s^P\rangle + \tfrac{1}{2}\operatorname{tr}\big[\nabla^2\phi(s+t,\omega \;\widetilde{\otimes}_t\; X)a_s^P\big]\big)ds\Big].
\end{aligned}
\tag{7.6}
$$

For $(u,\alpha) \in [\![0,T-t]\!]$, we set

$$
\begin{aligned}
\mathfrak{G}(u,\alpha) := \dot\partial\phi(u+t,\omega \;\widetilde{\otimes}_t\; \alpha) + \sup\Big\{ &\langle\nabla\phi(u+t,\omega \;\widetilde{\otimes}_t\; \alpha),b(f,u+t,\omega \;\widetilde{\otimes}_t\; \alpha)\rangle \\
&+ \tfrac{1}{2}\operatorname{tr}\big[\nabla^2\phi(u+t,\omega \;\widetilde{\otimes}_t\; \alpha)a(f,u+t,\omega \;\widetilde{\otimes}_t\; \alpha)\big] : f \in F\Big\}.
\end{aligned}
$$

From (7.5) and (7.6), we get that

$$0 \le \frac{1}{u}\int_0^u \sup_{P\in\mathcal{R}(t,\omega)} E^P\big[\mathfrak{G}(s,X)\big]ds, \quad u > 0. \tag{7.7}$$

We investigate the right hand side when $u \searrow 0$.

**Lemma 7.10.**

$$\frac{1}{u}\int_0^u \sup_{P\in\mathcal{R}(t,\omega)} E^P\big[\mathfrak{G}(s,X)\big]ds \to \sup_{P\in\mathcal{R}(t,\omega)} E^P\big[\mathfrak{G}(0,X)\big], \quad u \searrow 0.$$

*Proof.* By Lebesgue's differentiation theorem, it suffices to prove that the function $s \mapsto \sup_{P\in\mathcal{R}(t,\omega)} E^P[\mathfrak{G}(s,X)]$ is continuous. By the compactness of $\mathcal{R}(t,\omega)$, which follows from Theorem 4.4, and Berge's maximum theorem ([1, Theorem 17.31]), continuity of $s \mapsto \sup_{P\in\mathcal{R}(t,\omega)} E^P[\mathfrak{G}(s,X)]$ is implied by the continuity of

$$\mathcal{R}(t,\omega) \times [0,T] \ni (P,s) \mapsto E^P\big[\mathfrak{G}(s,X)\big].$$

Take a sequence $(P^n,s^n)_{n\in\mathbb{Z}_+} \subset \mathcal{R}(t,\omega) \times [0,T]$ such that $(P^n,s^n) \to (P^0,s^0)$. By Skorokhod's coupling theorem, on some probability space, there are random variables $(X^n)_{n\in\mathbb{Z}_+}$ with laws $(P^n)_{n\in\mathbb{Z}_+}$ such $X^n \to X^0$ almost surely. Due to the compactness of $F$, the continuity assumptions on $b$ and $a$ and Corollary 7.7, we deduce from Berge's maximum theorem that $\mathfrak{G}$ is continuous. Thus, we get a.s. $\mathfrak{G}(s^n,X^n) \to \mathfrak{G}(s^0,X^0)$. Thanks to the linear growth conditions on $b$ and $a$, the polynomial growth assumptions on the derivatives of $\phi$, and Lemma 7.4, we notice that

$$\sup_{n\in\mathbb{N}} E\big[|\mathfrak{G}(s^n,X^n)|^2\big] \le C\Big(1 + \sup_{P\in\mathcal{R}(t,\omega)} E^P\Big[\sup_{r\in[0,T]} \|X_r\|^p\Big]\Big) < \infty,$$

where $p \ge 4$ is a suitable power (which depends on the polynomial bounds of the derivatives of $\phi$). This estimate shows that the sequence $(\mathfrak{G}(s^n,X^n))_{n\in\mathbb{N}}$ is uniformly integrable and we conclude that

$$E\big[\mathfrak{G}(s^n,X^n)\big] \to E\big[\mathfrak{G}(s^0,X^0)\big].$$

This is the claimed continuity and therefore, the proof is complete. $\qquad\square$

Finally, it follows from (7.7) and Lemma 7.10 that

$$0 \leq \sup_{P \in \mathcal{R}(t,\omega)} E^P\big[\mathfrak{G}(0,X)\big]$$

$$= \dot{\partial}\phi(t,\omega) + \sup\big\{\langle \nabla\phi(t,\omega), b(f,t,\omega)\rangle + \tfrac{1}{2}\operatorname{tr}\big[\nabla^2\phi(t,\omega)a(f,t,\omega)\big] \colon f \in F\big\}.$$

This completes the proof of the subsolution property. $\qquad\square$

### 7.3 Supersolution property

The next lemma is the central tool for the proof of the supersolution property.

**Lemma 7.11.** *Suppose that the Conditions 2.4 and 2.6 hold. For every* $(f,t,\omega) \in F \times [\![0,\infty[\![$, *there exists a probability measure* $\overline{P} \in \mathcal{R}(t,\omega)$ *such that the* $\overline{P}$-*characteristics of* $X$ *(for its right-continuous natural filtration) have Lebesgue densities* $(b(f,\cdot+t,\omega \otimes_t X)$, $a(f,\cdot+t,\omega \widetilde{\otimes}_t X))$.

*Proof.* Thanks to the Conditions 2.4 and 2.6, and Corollary 7.7, the functions $(s,\alpha) \mapsto b(f,s+t,\omega \widetilde{\otimes}_t \alpha)$ and $(s,\alpha) \mapsto a(f,s+t,\omega \widetilde{\otimes}_t \alpha)$ are continuous and of linear growth (in the sense of (4) on p. 258 in [13]). Hence, the claim follows from [13, Theorem 4, p. 265] and Lemma 6.4. $\qquad\square$

**Lemma 7.12.** *Suppose that the Conditions 2.4, 2.6 and 2.7 hold. The value function* $v$ *is a weak sense viscosity supersolution to* (4.2).

*Proof.* Similar to the proof of Lemma 7.9, it suffices to prove the supersolution property. Take $(t,\omega) \in [\![0,T[\![$ and $\phi \in C^{1,2}_{pol}([\![t,T]\!];\mathbb{R})$ such that

$$0 = (u - \phi)(t,\omega) = \inf\{(u-\phi)(s,\omega') \colon (s,\omega') \in [\![t,T]\!]\}.$$

Theorem 3.1 and Corollary 7.3 show that

$$0 = \sup_{P \in \mathcal{C}(t,\omega)} E^P\big[v(u+t,X) - v(t,\omega)\big] \geq \sup_{P \in \mathcal{R}(t,\omega)} E^P\big[\phi(u+t,\omega \widetilde{\otimes}_t X) - \phi(t,\omega)\big]. \quad (7.8)$$

Let us fix an $f \in F$. By Lemma 7.11, there exists a probability measure $\overline{P} \in \mathcal{R}(t,\omega)$ such that the $\overline{P}$-characteristics of $X$ have Lebesgue densities $(b(f,\cdot+t,\omega \widetilde{\otimes}_t X), a(f,\cdot+t,\omega \widetilde{\otimes}_t X))$. As in the proof of Lemma 7.9, we get that

$$E^{\overline{P}}\big[\phi(u+t,\omega \widetilde{\otimes}_t X) - \phi(t,\omega)\big]$$

$$= E^{\overline{P}}\Big[\int_0^u \big(\dot{\partial}\phi(s+t,\omega \widetilde{\otimes}_t X) + \langle \nabla\phi(s+t,\omega \widetilde{\otimes}_t X), b(f,s+t,\omega \widetilde{\otimes}_t X)\rangle\big)ds \quad (7.9)$$

$$+ \int_0^u \tfrac{1}{2}\operatorname{tr}\big[\nabla^2\phi(s+t,\omega \widetilde{\otimes}_t X)a(f,s+t,\omega \widetilde{\otimes}_t X)\big]ds\Big].$$

For $(u,\alpha) \in [\![0,T-t]\!]$, we set

$$\mathfrak{K}(u,\alpha) := \dot{\partial}\phi(u+t,\omega \widetilde{\otimes}_t \alpha) + \langle \nabla\phi(u+t,\omega \widetilde{\otimes}_t \alpha), b(f,u+t,\omega \widetilde{\otimes}_t \alpha)\rangle$$

$$+ \tfrac{1}{2}\operatorname{tr}\big[\nabla^2\phi(u+t,\omega \widetilde{\otimes}_t \alpha)a(f,u+t,\omega \widetilde{\otimes}_t \alpha)\big].$$

With (7.8) and (7.9), we obtain that

$$0 \geq \frac{1}{u}\int_0^u E^{\overline{P}}\big[\mathfrak{K}(s,X)\big]ds, \quad u > 0.$$

As $s \mapsto E^{\overline{P}}[\mathfrak{K}(s, X)]$ is continuous (see the proof of Lemma 7.10), we conclude, with $u \searrow 0$, that

$$0 \geq \frac{1}{u} \int_0^u E^{\overline{P}}[\mathfrak{K}(s, X)]ds \to E^{\overline{P}}[\mathfrak{K}(0, X)]$$
$$= \dot{\partial}\phi(t, \omega) + \langle \nabla\phi(t, \omega), b(f, t, \omega)\rangle + \tfrac{1}{2} \operatorname{tr}\left[\nabla^2\phi(t, \omega)a(f, t, \omega)\right].$$

As $f \in F$ is arbitrary, taking the $\sup$ over all $f \in F$ shows that $v$ is a weak sense viscosity supersolution. The proof is complete. $\qquad\square$

### 7.4 Proof of Theorem 4.3

It follows from Lemmata 7.9 and 7.12 that $v$ is a weak sense viscosity solution to the PPDE (4.2). We now discuss its regularity properties. Recall that $\psi$ is bounded and continuous. Thanks to Corollary 7.7, the map

$$\Omega \times \Omega \times [0, T] \ni (\omega, \alpha, t) \mapsto \psi(\omega \,\widetilde{\otimes}_t\, \alpha)$$

is continuous. Thus, by [2, Theorem 8.10.61], the map

$$[\![0, T]\!] \times \mathfrak{P}(\Omega) \ni (t, \omega, P) \mapsto E^P\left[\psi(\omega \,\widetilde{\otimes}_t\, X)\right] \tag{7.10}$$

is continuous, too. Furthermore, by Theorem 4.4, $\mathcal{R}$ is compact-valued. From Corollary 7.3 we get that

$$v(t, \omega) = \sup_{P \in \mathcal{R}(t, \omega)} E^P\left[\psi(\omega \,\widetilde{\otimes}_t\, X)\right], \quad (t, \omega) \in [\![0, T]\!].$$

Hence, under the continuity hypothesis on $\mathcal{R}$, the continuity (upper, lower semicontinuity) of $v$ follows from the continuity of (7.10) and Berge's maximum theorem ([1, Theorem 17.31, Lemmata 17.29, 17.30]). The proof is complete. $\qquad\square$

## 8  The regularity of $\mathcal{R}$ and $v$: proofs of Theorems 4.4, 4.7 and 4.12

In the present section we prove the Theorems 4.4, 4.7 and 4.12. By a careful refinement of some key arguments, Theorem 4.4 extends [5, Proposition 3.8] beyond the Markovian case. Using the implicit function theorem, strong existence properties of stochastic differential equations with random coefficients and Gronwall arguments, we establish Theorem 4.7 that is seemingly the first result regarding lower hemicontinuity of the correspondence $(t, \omega) \mapsto \mathcal{R}(t, \omega)$ and thus the first result on lower semicontinuity of the value function in a fully path-dependent framework related to nonlinear stochastic processes. The proof for Theorem 4.12 on the uniform continuity of $v$ with respect to the metric d combines some observations made in the proof for the lower hemicontinuity of $\mathcal{R}$ with an application of the DPP and a Gronwall argument, cf. also the proof of [42, Lemma 3.6].

### 8.1 Proof of Theorem 4.4

To prove Theorem 4.4, we use the sequential characterization of upper hemicontinuity as given by [1, Theorem 17.20]. Namely, by [1, Theorem 17.20], $\mathcal{R}$ is upper hemicontinuous with compact values if and only if for all sequences $(t^n, \omega^n)_{n \in \mathbb{N}} \subset [\![0, \infty[\![$ and $(P^n)_{n \in \mathbb{N}} \subset \mathfrak{P}(\Omega)$ such that $(t^n, \omega^n) \to (t, \omega) \in [\![0, \infty[\![$ and $P^n \in \mathcal{R}(t^n, \omega^n)$, the sequence $(P^n)_{n \in \mathbb{N}}$ has an accumulation point in $\mathcal{R}(t, \omega)$.

From now on, take $(t^n, \omega^n)_{n \in \mathbb{N}} \subset [\![0, \infty[\![$ and $(P^n)_{n \in \mathbb{N}} \subset \mathfrak{P}(\Omega)$ such that $(t^n, \omega^n) \to (t, \omega) \in [\![0, \infty[\![$ and $P^n \in \mathcal{R}(t^n, \omega^n)$. Thanks to Lemma 7.4, the set $\bigcup_{n \in \mathbb{N}} \mathcal{R}(t^n, \omega^n)$ is relatively compact. Hence, up to passing to a subsequence, we can assume that the

sequence $(P^n)_{n\in\mathbb{N}}$ converges weakly to a limiting measure $P$. In the following we prove that $P \in \mathcal{R}(t,\omega)$. Clearly, as $P^n \circ X_0^{-1} \to P \circ X_0^{-1}$ and $P^n \circ X_0^{-1} = \delta_{\omega^n(t^n)} \to \delta_{\omega(t)}$, we have $P \circ X_0^{-1} = \delta_{\omega(t)}$. Thus, we need to show that $P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}$ with differential characteristics $(\lambda \otimes P)$-a.e. in $\Theta(\cdot + t, \omega \widetilde{\otimes}_t X)$. The proof of this is split into four steps.

Before we start our program, we need a last bit of notation. For each $n \in \mathbb{N}$, denote the $P^n$-characteristics of $X$ by $(B^n, C^n)$. Set $\Omega^* := \Omega \times \Omega \times C(\mathbb{R}_+; \mathbb{R}^{d\times d})$ and denote the coordinate process on $\Omega^*$ by $Y = (Y^{(1)}, Y^{(2)}, Y^{(3)})$. Further, set $\mathcal{F}^* := \sigma(Y_s, s \geq 0)$ and let $\mathbf{F}^* = (\mathcal{F}_s^*)_{s\geq 0}$ be the right-continuous filtration generated by $Y$.

*Step 1.* First, we show that the family $\{P^n \circ (X, B^n, C^n)^{-1}: n \in \mathbb{N}\}$ is tight (when seen as a sequence of probability measures on the measurable space $(\Omega^*, \mathcal{F}^*)$). Since $P^n \to P$, it suffices to prove tightness of $\{P^n \circ (B^n, C^n)^{-1}: n \in \mathbb{N}\}$. We use Aldous' tightness criterion ([23, Theorem VI.4.5]), i.e., we show the following two conditions:

(a) for every $T, \varepsilon > 0$, there exists a $K \in \mathbb{R}_+$ such that

$$\sup_{n\in\mathbb{N}} P^n \Big( \sup_{s\in[0,T]} \|B_s^n\| + \sup_{s\in[0,T]} \mathrm{tr}\,[C_s^n] \geq K \Big) \leq \varepsilon;$$

(b) for every $T, \varepsilon > 0$,

$$\lim_{\theta\searrow 0} \limsup_{n\to\infty} \sup \big\{ P^n(\|B_L^n - B_S^n\| + \|C_L^n - C_S^n\|_o \geq \varepsilon) \big\} = 0,$$

where the $\sup$ is taken over all stopping times $S, L \leq T$ such that $S \leq L \leq S + \theta$.

By Lemma 7.4, we have

$$\sup_{n\in\mathbb{N}} E^{P^n} \Big[ \sup_{s\in[0,T]} \|X_s\|^2 \Big] < \infty. \tag{8.1}$$

Using the linear growth assumption, we also obtain that $P^n$-a.s.

$$\sup_{s\in[0,T]} \|B_s^n\| + \sup_{s\in[0,T]} \mathrm{tr}\,[C_s^n] \leq \mathsf{C}\Big(1 + \sup_{s\in[0,T]} \|X_s\|^2\Big),$$

where the constant $\mathsf{C} > 0$ is independent of $n$. By virtue of (8.1), this bound yields (a). For (b), take two stopping times $S, L \leq T$ such that $S \leq L \leq S + \theta$ for some $\theta > 0$. Then, using again the linear growth assumptions, we get $P^n$-a.s.

$$\|B_L^n - B_S^n\| + \|C_L^n - C_S^n\|_o \leq \mathsf{C}(L-S)\Big(1 + \sup_{s\in[0,T]} \|X_s\|^2\Big) \leq \mathsf{C}\theta\Big(1 + \sup_{s\in[0,T]} \|X_s\|^2\Big),$$

which yields (b) by virtue of (8.1). We conclude that $\{P^n \circ (X, B^n, C^n)^{-1}: n \in \mathbb{N}\}$ is tight. Up to passing to a subsequence, from now on we assume that $P^n \circ (X, B^n, C^n)^{-1} \to Q$ weakly.

*Step 2.* Next, we show that $Y^{(2)}$ and $Y^{(3)}$ are $Q$-a.s. locally absolutely continuous. For $M > 0$ and $\omega \in \Omega$, define

$$\rho_M(\omega) := \inf\{s \geq 0 : \|\omega(s)\| \geq M\} \wedge M.$$

Furthermore, for $\omega = (\omega^{(1)}, \omega^{(2)}) \in \Omega \times \Omega$, we set

$$\zeta_M(\omega) := \sup\Big\{ \frac{\|\omega^{(2)}(u \wedge \rho_M(\omega)) - \omega^{(2)}(s \wedge \rho_M(\omega))\|}{u - s} : 0 \leq s < u \Big\}.$$

Similar to the proof of [5, Lemma 3.6], we obtain the existence of a dense set $D \subset \mathbb{R}_+$ such that, for every $M \in D$, the map $\zeta_M$ is $Q \circ (Y^{(1)}, Y^{(2)})^{-1}$-a.s. lower semicontinuous. By the linear growth conditions and the definition of $\rho_M$, for every $M \in D$, there exists

a constant $C = C(M) > 0$ such that $P^n(\zeta_M(X, B^n) \leq C) = 1$ for all $n \in \mathbb{N}$. As $\zeta_M$ is $Q \circ (Y^{(1)}, Y^{(2)})$-a.s. lower semicontinuous, [38, Example 17, p. 73] yields that

$$0 = \liminf_{n \to \infty} P^n(\zeta_M(X, B^n) > C) \geq Q(\zeta_M(Y^{(1)}, Y^{(2)}) > C).$$

Further, as $D$ is dense in $\mathbb{R}_+$, we obtain that $Q$-a.s. $Y^{(2)}$ is locally Lipschitz continuous, i.e., in particular locally absolutely continuous. Similarly, we get that $Y^{(3)}$ is $Q$-a.s. locally Lipschitz and hence, locally absolutely continuous.

*Step 3.* Define the map $\Phi \colon \Omega^* \to \Omega$ by $\Phi(\omega^{(1)}, \omega^{(2)}, \omega^{(3)}) := \omega^{(1)}$. Clearly, $Q \circ \Phi^{-1} = P$ and $Y^{(1)} = X \circ \Phi$. In this step, we prove that $(\lambda \otimes Q)$-a.e. $(dY^{(2)}/d\lambda, dY^{(3)}/d\lambda) \in \Theta(\cdot + t, \omega \widetilde{\otimes}_t X) \circ \Phi$. By [5, Lemma 3.2] and the Conditions 2.6 and 2.7, the correspondence $\Theta$ is continuous with compact and convex values. Furthermore, again by Condition 2.6, the set $\Theta([s, s + 1], \omega^o)$ is compact for every $(s, \omega^o) \in [\![0, \infty[\![$. Consequently, as in completely metrizable locally convex spaces the closed convex hull of a compact set is itself compact ([1, Theorem 5.35]), we conclude that $\overline{\mathrm{co}}\, \Theta([s, s+1], \omega^o)$ is compact and, by [5, Lemma 3.4], we get that

$$\bigcap_{m \in \mathbb{N}} \overline{\mathrm{co}}\, \Theta([s, s + 1/m], \omega^o) \subset \Theta(s, \omega^o) \tag{8.2}$$

for all $(s, \omega^o) \in [\![0, \infty[\![$. Here, $\overline{\mathrm{co}}$ denotes the closure of the convex hull. By virtue of [9, Corollary 8, p. 48], $P^n$-a.s. for all $s \in \mathbb{R}_+$, we have

$$m(B_{s+1/m}^n - B_s^n, C_{s+1/m}^n - C_s^n) \in \overline{\mathrm{co}}\, (dB^n/d\lambda, dC^n/d\lambda)([s, s + 1/m])$$
$$\subset \overline{\mathrm{co}}\, \Theta([t^n + s, t^n + s + 1/m], \omega^n \widetilde{\otimes}_{t_n} X). \tag{8.3}$$

Thanks to Skorokhod's coupling theorem, with a little abuse of notation, there exist random variables

$$(X^0, B^0, C^0), (X^1, B^1, C^1), (X^2, B^2, C^2), \dots,$$

defined on some probability space $(\Sigma, \mathcal{G}, R)$, such that $(X^0, B^0, C^0)$ has distribution $Q$, $(X^n, B^n, C^n)$ has distribution $P^n \circ (X, B^n, C^n)^{-1}$ and $R$-a.s. $(X^n, B^n, C^n) \to (X^0, B^0, C^0)$ in the local uniform topology. Thanks to Condition 2.6, [5, Lemmata 3.2, 3.3] and [1, Theorem 17.23], the correspondence $(s, \omega^o) \mapsto \Theta([s, s + 1/m], \omega^o)$ is continuous. Since, for every $(s, \omega^o) \in \mathbb{R}_+ \times \Omega$, the set $\overline{\mathrm{co}}\, \Theta([s, s + 1/m], \omega^o)$ is compact, we deduce from [1, Theorem 17.35] that the correspondence $(s, \omega^o) \mapsto \overline{\mathrm{co}}\, \Theta([s, s + 1/m], \omega^o)$ is upper hemicontinuous with compact values. Thus, (8.3), [1, Theorem 17.20] and Corollary 7.7 yield that, for all $m \in \mathbb{N}$, $R$-a.s. for all $s \in \mathbb{R}_+$

$$m(B_{s+1/m}^0 - B_s^0, C_{s+1/m}^0 - C_s^0) \in \overline{\mathrm{co}}\, \Theta([s + t, s + 1/m + t], \omega \widetilde{\otimes}_t X^0).$$

Notice that $(\lambda \otimes R)$-a.e.

$$(dB^0/d\lambda, dC^0/d\lambda) = \lim_{m \to \infty} m(B_{\cdot+1/m}^0 - B_\cdot^0, C_{\cdot+1/m}^0 - C_\cdot^0).$$

Now, with (8.2), we get that $R$-a.s. for $\lambda$-a.a. $s \in \mathbb{R}_+$

$$(dB^0/d\lambda, dC^0/d\lambda)(s) \in \bigcap_{m \in \mathbb{N}} \overline{\mathrm{co}}\, \Theta([s + t, s + 1/m + t], \omega \widetilde{\otimes}_t X^0) \subset \Theta(s + t, \omega \widetilde{\otimes}_t X^0).$$

This shows that $(\lambda \otimes Q)$-a.e. $(dY^{(2)}/d\lambda, dY^{(3)}/d\lambda) \in \Theta(\cdot + t, \omega \widetilde{\otimes}_t X) \circ \Phi$.

*Step 4.* In the final step of the proof, we show that $P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}$ and we relate $(Y^{(2)}, Y^{(3)})$ to the $P$-semimartingale characteristics of the coordinate process. Thanks to [41, Lemma 11.1.2], there exists a dense set $D \subset \mathbb{R}_+$ such that $\rho_M \circ \Phi$ is $Q$-a.s. continuous for all $M \in D$. Take some $M \in D$. Since $P^n \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}}$, it follows from the definition of the

first characteristic that the process $X_{\cdot \wedge \rho_M} - B^n_{\cdot \wedge \rho_M}$ is a local $P^n$-$\mathbf{F}_+$-martingale. Furthermore, by the definition of the stopping time $\rho_M$ and the linear growth Condition 2.4, we see that $X_{\cdot \wedge \rho_M} - B^n_{\cdot \wedge \rho_M}$ is $P^n$-a.s. bounded by a constant independent of $n$, which, in particular, implies that it is a true $P^n$-$\mathbf{F}_+$-martingale. Now, it follows from [23, Proposition IX.1.4] that $Y^{(1)}_{\cdot \wedge \rho_M \circ \Phi} - Y^{(2)}_{\cdot \wedge \rho_M \circ \Phi}$ is a $Q$-$\mathbf{F}^*$-martingale. Recalling that $Y^{(2)}$ is $Q$-a.s. locally absolutely continuous by Step 2, this means that $Y^{(1)}$ is a $Q$-$\mathbf{F}^*$-semimartingale with first characteristic $Y^{(2)}$. Similarly, we see that the second characteristic is given by $Y^{(3)}$. Finally, we need to relate these observations to the probability measure $P$ and the filtration $\mathbf{F}_+$. We denote by $A^{p,\Phi^{-1}(\mathbf{F}_+)}$ the dual predictable projection of a process $A$, defined on $(\Omega^*, \mathcal{F}^*)$, to the filtration $\Phi^{-1}(\mathbf{F}_+)$. Recall from [20, Lemma 10.42] that, for every $s \in \mathbb{R}_+$, a random variable $Z$ on $(\Omega^*, \mathcal{F}^*)$ is $\Phi^{-1}(\mathcal{F}_{s+})$-measurable if and only if it is $\mathcal{F}^*_s$-measurable and $Z(\omega^{(1)}, \omega^{(2)}, \omega^{(3)})$ does not depend on $(\omega^{(2)}, \omega^{(3)})$. Thanks to Stricker's theorem (see, e.g., [21, Lemma 2.7]), $Y^{(1)}$ is a $Q$-$\Phi^{-1}(\mathbf{F}_+)$-semimartingale. Notice that each $\rho_M \circ \Phi$ is a $\Phi^{-1}(\mathbf{F}_+)$-stopping time and recall from Step 3 that $(\lambda \otimes Q)$-a.e. $(dY^{(2)}/d\lambda, dY^{(3)}/d\lambda) \in \Theta(\cdot + t, \omega \widetilde{\otimes}_t X) \circ \Phi$. Hence, by definition of $\rho_M$ and the linear growth assumption, for every $M \in D$ and $i, j = 1, \ldots, d$, we have

$$E^Q\big[\operatorname{Var}(Y^{(2,i)})_{\rho_M \circ \Phi}\big] + E^Q\big[\operatorname{Var}(Y^{(3,ij)})_{\rho_M \circ \Phi}\big]$$
$$= E^Q\Big[\int_0^{\rho_M \circ \Phi}\Big(\Big|\frac{dY^{(2,i)}}{d\lambda}\Big| + \Big|\frac{dY^{(3,ij)}}{d\lambda}\Big|\Big)d\lambda\Big] < \infty,$$

where $\operatorname{Var}(\cdot)$ denotes the variation process. By virtue of this, we get from [20, Proposition 9.24] that the $Q$-$\Phi^{-1}(\mathbf{F}_+)$-characteristics of $Y^{(1)}$ are given by $((Y^{(2)})^{p,\Phi^{-1}(\mathbf{F}_+)}, (Y^{(3)})^{p,\Phi^{-1}(\mathbf{F}_+)})$. Hence, thanks to Lemma 6.4, the coordinate process $X$ is a $P$-$\mathbf{F}_+$-semimartingale whose characteristics $(B^P, C^P)$ satisfy $Q$-a.s.

$$(B^P, C^P) \circ \Phi = ((Y^{(2)})^{p,\Phi^{-1}(\mathbf{F}_+)}, (Y^{(3)})^{p,\Phi^{-1}(\mathbf{F}_+)}).$$

Consequently, we deduce from the Steps 2 and 3, and [16, Theorem 5.25], that $P$-a.s. $(B^P, C^P) \ll \lambda$ and

$$(\lambda \otimes P)\big((dB^P/d\lambda, dC^P/d\lambda) \notin \Theta(\cdot + t, \omega \widetilde{\otimes}_t X)\big)$$
$$= (\lambda \otimes Q \circ \Phi^{-1})\big((dB^P/d\lambda, dC^P/d\lambda) \notin \Theta(\cdot + t, \omega \widetilde{\otimes}_t X)\big)$$
$$= (\lambda \otimes Q)\big(E^Q[(dY^{(2)}/d\lambda, dY^{(3)}/d\lambda)|\Phi^{-1}(\mathbf{F}_+)_-] \notin \Theta(\cdot + t, \omega \widetilde{\otimes}_t Y^{(1)})\big) = 0,$$

where we use [9, Corollary 8, p. 48] for the final equality. $\qquad\square$

## 8.2  Proof of Theorem 4.7

To prove Theorem 4.7, we use the sequential characterization of lower hemicontinuity as given by [1, Theorem 17.21]. Namely, we prove that for every sequence $(t^n, \omega^n)_{n \in \mathbb{Z}_+} \subset [0, \infty[$ such that $(t^n, \omega^n) \to (t^0, \omega^0)$ and every $P \in \mathcal{R}(t^0, \omega^0)$, there exists a sequence $P^n \in \mathcal{R}(t^n, \omega^n)$, for $n \in \mathbb{N}$, such that $P^n \to P$ weakly. From now on, we fix $(t^n, \omega^n)_{n \in \mathbb{Z}_+} \subset [0, \infty[$ such that $(t^n, \omega^n) \to (t^0, \omega^0)$ and $P \in \mathcal{R}(t^0, \omega^0)$.

*Step 1: On the structure of $P$.* Recall that $\mathscr{P}$ denotes the predictable $\sigma$-field on $[0, \infty[$ and let $(b^P, a^P)$ be the Lebesgue densities of the $P$-$\mathbf{F}$-characteristics of the coordinate process $X$.

**Lemma 8.1.** *Assume that $F$ is a compact metrizable space and that Condition 2.5 holds. Then, for every $(t^0, \omega^0) \in [0, \infty[$ and every $P \in \mathcal{R}(t^0, \omega^0)$, there exists a predictable map $\mathfrak{f} = \mathfrak{f}(P) \colon [0, \infty[ \to F$ such that, for $(\lambda \otimes P)$-a.e. $(t, \omega) \in [0, \infty[$,*

$$(b^P, a^P)(t, \omega) = (b(\mathfrak{f}(t, \omega), t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega), a(\mathfrak{f}(t, \omega), t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega)),$$

*where $(b^P, a^P)$ denote the Lebesgue densities of the $P$-$\mathbf{F}$-characteristics of the coordinate process $X$.*

*Proof.* By virtue of [6, Theorem 97, p. 147], the maps $(t,\omega) \mapsto b(f, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega)$ and $(t,\omega) \mapsto a(f, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega)$ are predictable for every $f \in F$. Hence, by Lemma 2.9 and Condition 2.5, the correspondence $(t,\omega) \mapsto \Theta^*(t,\omega) := \Theta(t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega)$ has a $\mathscr{P} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}_+)$-measurable graph and consequently,

$$G := \left\{ (t,\omega) \in [\![0,\infty[\![ : (b_t^P(\omega), a_t^P(\omega)) \in \Theta(t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega) \right\}$$
$$= \left\{ (t,\omega) \in [\![0,\infty[\![ : (t, \omega, b_t^P(\omega), a_t^P(\omega)) \in \mathrm{gr}\,\Theta^* \right\} \in \mathscr{P}.$$

For some fixed, but arbitrary, $f_0 \in F$, define

$$\pi(t,\omega) := \begin{cases} (b(f_0, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega), a(f_0, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega)), & \text{if } (t,\omega) \notin G, \\ (b_t^P(\omega), a_t^P(\omega)), & \text{if } (t,\omega) \in G. \end{cases}$$

As the function $(f, t, \omega) \mapsto (b(f, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega), a(f, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega))$ is continuous in the $F$ and $\mathscr{P}$-measurable in the $[\![0,\infty[\![$ variable, we deduce from the measurable implicit function theorem [1, Theorem 18.17] that the correspondence $\gamma \colon [\![0,\infty[\![ \twoheadrightarrow F$ defined by

$$\gamma(t,\omega) := \left\{ f \in F : (b(f, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega), a(f, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega)) = \pi(t,\omega) \right\}$$

is $\mathscr{P}$-measurable and that it admits a measurable selector, i.e., there exists a $\mathscr{P}$-measurable function $\mathfrak{f} \colon [\![0,\infty[\![ \to F$ such that

$$\pi(t,\omega) = (b(\mathfrak{f}(t,\omega), t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega), a(\mathfrak{f}(t,\omega), t + t^0, \omega^0 \widetilde{\otimes}_{t^0} \omega))$$

for all $(t,\omega) \in [\![0,\infty[\![$. Since $P \in \mathcal{R}(t,\omega)$, we have $(\lambda \otimes P)$-a.e. $\pi = (b^P, a^P)$. This completes the proof. $\square$

Let $\mathfrak{f} = \mathfrak{f}(P)$ be as in Lemma 8.1. As Condition 4.6 is assumed in Theorem 4.7, we have a decomposition $a = \sigma\sigma^*$ for an $\mathbb{R}^{d \times r}$-valued function $\sigma$. By a standard integral representation result for continuous local martingales (see, e.g., [19, Theorem II.7.1′]), possibly on an extension of the filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, P)$, there exists an $r$-dimensional standard Brownian motion $W$ such that $P$-a.s.

$$dX_t = b(\mathfrak{f}(t,X), t + t^0, \omega^0 \widetilde{\otimes}_{t^0} X)dt + \sigma(\mathfrak{f}(t,X), t + t^0, \omega^0 \widetilde{\otimes}_{t^0} X)dW_t, \quad X_0 = \omega^0(t^0).$$

For simplicity, we ignore the standard extension in our notation.

*Step 2: A candidate for the approximating sequence.* By Condition 4.6, we obtain from a short computation that, for every $n \in \mathbb{Z}_+$, $T, M > 0$ and every $\alpha^0 \in \Omega$, there exists a constant $C > 0$, that depends on $T, M, \sup_{n \in \mathbb{Z}_+} t^n$ and $\sup_{n \in \mathbb{Z}_+} \sup_{s \in [0,t^n]} \|\omega^n(s)\|$, such that

$$\|b(\mathfrak{f}(t,\alpha^0), t + t^n, \omega^n \widetilde{\otimes}_{t^n} \omega) - b(\mathfrak{f}(t,\alpha^0), t + t^n, \omega^n \widetilde{\otimes}_{t^n} \alpha)\| \leq C \sup_{s \in [0,t]} \|\omega(s) - \alpha(s)\|,$$

$$\|\sigma(\mathfrak{f}(t,\alpha^0), t + t^n, \omega^n \widetilde{\otimes}_{t^n} \omega) - \sigma(\mathfrak{f}(t,\alpha^0), t + t^n, \omega^n \widetilde{\otimes}_{t^n} \alpha)\|_o \leq C \sup_{s \in [0,t]} \|\omega(s) - \alpha(s)\|,$$

$$\tag{8.4}$$

for all $\omega, \alpha \in \Omega$: $\sup_{s \in [0,t]} \|\omega(s)\| \vee \|\alpha(s)\| \leq M$ and $t \in [0,T]$. Furthermore, we deduce from Condition 2.4 that, for every $T > 0$, there exists a constant $C > 0$, that depends on $T, \sup_{n \in \mathbb{Z}_+} t^n$ and $\sup_{n \in \mathbb{Z}_+} \sup_{s \in [0,t^n]} \|\omega^n(s)\|$, such that

$$\|b(\mathfrak{f}(t,\alpha^0), t + t^n, \omega^n \widetilde{\otimes}_{t^n} \omega)\| + \|\sigma(\mathfrak{f}(t,\alpha^0), t + t^n, \omega^n \widetilde{\otimes}_{t^n} \omega)\|_o \leq C\Big(1 + \sup_{s \in [0,t]} \|\omega(s)\|\Big),$$

for all $\omega, \alpha \in \Omega$ and $t \in [0,T]$. Hence, by a standard existence result for stochastic differential equations with random locally Lipschitz coefficients of linear growth (see,

e.g., [20, Theorem 14.30] or [22, Theorem 4.5]), for every $n \in \mathbb{N}$, there exists a continuous adapted process $Y^n$ such that $P$-a.s.

$$dY_t^n = b(\mathfrak{f}(t, X), t + t^n, \omega^n \widetilde{\otimes}_{t^n} Y^n)dt + \sigma(\mathfrak{f}(t, X), t + t^n, \omega^n \widetilde{\otimes}_{t^n} Y^n)dW_t, \quad Y_0^n = \omega^n(t^n).$$

Next, we show that the laws of $\{Y^n : n \in \mathbb{N}\}$ form a candidate for an approximation sequence of $P$.

**Lemma 8.2.** *Suppose that the Conditions 2.4 and 2.7 hold and take $(t^0, \omega^0) \in [\![0, \infty[\![$. On some filtered probability space, let $\zeta$ be an $F$-valued measurable process and let $Y$ be a continuous semimartingale starting at $\omega^0(t^0)$ those semimartingale characteristics are absolutely continuous with densities $(b(\zeta, \cdot + t^0, \omega^0 \widetilde{\otimes}_{t^0} Y), a(\zeta, \cdot + t^0, \omega^0 \widetilde{\otimes}_{t^0} Y))$. Then, the law of $Y$ is an element of $\mathcal{R}(t^0, \omega^0)$.*

*Proof.* Let $P$ be the probability measure of the underlying filtered space and let $Q := P \circ Y^{-1}$ be the law of $Y$. Further, denote the natural right-continuous filtration of $Y$ by $(\mathcal{G}_t)_{t \geq 0}$. Thanks to Stricker's theorem (see, e.g., [21, Lemma 2.7]), $Y$ is a $P$-$(\mathcal{G}_t)_{t \geq 0}$-semimartingale. Furthermore, by virtue of Condition 2.4, we deduce from [20, Proposition 9.24] and [16, Theorem 5.25] that the $(\mathcal{G}_t)_{t \geq 0}$-characteristics of $Y$ are given by

$$\left( \int_0^\cdot E^P [b(\zeta, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} Y)|\mathcal{G}_{t-}]dt, \int_0^\cdot E^P [a(\zeta, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} Y)|\mathcal{G}_{t-}]dt \right).$$

By Lemma 6.4, the coordinate process $X$ is a $Q$-$\mathbf{F}_+$-semimartingale and its characteristics $(B, C)$ are $Q$-a.s. absolutely continuous and such that $P$-a.s. for $\lambda$-a.a. $t \in \mathbb{R}_+$

$$\left( \frac{dB}{d\lambda}, \frac{dC}{d\lambda} \right)(t, Y) = E^P [(b(\zeta, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} Y), a(\zeta, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} Y))|\mathcal{G}_{t-}].$$

Using the convexity assumption given by Condition 2.7 and [9, Corollary 8, p. 48], we obtain that

$$\begin{aligned}
(\lambda \otimes Q)&((dB/d\lambda, dC/d\lambda) \notin \Theta(\cdot + t^0, \omega^0 \widetilde{\otimes}_{t^0} X)) \\
&= \iint \mathbb{1}\{E^P [(b(\zeta, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} Y), \\
&\qquad a(\zeta, t + t^0, \omega^0 \widetilde{\otimes}_{t^0} Y))|\mathcal{G}_{t-}] \notin \Theta(t + t^0, \omega^0 \widetilde{\otimes}_{t^0} Y)\}d(\lambda \otimes P) \\
&= 0.
\end{aligned}$$

This completes the proof. $\qquad\square$

Thanks to Lemma 8.2, for every $n \in \mathbb{N}$, the law of $Y^n$ is an element of $\mathcal{R}(t^n, \omega^n)$. Consequently, the laws of $\{Y^n : n \in \mathbb{N}\}$ are candidates for an approximation sequence of the measure $P$.

*Step 3: ucp convergence of $Y^n$ to $X$:* In this step we prove that $Y^n \to X$ in the ucp topology, i.e., uniformly on compact time sets in probability. Hereby, we use some ideas we learned from the proof of [29, Theorem, p. 88]. Take $T, \varepsilon > 0$ and define

$$\tau_n^m := \inf\{t \geq 0 : \|X_t\| \vee \|Y_t^n\| \geq m\}, \qquad n, m > 0.$$

Thanks to the moment bound from Lemma 7.4, we obtain that

$$P(T_n^m \leq T) = P\left( \sup_{s \in [0, T]} \|X_s\| \vee \|Y_s^n\| \geq m \right) \leq \frac{C}{m},$$

where the constant $C > 0$ is independent of $n$ and $m$. Thus, we can take $M = M_\varepsilon > 0$ large enough such that

$$\sup_{n \in \mathbb{N}} P(T_n^M \leq T) \leq \varepsilon. \tag{8.5}$$

Next, using the Burkholder–Davis–Gundy inequality and Hölder's inequality, and (8.4), we obtain, for every $t \in [0, T]$ and $n \in \mathbb{N}$, that

$$E^P\Big[\sup_{s \in [0, t \wedge T_n^M]} \|X_s - Y_s^n\|^2\Big]$$

$$\leq C\Big(\|\omega^0(t^0) - \omega^n(t^n)\|^2$$

$$+ E^P\Big[\int_0^{t \wedge T_n^M} \big(\|b(\mathfrak{f}(s, X), s + t^0, \omega^0 \widetilde{\otimes}_{t^0} X) - b(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} Y^n)\|^2$$

$$+ \|\sigma(\mathfrak{f}(s, X), s + t^0, \omega^0 \widetilde{\otimes}_{t^0} X) - \sigma(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} Y^n)\|_o^2\big)ds\Big]\Big)$$

$$\leq C\Big(\|\omega^0(t^0) - \omega^n(t^n)\|^2$$

$$+ E^P\Big[\int_0^{t \wedge T_n^M} \big(\|b(\mathfrak{f}(s, X), s + t^0, \omega^0 \widetilde{\otimes}_{t^0} X) - b(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} X)\|^2$$

$$+ \|b(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} X) - b(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} Y^n)\|^2$$

$$+ \|\sigma(\mathfrak{f}(s, X), s + t^0, \omega^0 \widetilde{\otimes}_{t^0} X) - \sigma(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} X)\|_o^2$$

$$+ \|\sigma(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} X) - \sigma(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} Y^n)\|_o^2\big)ds\Big]\Big)$$

$$\leq C\Big(\|\omega^0(t^0) - \omega^n(t^n)\|^2$$

$$+ E^P\Big[\int_0^{t \wedge T_n^M} \big(\|b(\mathfrak{f}(s, X), s + t^0, \omega^0 \widetilde{\otimes}_{t^0} X) - b(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} X)\|^2$$

$$+ \|\sigma(\mathfrak{f}(s, X), s + t^0, \omega^0 \widetilde{\otimes}_{t^0} X) - \sigma(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} X)\|_o^2\big)ds\Big]$$

$$+ E^P\Big[\sup_{s \in [0, t \wedge T_n^M]} \|X_s - Y_s^n\|^2\Big]\Big).$$

Hence, from Gronwall's lemma, we get, for all $n \in \mathbb{N}$, that

$$E^P\Big[\sup_{s \in [0, T \wedge T_n^M]} \|X_s - Y_s^n\|^2\Big]$$

$$\leq C\Big(\|\omega^0(t^0) - \omega^n(t^n)\|^2$$

$$+ E^P\Big[\int_0^{T \wedge T_n^M} \big(\|b(\mathfrak{f}(s, X), s + t^0, \omega^0 \widetilde{\otimes}_{t^0} X) - b(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} X)\|^2$$

$$+ \|\sigma(\mathfrak{f}(s, X), s + t^0, \omega^0 \widetilde{\otimes}_{t^0} X) - \sigma(\mathfrak{f}(s, X), s + t^n, \omega^n \widetilde{\otimes}_{t^n} X)\|_o^2\big)ds\Big]\Big),$$

where the constant $C > 0$ is independent of $n$. Clearly, the first term on the r.h.s. converges to zero as $n \to \infty$. The second term converges to zero by Condition 2.6, Corollary 7.7 and the dominated convergence theorem, which can be applied thanks to the linear growth Condition 2.4 and either the definition of the sequence $(T_n^M)_{n \in \mathbb{N}}$, or the uniform moment bound from Lemma 7.4. We conclude that

$$E^P\Big[\sup_{s \in [0, T \wedge T_n^M]} \|X_s - Y_s^n\|^2\Big] \to 0 \text{ as } n \to \infty. \tag{8.6}$$

We are in the position to complete the proof for ucp convergence. Namely, using (8.5) and (8.6), we get, for every $\delta > 0$, that

$$P\Big(\sup_{s \in [0, T]} \|X_s - Y_s^n\| \geq \delta\Big) \leq P\Big(\sup_{s \in [0, T]} \|X_s - Y_s^n\| \geq \delta, T_n^M > T\Big) + \varepsilon$$

$$\leq \frac{1}{\delta^2} E^P\Big[\sup_{s \in [0, T \wedge T_n^M]} \|X_s - Y_s^n\|^2\Big] + \varepsilon \to \varepsilon$$

as $n \to \infty$. This proves that $Y^n \to X$ in the ucp topology. For continuous processes, ucp convergence implies weak convergence. Hence, the proof of lower hemicontinuity is complete. $\qquad\square$

### 8.3 Proof of Theorem 4.12

Fix $T > 0$ and let $\psi$ be a bounded Lipschitz continuous function. We fix $\omega^0, \alpha^0 \in \Omega$ and $t^0 \in [0, T]$. Take an arbitrary measure $P \in \mathcal{R}(t^0, \omega^0)$. By Lemma 8.1 and some classical representation result for continuous local martingales, there exists a predictable function $\mathfrak{f} = \mathfrak{f}(P) \colon [0, \infty[\, \to F$ such that, possibly on an extension of the filtered probability space $(\Omega, \mathcal{F}, \mathbf{F}, P)$, there exists an $r$-dimensional standard Brownian motion $W$ such that $P$-a.s.

$$dX_t = b(\mathfrak{f}(t, X), t + t^0, \omega^0 \,\widetilde{\otimes}_{t^0} X)dt + \sigma(\mathfrak{f}(t, X), t + t^0, \omega^0 \,\widetilde{\otimes}_{t^0} X)dW_t, \quad X_0 = \omega^0(t^0).$$

Thanks to Condition 4.11, by a standard existence result for stochastic differential equations with random locally Lipschitz coefficients of linear growth, there exists a continuous adapted process $Y$ such that $P$-a.s.

$$dY_t = b(\mathfrak{f}(t, X), t + t^0, \alpha^0 \,\widetilde{\otimes}_{t^0} Y)dt + \sigma(\mathfrak{f}(t, X), t + t^0, \alpha^0 \,\widetilde{\otimes}_{t^0} Y)dW_t, \quad Y_0 = \alpha^0(t^0).$$

Using the Burkholder–Davis–Gundy inequality, Hölder's inequality and the global Lipschitz part from Condition 4.11, we obtain that

$$E^P\Big[\sup_{s \in [0,T]} \|X_s - Y_s\|^2\Big]$$
$$\leq C\Big(\|\omega(t^0) - \alpha(t^0)\|^2$$
$$\qquad + \int_0^T E^P\big[\|b(\mathfrak{f}(t, X), t + t^0, \omega^0 \,\widetilde{\otimes}_{t^0} X) - b(\mathfrak{f}(t, X), t + t^0, \alpha^0 \,\widetilde{\otimes}_{t^0} Y)\|^2\big]ds$$
$$\qquad + \int_0^T E^P\big[\|\sigma(\mathfrak{f}(t, X), t + t^0, \omega^0 \,\widetilde{\otimes}_{t^0} X) - \sigma(\mathfrak{f}(t, X), t + t^0, \alpha^0 \,\widetilde{\otimes}_{t^0} Y)\|_o^2\big]ds\Big)$$
$$\leq C\Big(\sup_{s \in [0,t^0]} \|\omega^0(s) - \alpha^0(s)\|^2 + \int_0^T E^P\Big[\sup_{s \in [0,t]} \|X_s - Y_s\|^2\Big]ds\Big).$$

Now, Gronwall's lemma yields that

$$E^P\Big[\sup_{s \in [0,T]} \|X_s - Y_s\|^2\Big] \leq C \sup_{s \in [0,t^0]} \|\omega^0(s) - \alpha^0(s)\|^2,$$

and, by Jensen's inequality,

$$E^P\Big[\sup_{s \in [0,T]} \|X_s - Y_s\|\Big] \leq C \sup_{s \in [0,t^0]} \|\omega^0(s) - \alpha^0(s)\|. \tag{8.7}$$

Thanks to Lemma 8.2, we have $P \circ Y^{-1} \in \mathcal{R}(t^0, \alpha^0)$. Thus, in case $v(t^0, \omega^0) \geq v(t^0, \alpha^0)$, we get, from Corollary 7.3, the Lipschitz continuity of $\psi$ and (8.7), that

$$v(t^0, \omega^0) - v(t^0, \alpha^0)$$
$$\leq \sup_{P \in \mathcal{R}(t^0, \omega^0)} \Big(E^P\big[|\psi(\omega^0 \,\widetilde{\otimes}_{t^0} X) - \psi(\alpha^0 \,\widetilde{\otimes}_{t^0} Y)|\big] + E^P\big[\psi(\alpha^0 \,\widetilde{\otimes}_{t^0} Y)\big]\Big) - v(t^0, \alpha^0)$$
$$\leq C\Big(\sup_{s \in [0,t^0]} \|\omega^0(s) - \alpha^0(s)\| + \sup_{P \in \mathcal{R}(t^0, \omega)} E^P\Big[\sup_{s \in [0,T]} \|X_s - Y_s\|\Big]\Big) + v(t^0, \alpha^0) - v(t^0, \alpha^0)$$
$$\leq C \sup_{s \in [0,t^0]} \|\omega^0(s) - \alpha^0(s)\|.$$

By symmetry, this inequality yields that

$$|v(t^0, \omega^0) - v(t^0, \alpha^0)| \leq C \sup_{s \in [0, t^0]} \|\omega^0(s) - \alpha^0(s)\|. \tag{8.8}$$

Next, let $0 \leq s_0 \leq t_0$. Using the DPP as given by Theorem 3.1 for the first line and (8.8) for the third line, we obtain that

$$
\begin{aligned}
\left| v(s^0, \omega^0) - v(t^0, \omega^0) \right| &= \left| \sup_{P \in \mathcal{C}(s^0, \omega^0)} E^P\left[ v(t^0, X) \right] - v(t^0, \omega^0) \right| \\
&\leq \sup_{P \in \mathcal{C}(s^0, \omega^0)} E^P\left[ |v(t^0, X) - v(t^0, \omega^0)| \right] \\
&\leq \sup_{P \in \mathcal{C}(s^0, \omega^0)} E^P\left[ \sup_{s \in [0, t^0]} \|X_s - \omega^0(s)\| \right] \\
&= \sup_{P \in \mathcal{C}(s^0, \omega^0)} E^P\left[ \sup_{s \in [s^0, t^0]} \|X_s - \omega^0(s)\| \right] \\
&\leq \sup_{P \in \mathcal{C}(s^0, \omega^0)} E^P\left[ \sup_{s \in [s^0, t^0]} \|X_s - \omega^0(s^0)\| \right] \\
&\qquad\qquad\qquad + \sup_{s \in [s^0, t^0]} \|\omega^0(s) - \omega^0(s^0)\|.
\end{aligned}
\tag{8.9}
$$

Using Condition 2.4, it follows as in the solution to [24, Problem 3.3.15] that

$$\sup_{P \in \mathcal{C}(s^0, \omega^0)} E^P\left[ \sup_{s \in [s^0, t^0]} \|X_s - \omega^0(s^0)\| \right] \leq C\left(1 + \sup_{s \in [0, s^0]} \|\omega^0(s)\|\right) |t^0 - s^0|^{1/2}. \tag{8.10}$$

Finally, using (8.8), (8.9) and (8.10), we obtain that

$$
\begin{aligned}
|v(t^0, &\omega^0) - v(s^0, \alpha^0)| \\
&\leq |v(t^0, \omega^0) - v(s^0, \omega^0)| + |v(s^0, \omega^0) - v(s^0, \alpha^0)| \\
&\leq C\Big[\Big(1 + \sup_{s \in [0, s^0]} \|\omega^0(s)\|\Big)|t^0 - s^0|^{1/2} \\
&\qquad\qquad + \sup_{s \in [s^0, t^0]} \|\omega^0(s) - \omega^0(s^0)\| + \sup_{s \in [0, s^0]} \|\omega^0(s) - \alpha^0(s)\|\Big] \\
&\leq C\Big[\Big(1 + \sup_{s \in [0, s^0]} \|\omega^0(s)\|\Big)|t^0 - s^0|^{1/2} \\
&\qquad\qquad + \sup_{s \in [s^0, t^0]} \|\omega^0(s) - \alpha^0(s^0)\| + 2 \sup_{s \in [0, s^0]} \|\omega^0(s) - \alpha^0(s)\|\Big] \\
&\leq C\Big[\Big(1 + \sup_{s \in [0, s^0]} \|\omega^0(s)\|\Big)|t^0 - s^0|^{1/2} + \sup_{s \in [0, T]} \|\omega^0(s \wedge t^0) - \alpha^0(s \wedge s^0)\|\Big] \\
&\leq C\mathsf{d}_T((t^0, \omega^0), (s^0, \alpha^0)).
\end{aligned}
$$

The proof is complete. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \square$

## 9 Proof of the nonlinear martingale problem: Theorem 5.2

In this section we work under the assumptions of Theorem 5.2, i.e., we assume that $d = 1$, that $F = F_0 \times F_1$ for two compact metrizable spaces $F_0$ and $F_1$, that $b$ depends on $F$ only through the $F_0$ variable, that $a$ depends on $F$ only through the $F_1$ variable, and that the Conditions 2.4 and 2.6 hold.

**Lemma 9.1.** *For every $(t, \omega) \in [\![0, \infty[\![$, there exists a $\overline{P} \in \mathcal{C}(t, \omega)$ such that $(\lambda \otimes \overline{P})$-a.e.*

$$b^{\overline{P}}_{\cdot + t} = \sup\left\{ b(f_0, \cdot + t, \omega \otimes_t X) \colon f_0 \in F_0 \right\}, \qquad a^{\overline{P}}_{\cdot + t} = \sup\left\{ a(f_1, \cdot + t, \omega \otimes_t X) \colon f_1 \in F_1 \right\},$$

*where $(b^{\overline{P}}_{\cdot + t}, a^{\overline{P}}_{\cdot + t})$ denote the Lebesgue densities of the $\overline{P}$-characteristics of $X_{\cdot + t}$ (for its right-continuous natural filtration).*

*Proof.* Thanks to the Conditions 2.4 and 2.6, Corollary 7.7 and Berge's maximum theorem ([1, Theorem 17.31]), the functions

$$(s, \alpha) \mapsto \sup\{b(f_0, s + t, \omega \otimes_t \alpha) \colon f_0 \in F_0\}, \quad (s, \alpha) \mapsto \sup\{a(f_1, s + t, \omega \otimes_t \alpha) \colon f_1 \in F_1\}$$

are continuous and of linear growth (in the sense of (4) on p. 258 in [13]). Hence, the claim follows from [13, Theorem 4, p. 265] and Lemma 6.4. □

*Proof of Theorem 5.2.* Fix $\omega \in \Omega, n \in \mathbb{N}, \phi \in \mathbb{M}_{icx}$ and $s, h \geq 0$. Take $P \in \mathcal{C}(s, \omega)$, denote the Lebesgue densities of the $P$-characteristics of $X_{\cdot+s}$ by $(b^P_{\cdot+s}, a^P_{\cdot+s})$ and denote the local $P$-martingale part of $X_{\cdot+s}$ by $M^P_{\cdot+s}$. Itô's formula yields that

$$\phi(X_{s+h}) - \phi(X_s) = \int_0^h \phi'(X_{r+s}) dM^P_{r+s} + \int_s^{s+h} \left(\phi'(X_r)b^P_r + \frac{1}{2}\phi''(X_r)a^P_r\right) dr.$$

Hence, using the definition of $\rho_n = \inf\{r \geq 0 \colon |X_r| \geq n\}$ and Condition 2.4, the stopped local martingale part is a true martingale and we obtain that

$$E^P\left[\phi(X_{(s+h)\wedge\rho_n}) - \phi(X_{s\wedge\rho_n}) - \int_{s\wedge\rho_n}^{(s+h)\wedge\rho_n} G(r, X, \phi) dr\right]$$
$$= E^P\left[\int_{s\wedge\rho_n}^{(s+h)\wedge\rho_n} \left(\left(\phi'(X_r)b^P_r + \frac{1}{2}\phi''(X_r)a^P_r\right) - G(r, X, \phi)\right) dr\right].$$

By definition of $G$, we clearly have

$$E^P\left[\int_{s\wedge\rho_n}^{(s+h)\wedge\rho_n} \left(\left(\phi'(X_r)b^P_r + \frac{1}{2}\phi''(X_r)a^P_r\right) - G(r, X, \phi)\right) dr\right] \leq 0,$$

which implies

$$\mathcal{E}_s\left(\phi(X_{(s+h)\wedge\rho_n}) - \phi(X_{s\wedge\rho_n}) - \int_{s\wedge\rho_n}^{(s+h)\wedge\rho_n} G(r, X, \phi) dr\right)(\omega) \leq 0.$$

Let us now prove the converse inequality. By Lemma 9.1, there exists a probability measure $\overline{P} \in \mathcal{C}(s, \omega)$ such that $(\lambda \otimes \overline{P})$-a.e.

$$b^{\overline{P}}_{\cdot+s} = \sup\{b(f_0, \cdot + s, \omega \otimes_s X) \colon f_0 \in F_0\}, \qquad a^{\overline{P}}_{\cdot+s} = \sup\{a(f_1, \cdot + s, \omega \otimes_s X) \colon f_1 \in F_1\}.$$

Now, as $\phi', \phi'' \geq 0$, we get

$$E^{\overline{P}}\left[\int_{s\wedge\rho_n}^{(s+h)\wedge\rho_n} \left(\left(\phi'(X_r)b^{\overline{P}}_r + \frac{1}{2}\phi''(X_r)a^{\overline{P}}_r\right) - G(r, X, \phi)\right) dr\right]$$
$$= E^{\overline{P}}\left[\int_{s\wedge\rho_n}^{(s+h)\wedge\rho_n} \left(G(r, \omega \otimes_t X, \phi) - G(r, X, \phi)\right) dr\right] = 0,$$

which implies

$$\mathcal{E}_s\left(\phi(X_{(s+h)\wedge\rho_n}) - \phi(X_{s\wedge\rho_n}) - \int_{s\wedge\rho_n}^{(s+h)\wedge\rho_n} G(r, X, \phi) dr\right)(\omega) \geq 0.$$

The proof is complete. □

# A  Continuity of $\mathcal{R}$ in the Lévy case

This appendix is dedicated to the fact that the correspondence $(t, \omega) \mapsto \mathcal{R}(t, \omega)$ is continuous if the uncertainty set $\Theta(t, \omega) \equiv \Theta$ is independent of $(t, \omega)$ and convex and compact. This result is covered by our more general Theorems 4.4 and 4.7. The purpose of this section is to explain a substantially simpler proof for this Lévy setting.

**Proposition A.1.** *Let $\Theta \subset \mathbb{R}^d \times \mathbb{S}_+^d$ be a convex and compact set. Then, the correspondence*

$$(t, \omega) \mapsto \mathcal{R}(t, \omega) = \big\{ P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}} \colon P \circ X_0^{-1} = \delta_{\omega(t)}, \ (\lambdabar \otimes P)\text{-a.e. } (dB^P/d\lambdabar, dC^P/d\lambdabar) \in \Theta \big\}$$

*is continuous.*

*Proof.* Define

$$\mathfrak{P}(\Theta) := \big\{ P \in \mathfrak{P}_{\mathrm{sem}}^{\mathrm{ac}} \colon P \circ X_0^{-1} \in \{\delta_x \colon x \in \mathbb{R}^d\}, \ (\lambdabar \otimes P)\text{-a.e. } (dB^P/d\lambdabar, dC^P/d\lambdabar) \in \Theta \big\}.$$

By Theorem 4.4 and [1, Theorems 15.8, 17.20], the set $\mathfrak{P}(\Theta)$ is closed, cf. also [5, Proposition 3.8] and [28, Proposition 4.4]. Furthermore, the function

$$\Omega \times [\![0, \infty[\![ \ni (\omega', t, \omega) \mapsto \omega(t) + \omega' - \omega'(0) = \omega(t) \otimes_0 \omega'$$

is continuous, and hence, by [2, Theorem 8.10.61], so is the map

$$[\![0, \infty[\![ \times \mathfrak{P}(\Omega) \ni (t, \omega, P) \mapsto P \circ (\omega(t) \otimes_0 X)^{-1}.$$

We claim that

$$\mathcal{R}(t, \omega) = \big\{ P \circ (\omega(t) \otimes_0 X)^{-1} \colon P \in \mathfrak{P}(\Theta) \big\}.$$

Indeed, Lemmata 6.3 and 6.4 show that the map $P \mapsto P \circ (\omega(t) \otimes_0 X)^{-1}$ leaves $\mathfrak{P}(\Theta)$ invariant. We stress that this part of the argument is special for the Lévy case. Providing an intuition, it corresponds to the fact that Lévy processes have an additive structure in their initial values, i.e., if $L^x$ is a Lévy process with starting value $x$, then $L^x - x$ is a Lévy process (with the same Lévy–Khinchine triplet) starting at zero.

Finally, we are in the position to prove that $(t, \omega) \mapsto \mathcal{R}(t, \omega)$ is continuous. By virtue of [1, Theorem 17.23], $\mathcal{R}$ is lower hemicontinuous, being the composition of the lower hemicontinuous correspondence $(t, \omega) \mapsto \varphi(t, \omega) := \{(t, \omega)\} \times \mathfrak{P}(\Theta)$ and the (single-valued) continuous correspondence $(t, \omega, P) \mapsto \{P \circ (\omega(t) \otimes_0 X)^{-1}\}$. Here, lower hemicontinuity of $\varphi$ follows from [1, Theorem 17.28]. We would like to point out that this part of the argument hinges on the fact that $\Theta$ is constant. For the upper hemicontinuity of $\mathcal{R}$ we need to argue separately. Let $F \subset \mathfrak{P}(\Omega)$ be closed. We need to show that the lower inverse $\mathcal{R}^l(F) = \{(t, \omega) \colon \mathcal{R}((t, \omega)) \cap F \neq \varnothing\}$ is closed. Assume that $(t^n, \omega^n)_{n \in \mathbb{N}} \subset \mathcal{R}^l(F)$ converges to $(t, \omega) \in [\![0, \infty[\![$. For each $n \in \mathbb{N}$, there exists a probability measure $P^n \in \mathcal{R}(t^n, \omega^n) \cap F$. As $\{\omega^n(t^n) \colon n \in \mathbb{N}\} \subset \mathbb{R}^d$ is bounded, the set $\{P^n \colon n \in \mathbb{N}\}$ is relatively compact by Lemma 7.4. Hence, passing to a subsequence if necessary, we can assume that $P^n \to P$ weakly for some $P \in \mathrm{cl}(\mathfrak{P}(\Theta) \cap F) = \mathfrak{P}(\Theta) \cap F$. As $P \circ X_0^{-1} = \lim_{n \to \infty} P^n \circ X_0^{-1} = \lim_{n \to \infty} \delta_{\omega^n(t^n)} = \delta_{\omega(t)}$, we conclude that $P \in \mathcal{R}(t, \omega) \cap F$, which implies $(t, \omega) \in \mathcal{R}^l(F)$. $\qquad\square$

# References

[1] C. D. Aliprantis and K. B. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide.* Springer Berlin Heidelberg, 3rd ed., 2006. MR2378491

[2] V. I. Bogachev. *Measure Theory.* Springer Berlin Heidelberg, 2007. MR2267655

[3] A. Cosso and F. Russo. Crandall–Lions viscosity solutions for path-dependent PDEs: The case of heat equation. *Bernoulli*, 28(1):481–503, 2022. MR4337713

[4] A. Cosso, F. Gozzi, M. Rosestolato and F. Russo. Path-dependent Hamilton-Jacobi-Bellman equation: Uniqueness of Crandall-Lions viscosity solutions. arXiv:2107.05959v2, 2022. MR4337713

[5] D. Criens and L. Niemann. Markov selections and Feller properties of nonlinear diffusions. arXiv:2205.15200v4, 2022.

[6] C. Dellacherie and P. A. Meyer. *Probability and Potential*. North-Holland Publishing Company – Amsterdam, New York, Oxford, 1978. MR0521810

[7] R. Denk, M. Kupper, and M. Nendel. A semigroup approach to nonlinear Lévy processes. *Stochastic Processes and their Applications*, 130:1616–1642, 2020. MR4058284

[8] M. Denuit and A. Müller. Smooth generators of integral stochastic orders. *The Annals of Applied Probability*, 12(4):1174–1184, 2002. MR1936588

[9] J. Diestel and J. J. Uhl, Jr. *Vector Measures*. American Mathematical Society, 1977. MR0453964

[10] N. El Karoui, D. Nguyen and M. Jeanblanc-Picqué. Compactification methods in the control of degenerate diffusions: existence of an optimal control. *Stochastics*, 20(3):169–219, 1987. MR0878312

[11] N. El Karoui and X. Tan. Capacities, measurable selection and dynamic programming, part II: Application in stochastic control problems. arXiv:1310.3364v2, 2015.

[12] T. Fadina, A. Neufeld, and T. Schmidt. Affine processes under parameter uncertainty. *Probability, Uncertainty and Quantitative Risk*, 4(5), 2019. MR3955321

[13] I. I. Gikhman and A. V. Skorokhod. *The Theory of Stochastic Processes III*. Springer Berlin Heidelberg, reprint of the 1974 ed., 2007. MR2304859

[14] F. Gozzi, C. Marinelli and S. Savin. On controlled linear diffusions with delay in a model of optimal advertising under uncertainty with memory effects. *Journal of Optimization Theory and Applications*, 142:291–321, 2009. MR2525792

[15] X. Guo, C. Pan and S. Peng. Martingale problem under nonlinear expectations. *Mathematics and Financial Economics*, 12:135–164, 2018. MR3766055

[16] S.-W. He, J.-G. Wang and J.-A- Yan. *Semimartingale Theory and Stochastic Calculus*. Science Press and CRC Press, Inc., 1992. MR1219534

[17] J. Hollender. *Lévy-Type Processes under Uncertainty and Related Nonlocal Equations*. PhD thesis, TU Dresden, 2016.

[18] M. Hu and S. Peng. G-Lévy processes under sublinear expectations. *Probability, Uncertainty and Quantitative Risk*, 6(1), 2021. MR4251961

[19] N. Ikeda and S. Watanabe. *Stochastic Differential Equatrions and Diffusion Processes*. North-Holland Publishing Company Amsterdam Oxford New York, 2nd ed., 1989. MR1011252

[20] J. Jacod. *Calcul stochastique et problèmes de martingales*. Springer Berlin Heidelberg New York, 1979. MR0542115

[21] J. Jacod. Weak and strong solutions of stochastic differential equations. *Stochastics*, 3:171–191, 1980. MR0573202

[22] J. Jacod and J. Mémin. Weak and strong solutions of stochastic differential equations: Existence and stability. In D. Williams, editor, *Stochastic Integrals*, volume 851 of *Lecture Notes in Mathematics*, pages 169–212. Springer Berlin Heidelberg, 1981. MR0620991

[23] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer Berlin Heidelberg, 2nd ed., 2003. MR1943877

[24] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer New York, 2nd ed., 1991. MR1121940

[25] A. Klenke. *Probability Theory*. Springer London, 2nd ed., 2014. MR3112259

[26] F. Kühn. Viscosity solutions to Hamilton–Jacobi–Bellman equations associated with sublinear Lévy(-type) processes. *ALEA Latin American Journal of Probability and Mathematical Statistics*, 16:531–559, 2019. MR3941868

[27] F. Kühn. On infinitesimal generators of sublinear Markov semigroups. *Osaka Journal of Mathematics*, 58(3):487–508, 2021. MR4350042

[28] C. Liu and A. Neufeld. Compactness criterion for semimartingale laws and semimartingale optimal transport. *Transactions of the American Mathematical Society*, 372(1):187–231, 2019. MR3968767

[29] M. Metivier and J. Pellaumail. *Stochastic Integration*. Academic Press, 1980. MR0578177

[30] A. Müller. Stochastic orders generated by integrals: a unified study. *Advances in Applied Probability*, 29(2):414–428, 1997. MR1450937

[31] A. Neufeld and M. Nutz. Measurability of semimartingale characteristics with respect to the probability law. *Stochastic Processes and their Applications*, 124:3819–3845, 2014. MR3249357

[32] A. Neufeld and M. Nutz. Nonlinear Lévy processes and their characteristics. *Transactions of the American Mathematical Society*, 369:69–95, 2017. MR3557768

[33] M. Nutz. Random G-expectations. *The Annals of Applied Probability*, 23(5):1755–1777, 2013. MR3114916

[34] M. Nutz and R. van Handel. Constructing sublinear expectations on path space. *Stochastic Processes and their Applications*, 123(8):3100–3121, 2013. MR3062438

[35] S. G. Peng. G-expectation, G-Brownian motion and related stochastic calculus of Itô type. In F. E. Benth et al., editors, *Stochastic Analysis and Applications: The Abel Symposium 2005*, pages 541–567, Springer Berlin Heidelberg, 2007. MR2397805

[36] S. G. Peng. Multi-dimensional G-Brownian motion and related stochastic calculus under G-expectation. *Stochastic Processes and their Applications*, 118:2223–2253, 2008. MR2474349

[37] S. G. Peng. Nonlinear expectations and stochastic calculus under uncertainty. arXiv:1002.4546, 2010. MR3970247

[38] D. Pollard. *Convergence of Stochastic Processes*. Springer New York, 1984. MR0762984

[39] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer Berlin Heidelberg, 2nd ed., 2004.

[40] D. W. Stroock. *Probability Theory: An Analytic View*. Cambridge University Press, 2nd ed., 2011. MR2760872

[41] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*. Springer Berlin Heidelberg, reprint of 1997 ed., 2006. MR2190038

[42] S. Tang and F. Zhang. Path-dependent optimal stochastic control and viscosity solution of associated Bellman equations. *Discrete and Continuous Dynamical Systems*, 35(11):5521-5553, 2015. MR3392684

[43] J. Zhou. Viscosity solutions to second order path-dependent Hamilton-Jacobi-Bellman equations and applications. arXiv:2005.05309v3, 2022.