

STOCHASTIC CONTROL FOR A CLASS OF NONLINEAR KERNELS AND APPLICATIONS¹

BY DYLAN POSSAMAÏ^{*,2}, XIAOLU TAN^{*,3} AND CHAO ZHOU^{†,4}

University of Paris-Dauphine, PSL Research University and National University of Singapore[†]*

We consider a stochastic control problem for a class of nonlinear kernels. More precisely, our problem of interest consists in the optimization, over a set of possibly nondominated probability measures, of solutions of backward stochastic differential equations (BSDEs). Since BSDEs are nonlinear generalizations of the traditional (linear) expectations, this problem can be understood as stochastic control of a family of nonlinear expectations, or equivalently of nonlinear kernels. Our first main contribution is to prove a dynamic programming principle for this control problem in an abstract setting, which we then use to provide a semimartingale characterization of the value function. We next explore several applications of our results. We first obtain a wellposedness result for second order BSDEs (as introduced in Soner, Touzi and Zhang [*Probab. Theory Related Fields* **153** (2012) 149–190]) which does not require any regularity assumption on the terminal condition and the generator. Then we prove a nonlinear optional decomposition in a robust setting, extending recent results of Nutz [*Stochastic Process. Appl.* **125** (2015) 4543–4555], which we then use to obtain a super-hedging duality in uncertain, incomplete and nonlinear financial markets. Finally, we relate, under additional regularity assumptions, the value function to a viscosity solution of an appropriate path-dependent partial differential equation (PPDE).

CONTENTS

1. Introduction	552
2. Stochastic control for a class of nonlinear stochastic kernels	556
2.1. Probabilistic framework	556
2.1.1. Canonical space	556

Received January 2016; revised January 2017.

¹This work was started while the first and second authors were visiting the National University of Singapore, whose hospitality is kindly acknowledged.

²Supported by the ANR Pacman ANR-16-CE05-0027.

³Supported in part by the ERC 321111 Rofirm, the ANR Isotace, the Chairs Financial Risks (Risk Foundation, sponsored by Société Générale) and Finance and Sustainable Development (IEF sponsored by EDF and CA).

⁴Supported by NUS Grant R-146-000-179-133 and Singapore MOE AcRF Grant R-146-000-219-112.

MSC2010 subject classifications. 60H10, 60H30.

Key words and phrases. Stochastic control, measurable selection, nonlinear kernels, second-order BSDEs, path-dependent PDEs, robust super-hedging.

2.1.2. Semimartingale measures	557
2.1.3. Conditioning and concatenation of probability measures	558
2.1.4. Hypotheses	559
2.2. Spaces and norms	560
2.3. Control on a class of nonlinear stochastic kernels and the dynamic programming principle	561
2.4. Proof of Theorem 2.1	563
2.4.1. An equivalent formulation on enlarged canonical space	563
2.4.2. An iterative construction of the solution to BSDE (2.5)	565
2.4.3. On the measurability issues of the iteration	567
2.4.4. End of the Proof of Theorem 2.1	569
2.4.5. Further discussions	571
3. Path regularization of the value function	572
4. Application to 2BSDEs	582
4.1. Definition	582
4.2. Uniqueness, stochastic control representation and comparison	583
4.3. A priori estimates and stability	587
4.4. Existence through dynamic programming	589
5. Nonlinear optional decomposition and super-hedging duality	590
5.1. Saturated 2BSDEs	590
5.2. A super-hedging duality in uncertain, incomplete and nonlinear markets	591
6. Path-dependent PDEs	594
Appendix	596
A.1. Technical results for BSDEs	597
References	599

1. Introduction. The dynamic programming principle (DPP for short) has been a major tool in the control theory, since the latter took off in the 1970s. Informally speaking, this principle simply states that a global optimization problem can be split into a series of local optimization problems. Although such a principle is extremely intuitive, its rigorous justification has proved to be a surprisingly difficult issue. Hence, for stochastic control problems, the dynamic programming principle is generally based on the stability of the controls with respect to conditioning and concatenation, together with a measurable selection argument, which roughly speaking, allow to prove the measurability of the associated value function, as well as constructing almost optimal controls through “pasting”. This is exactly the approach followed by Bertsekas and Shreve [6], and Dellacherie [24] for discrete time stochastic control problems. In continuous time, a comprehensive study of the dynamic programming principle remained more elusive. Thus, El Karoui, in [33], established the dynamic programming principle for the optimal stopping problem in a continuous time setting, using crucially the strong stability properties of stopping times, as well as the fact that the measurable selection argument can be avoided in this context, since an essential supremum over stopping times can be approximated by a supremum over a countable family of random variables. Later, for general controlled Markov processes (in continuous time) problems, El Karoui, Huu Nguyen and Jeanblanc [35] provided a framework to derive

the dynamic programming principle using the measurable selection theorem, by interpreting the controls as probability measures on the canonical trajectory space (see, e.g., Theorems 6.2, 6.3 and 6.4 of [35]). Another commonly used approach to derive the DPP was to bypass the measurable selection argument by proving, under additional assumptions, *a priori* regularity of the value function. This was the strategy adopted, among others, by Fleming and Soner [42], and in the so-called weak DPP of Bouchard and Touzi [14], which has then been extended by Bouchard and Nutz [9, 11] and Bouchard, Moreau and Nutz [8] to optimal control problems with state constraints as well as to differential games (see also Dumitrescu, Quenez and Sulem [29] for a combined stopping-control problem on BSDEs). One of the main motivations of this weak DPP is that it is generally enough to characterize the value function as a viscosity solution of the associated Hamilton–Jacobi–Bellman partial differential equation (PDE). Let us also mention the so-called stochastic Perron’s method, which has been developed by Bayraktar and Sîrbu (see, e.g., [4, 5]), which allows for Markov problems to obtain the viscosity solution characterization of the value function without using the DPP, and then to prove the latter *a posteriori*. Recently, motivated by the emerging theory of robust finance, Nutz et al. [65, 71] gave a framework, which allowed to prove the dynamic programming principle for sub-linear expectations (or equivalently a non-Markovian stochastic control problem), where the essential arguments are close to those in [35], though the presentation is more modern, pedagogic and accessible. The problem in continuous-time has also been studied by El Karoui and Tan [40, 41], in a more general context than the previous references, but still based on the same arguments as in [35] and [65].

However, all the above works consider only what needs to be qualified as the sublinear case. Indeed, the control problems considered consists generically in the maximization of a family of expectations over the set of controls. Nonetheless, so-called nonlinear expectations on a given probability space (that is to say operators acting on random variables which preserve all the properties of expectations but linearity) have now a long history, be it from the capacity theory, used in economics to axiomatize preferences of economic agents which do not satisfy the usual axiom’s of von Neumann and Morgenstern, or from the seminal g -expectations (or BSDEs) introduced by Peng [74]. Before pursuing, let us just recall that in the simple setting of a probability space carrying a Brownian motion W , with its (completed) natural filtration \mathbb{F} , finding the solution of a BSDE with generator g and terminal condition $\xi \in \mathcal{F}_T$ amounts to finding a pair of \mathbb{F} -progressively measurable processes (Y, Z) such that

$$Y_t = \xi - \int_t^T g_s(Y_s, Z_s) ds - \int_t^T Z_s \cdot dW_s, \quad t \in [0, T], \text{ a.s.}$$

This theory is particularly attractive from the point of view of stochastic control, since it is constructed to be filtration (or time) consistent, that is to say that its conditional version satisfies a tower property similar to that of linear expectations, which is itself a kind of dynamic programming principle. Furthermore, it has been

proved by Coquet *et al.* [18] that essentially all filtration consistent nonlinear expectations satisfying appropriate domination properties could be represented with BSDEs (we refer the reader to [49] and [17] for more recent extensions of this result). Our first contribution in this paper, in Section 2, is therefore to generalize the measurable selection argument to derive the dynamic programming principle in the context of optimal stochastic control of nonlinear expectations (or kernels), which can be represented by BSDEs (which as mentioned above is not such a stringent assumption). We emphasize that such an extension is certainly not straightforward. Indeed, in the context of linear expectations, there is a very well established theory studying how the measurability properties of a given map are impacted by its integration with respect to a so-called stochastic kernel (roughly speaking, one can see this as a regular version of a conditional expectation in our context; see, for instance, [6], Chapter 7). For instance, integrating a Borel map with respect to a Borel stochastic kernel preserves the Borel measurability. However, in the context of BSDEs, one has to integrate with respect to nonlinear stochastic kernels, for which, as far as we know, no such theory of measurability exists. Moreover, we also obtain a semimartingale decomposition for the value function of our control problem. This is the objective of Section 3.

Let us now explain where our motivation for studying this problem stems from. The problem of studying a controlled system of BSDEs is not new. For instance, it was shown by El Karoui and Quenez [38, 39] and Hamadène and Lepeltier [46] (see also [37] and the references therein) that a stochastic control problem with control on the drift only could be represented via a controlled family of BSDEs (which can actually be itself represented by a unique BSDE with convex generator). More recently, motivated by obtaining probabilistic representations for fully nonlinear PDEs, Soner, Touzi and Zhang [84, 85] (see also the earlier works [15] and [83]) introduced a notion of second-order BSDEs (2BSDEs for short), whose solutions could actually be written as a supremum, over a family of nondominated probability measures (unlike in [39] where the family is dominated), of standard BSDEs. Therefore, the 2BSDEs fall precisely in the class of problem that we want to study, that is, stochastic control of nonlinear kernels. The authors of [84, 85] managed to obtain the dynamic programming principle, but under very strong continuity assumptions w.r.t. ω on the terminal condition and the generator of the BSDEs, and obtained a semimartingale decomposition of the value function of the corresponding stochastic control problem, which ensured well-posedness of the associated 2BSDE. Again, these regularity assumptions are made to obtain the continuity of the value function *a priori*, which allows to avoid completely the use of the measurable selection theorem. Since then, the 2BSDE theory has been extended by allowing more general generators, filtrations and constraints (see [52, 53, 62, 63, 77, 79]), but no progress has been made concerning the regularity assumptions. However, the 2BSDEs (see, for instance, [64]) have proved to provide a particularly nice framework to study the so-called robust problems in finance,

which were introduced by [2, 60] and in a more rigorous setting by [26]. However, the regularity assumptions put strong limitations to the range of the potential applications of the theory.

We also would like to mention a related theory introduced by Peng [76], and developed around the notion of G -expectations, which lead to the so-called G -BSDEs (see [47, 48]). Instead of working on a fixed probability space carrying different probability measures corresponding to the controls, they work directly on a so-called sublinear expectation space in which the canonical process already incorporates the different measures, without having to refer to a probabilistic setting. Although their method of proof is different, since they mainly use PDE arguments to construct a solution in the Markovian case and then a closure argument, the final objects are extremely close to 2BSDEs, with similar restrictions in terms of regularity. Moreover, the PDE approach they use is unlikely to be compatible with a theory without any regularity, since the PDEs they consider need at the very least to have a continuous solution. On the other hand, there is more hope for the probabilistic approach of the 2BSDEs, since, as shown in [71] in the case of linear expectations (i.e., when the generator of the BSDEs is 0), everything can be well defined by assuming only that the terminal condition is (Borel) measurable.

There is a third theory which shares deep links with 2BSDEs, namely that of viscosity solutions of fully nonlinear path dependent PDEs (PPDEs for short), which has been introduced recently by Ekren, Keller, Touzi and Zhang [30–32]. Indeed, they showed that the solution of a 2BSDE, with a generator and a terminal condition uniformly continuous (in ω), was nothing else than the viscosity solution of a particular PPDE, making the previous theory of 2BSDEs a special case of the theory of PPDEs. The second contribution of our paper is therefore that we show (a suitable version of) the value function for which we have obtained the dynamic programming principle provides a solution to a 2BSDE without requiring any regularity assumption, a case which cannot be covered by the PPDE theory. This takes care of the existence problem, while we tackle, as usual, the uniqueness problem through *a priori* \mathbb{L}^p estimates on the solution, for any $p > 1$. We emphasize that in the very general setting that we consider, the classical method of proof fails (in particular since the filtration we work with is not quasi-left continuous in general), and the estimates follow from a general result that we prove in our accompanying paper [13]. In particular, our wellposedness results contains as a special case the theory of BSDEs, which was not the case neither for the 2BSDEs of [84], nor the G -BSDE. Moreover, the class of probability measures that we can consider is much more general than the ones considered in the previous literature, even allowing for degeneracy of the diffusion coefficient. This is the objective of Section 4.

The rest of the paper is mainly concerned with applications of the previous theory. First, in Section 5, we use our previous results to obtain a nonlinear and robust generalization of the so-called optional decomposition for supermartingales (see, for instance, [39, 56] and the other references given in Section 5 for more details),

which is new in the literature. This allows us to introduce, under an additional assumption stating that the family of measures is roughly speaking rich enough, a new notion of solutions, which we coined saturated 2BSDEs. This new formulation has the advantage that it allows us to get rid of the orthogonal martingales, which generically appear in the definition of a 2BSDE (see Definitions 4.1 and 5.2 for more details). This is particularly important in some applications; see, for instance, the general principal-agent problem studied in [21]. We then give a duality result for the robust pricing of contingent claims in nonlinear and incomplete financial markets. Finally, in Section 6, we recall in our context the link between 2BSDEs and PPDEs when we work under additional regularity assumptions. Compared to [31], our result can accommodate degenerate diffusions.

To conclude this Introduction, we really want to insist on the fact that our new results have much more far-reaching applications, and are not a mere mathematical extension. Indeed, in the paper [21], the well-posedness theory of 2BSDEs we have obtained is used crucially to solve general principal-agent problems in contracting theory, when the agent controls both the drift and the volatility of the corresponding output process (we refer the reader to the excellent monograph [22] for more details on contract theory), a problem which could not be treated with the technics prevailing in the previous literature. Such a result has potential applications in many fields, ranging from economics (see, for instance, [20, 61]) to energy management (see [1]).

Notations: Throughout this paper, we fix a constant $p > 1$. Let $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$ and let \mathbb{R}_+^* be the set of real positive numbers. For every d -dimensional vector b with $d \in \mathbb{N}^*$, we denote by b^1, \dots, b^d its coordinates and for $\alpha, \beta \in \mathbb{R}^d$ we denote by $\alpha \cdot \beta$ the usual inner product, with associated norm $\|\cdot\|$, which we simplify to $|\cdot|$ when d is equal to 1. We also let $\mathbf{1}_d$ be the vector whose coordinates are all equal to 1. For any $(\ell, c) \in \mathbb{N}^* \times \mathbb{N}^*$, $\mathcal{M}_{\ell,c}(\mathbb{R})$ will denote the space of $\ell \times c$ matrices with real entries. Elements of the matrix $M \in \mathcal{M}_{\ell,c}$ will be denoted by $(M^{i,j})_{1 \leq i \leq \ell, 1 \leq j \leq c}$, and the transpose of M will be denoted by M^\top . When $\ell = c$, we let $\mathcal{M}_\ell(\mathbb{R}) := \mathcal{M}_{\ell,\ell}(\mathbb{R})$. We also identify $\mathcal{M}_{\ell,1}(\mathbb{R})$ and \mathbb{R}^ℓ . The identity matrix in $\mathcal{M}_\ell(\mathbb{R})$ will be denoted by I_ℓ . Let $\mathbb{S}_d^{\geq 0}$ denote the set of all symmetric positive semi-definite $d \times d$ matrices. We fix a map $\psi : \mathbb{S}_d^{\geq 0} \rightarrow \mathcal{M}_d(\mathbb{R})$ which is (Borel) measurable and satisfies $\psi(a)\psi(a)^\top = a$ for all $a \in \mathbb{S}_d^{\geq 0}$, and denote $a^{\frac{1}{2}} := \psi(a)$. Finally, we denote by a^\oplus the Moore–Penrose pseudo-inverse of $a \in \mathbb{S}_d^{\geq 0}$. In particular, we know that $a \mapsto a^\oplus = \lim_{\delta \searrow 0} (a^\top a + \delta I_d)^{-1} a^\top$ is Borel measurable.

2. Stochastic control for a class of nonlinear stochastic kernels.

2.1. *Probabilistic framework.*

2.1.1. *Canonical space.* Let $d \in \mathbb{N}^*$. We denote by $\Omega := C([0, T], \mathbb{R}^d)$ the canonical space of all \mathbb{R}^d -valued continuous paths ω on $[0, T]$ such that $\omega_0 =$

0, equipped with the canonical process X , that is, $X_t(\omega) := \omega_t$, for all $\omega \in \Omega$. Denote by $\mathbb{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ the canonical filtration generated by X , and by $\mathbb{F}_+ = (\mathcal{F}_t^+)_{0 \leq t \leq T}$ the right limit of \mathbb{F} with $\mathcal{F}_t^+ := \bigcap_{s>t} \mathcal{F}_s$ for all $t \in [0, T)$ and $\mathcal{F}_T^+ := \mathcal{F}_T$. We equip Ω with the uniform convergence norm $\|\omega\|_\infty := \sup_{0 \leq t \leq T} \|\omega_t\|$, so that the Borel σ -field of Ω coincides with \mathcal{F}_T . Let \mathbb{P}_0 denote the Wiener measure on Ω under which X is a Brownian motion.

Let \mathbb{M}_1 denote the collection of all probability measures on (Ω, \mathcal{F}_T) . Notice that \mathbb{M}_1 is a Polish space equipped with the weak convergence topology. We denote by \mathfrak{B} its Borel σ -field. Then for any $\mathbb{P} \in \mathbb{M}_1$, denote by $\mathcal{F}_t^\mathbb{P}$ the completed σ -field of \mathcal{F}_t under \mathbb{P} . Denote also the completed filtration by $\mathbb{F}^\mathbb{P} = (\mathcal{F}_t^\mathbb{P})_{t \in [0, T]}$ and $\mathbb{F}_+^\mathbb{P}$ the right limit of $\mathbb{F}^\mathbb{P}$, so that $\mathbb{F}_+^\mathbb{P}$ satisfies the usual conditions. Moreover, for $\mathcal{P} \subset \mathbb{M}_1$, we introduce the universally completed filtration $\mathbb{F}^U := (\mathcal{F}_t^U)_{0 \leq t \leq T}$, $\mathbb{F}^\mathcal{P} := (\mathcal{F}_t^\mathcal{P})_{0 \leq t \leq T}$, and $\mathbb{F}^{\mathcal{P}+} := (\mathcal{F}_t^{\mathcal{P}+})_{0 \leq t \leq T}$, defined as follows:

$$\begin{aligned} \mathcal{F}_t^U &:= \bigcap_{\mathbb{P} \in \mathbb{M}_1} \mathcal{F}_t^\mathbb{P}, & \mathcal{F}_t^\mathcal{P} &:= \bigcap_{\mathbb{P} \in \mathcal{P}} \mathcal{F}_t^\mathbb{P}, & t \in [0, T], \\ \mathcal{F}_t^{\mathcal{P}+} &:= \mathcal{F}_{t+}^\mathcal{P}, & t \in [0, T), & \text{ and } & \mathcal{F}_T^{\mathcal{P}+} &:= \mathcal{F}_T^\mathcal{P}. \end{aligned}$$

We also introduce an enlarged canonical space $\bar{\Omega} := \Omega \times \Omega'$, where Ω' is identical to Ω . By abuse of notation, we denote by (X, B) its canonical process, that is, $X_t(\bar{\omega}) := \omega_t$, $B_t(\bar{\omega}) := \omega'_t$ for all $\bar{\omega} := (\omega, \omega') \in \bar{\Omega}$, by $\bar{\mathbb{F}} = (\bar{\mathcal{F}}_t)_{0 \leq t \leq T}$ the canonical filtration generated by (X, B) , and by $\bar{\mathbb{F}}^X = (\bar{\mathcal{F}}_t^X)_{0 \leq t \leq T}$ the filtration generated by X . Similarly, we denote the corresponding right-continuous filtrations by $\bar{\mathbb{F}}_+^X$ and $\bar{\mathbb{F}}_+$, and the augmented filtration by $\bar{\mathbb{F}}_+^{X, \mathbb{P}}$ and $\bar{\mathbb{F}}_+^\mathbb{P}$, given a probability measure \mathbb{P} on $\bar{\Omega}$.

2.1.2. Semimartingale measures. We say that a probability measure \mathbb{P} on (Ω, \mathcal{F}_T) is a semimartingale measure if X is a semimartingale under \mathbb{P} . Then on the canonical space Ω , there is some \mathbb{F} -progressively measurable nondecreasing process (see, e.g., Karandikar [51]), denoted by $\langle X \rangle = (\langle X \rangle_t)_{0 \leq t \leq T}$, which coincides with the quadratic variation of X under each semimartingale measure \mathbb{P} . Denote further

$$\widehat{a}_t := \limsup_{\varepsilon \searrow 0} \frac{\langle X \rangle_t - \langle X \rangle_{t-\varepsilon}}{\varepsilon}.$$

For every $t \in [0, T]$, let \mathcal{P}_t^W denote the collection of all probability measures \mathbb{P} on (Ω, \mathcal{F}_T) such that:

- $(X_s)_{s \in [t, T]}$ is a (\mathbb{P}, \mathbb{F}) -semimartingale admitting the canonical decomposition (see, e.g., [50], Theorem I.4.18)

$$X_s = \int_t^s b_r^\mathbb{P} dr + X_s^{c, \mathbb{P}}, \quad s \in [t, T], \mathbb{P}\text{-a.s.},$$

where $b^{\mathbb{P}}$ is a $\mathbb{F}^{\mathbb{P}}$ -predictable \mathbb{R}^d -valued process, and $X^{c,\mathbb{P}}$ is the continuous local martingale part of X under \mathbb{P} .

- $((X)_s)_{s \in [t, T]}$ is absolutely continuous in s with respect to the Lebesgue measure, and \hat{a} takes values in $\mathbb{S}_d^{\geq 0}$, \mathbb{P} -a.s.

Given a random variable or process λ defined on Ω , we can naturally define its extension on $\bar{\Omega}$ (which, abusing notation slightly, we still denote by λ) by

$$(2.1) \quad \lambda(\bar{\omega}) := \lambda(\omega), \quad \forall \bar{\omega} := (\omega, \omega') \in \bar{\Omega}.$$

In particular, the process \hat{a} can be extended on $\bar{\Omega}$. Given a probability measure $\mathbb{P} \in \mathcal{P}_t^W$, we define a probability measure $\bar{\mathbb{P}}$ on the enlarged canonical space $\bar{\Omega}$ by $\bar{\mathbb{P}} := \mathbb{P} \otimes \mathbb{P}_0$, so that X in $(\bar{\Omega}, \bar{\mathcal{F}}_T, \bar{\mathbb{P}}, \bar{\mathbb{F}})$ is a semimartingale with the same triplet of characteristics as X in $(\Omega, \mathcal{F}_T, \mathbb{P}, \mathbb{F})$, B is a $\bar{\mathbb{F}}$ -Brownian motion, and X is independent of B . Then for every $\mathbb{P} \in \mathcal{P}_t^W$, there is some \mathbb{R}^d -valued, $\bar{\mathbb{F}}$ -Brownian motion $W^{\mathbb{P}} = (W_r^{\mathbb{P}})_{t \leq r \leq s}$ such that (see, e.g., Theorem 4.5.2 of [87])

$$(2.2) \quad X_s = \int_t^s b_r^{\mathbb{P}} dr + \int_t^s \hat{a}_r^{1/2} dW_r^{\mathbb{P}}, \quad s \in [t, T], \bar{\mathbb{P}}\text{-a.s.},$$

where we extend the definition of $b^{\mathbb{P}}$ and \hat{a} on $\bar{\Omega}$ as in (2.1), and where we recall that $\hat{a}^{1/2}$ has been defined in Notation above.

Notice that when $\hat{a}_r^{1/2}$ is nondegenerate \mathbb{P} -a.s., for all $r \in [t, T]$, we can construct the Brownian motion $W^{\mathbb{P}}$ on Ω as follows:

$$W_t^{\mathbb{P}} := \int_0^t \hat{a}_s^{-1/2} dX_s^{c,\mathbb{P}}, \quad t \in [0, T], \mathbb{P}\text{-a.s.},$$

and do not need to consider the above enlarged space equipped with an independent Brownian motion to construct $W^{\mathbb{P}}$.

REMARK 2.1 (On the choice of $\hat{a}^{1/2}$). The measurable map $a \mapsto a^{1/2}$ is fixed throughout the paper. A first choice is to take $a^{1/2}$ as the unique nonnegative symmetric square root of a (see, e.g., Lemma 5.2.1 of [87]). One can also use the Cholesky decomposition to obtain $a^{1/2}$ as a lower triangular matrix. Finally, when $d = m + n$ for $m, n \in \mathbb{N}^*$, and \hat{a} has the specific structure of Remark 2.2 below, one can take $a^{1/2}$ in the following way:

$$(2.3) \quad a = \begin{pmatrix} \sigma \sigma^\top & \sigma \\ \sigma^\top & I_n \end{pmatrix} \quad \text{and} \quad a^{1/2} = \begin{pmatrix} \sigma & 0 \\ I_n & 0 \end{pmatrix}, \quad \text{for some } \sigma \in \mathcal{M}_{m,n}.$$

2.1.3. *Conditioning and concatenation of probability measures.* We also recall that for every probability measure \mathbb{P} on Ω and \mathbb{F} -stopping time τ taking value in $[0, T]$, there exists a family of regular conditional probability distribution (r.c.p.d. for short) $(\mathbb{P}_\omega^\tau)_{\omega \in \Omega}$ (see, e.g., Stroock and Varadhan [87]), satisfying:

- (i) For every $\omega \in \Omega$, \mathbb{P}_ω^τ is a probability measure on (Ω, \mathcal{F}_T) .

(ii) For every $E \in \mathcal{F}_T$, the mapping $\omega \mapsto \mathbb{P}_\omega^\tau(E)$ is \mathcal{F}_τ -measurable.

(iii) The family $(\mathbb{P}_\omega^\tau)_{\omega \in \Omega}$ is a version of the conditional probability measure of \mathbb{P} on \mathcal{F}_τ , that is, for every integrable \mathcal{F}_T -measurable random variable ξ we have $\mathbb{E}^{\mathbb{P}}[\xi | \mathcal{F}_\tau](\omega) = \mathbb{E}^{\mathbb{P}_\omega^\tau}[\xi]$, for \mathbb{P} -a.e. $\omega \in \Omega$.

(iv) For every $\omega \in \Omega$, $\mathbb{P}_\omega^\tau(\Omega_\tau^\omega) = 1$, where $\Omega_\tau^\omega := \{\bar{\omega} \in \Omega : \bar{\omega}(s) = \omega(s), 0 \leq s \leq \tau(\omega)\}$.

Furthermore, given some \mathbb{P} and a family $(\mathbb{Q}_\omega)_{\omega \in \Omega}$ such that $\omega \mapsto \mathbb{Q}_\omega$ is \mathcal{F}_τ -measurable and $\mathbb{Q}_\omega(\Omega_\tau^\omega) = 1$ for all $\omega \in \Omega$, one can then define a concatenated probability measure $\mathbb{P} \otimes_\tau \mathbb{Q}$ by

$$\mathbb{P} \otimes_\tau \mathbb{Q}[A] := \int_\Omega \mathbb{Q}_\omega[A] \mathbb{P}(d\omega) \quad \forall A \in \mathcal{F}_T.$$

2.1.4. *Hypotheses.* We shall consider a random variable $\xi : \Omega \rightarrow \mathbb{R}$ and a generator function

$$f : (t, \omega, y, z, a, b) \in [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}_d^{\geq 0} \times \mathbb{R}^d \rightarrow \mathbb{R}.$$

Define for simplicity

$$\widehat{f}_s^\mathbb{P}(y, z) := f(s, X_{\cdot \wedge s}, y, z, \widehat{a}_s, b_s^\mathbb{P}) \quad \text{and} \quad \widehat{f}_s^{\mathbb{P}, 0} := f(s, X_{\cdot \wedge s}, 0, 0, \widehat{a}_s, b_s^\mathbb{P}).$$

Moreover, we are given a family $(\mathcal{P}(t, \omega))_{(t, \omega) \in [0, T] \times \Omega}$ of sets of probability measures on (Ω, \mathcal{F}_T) , where $\mathcal{P}(t, \omega) \subset \mathcal{P}_t^W$ for all $(t, \omega) \in [0, T] \times \Omega$. Denote also $\mathcal{P}_t := \bigcup_{\omega \in \Omega} \mathcal{P}(t, \omega)$. We make the following assumption on ξ, f and the family $(\mathcal{P}(t, \omega))_{(t, \omega) \in [0, T] \times \Omega}$.

ASSUMPTION 2.1. (i) The random variable ξ is \mathcal{F}_T -measurable, the generator function f is jointly Borel measurable and such that for every $(t, \omega, y, y', z, z', a, b) \in [0, T] \times \Omega \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{S}_d^{\geq 0} \times \mathbb{R}^d$

$$|f(t, \omega, y, z, a, b) - f(t, \omega, y', z', a, b)| \leq C(|y - y'| + \|z - z'\|),$$

and for every fixed (y, z, a, b) , the map $(t, \omega) \mapsto f(t, \omega, y, z, a, b)$ is \mathbb{F} -progressively measurable.

(ii) For the fixed constant $p > 1$, one has, for every $(t, \omega) \in [0, T] \times \Omega$,

$$(2.4) \quad \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \mathbb{E}^\mathbb{P} \left[|\xi|^p + \int_t^T |f(s, X_{\cdot \wedge s}, 0, 0, \widehat{a}_s, b_s^\mathbb{P})|^p ds \right] < +\infty.$$

(iii) For every $(t, \omega) \in [0, T] \times \Omega$, one has $\mathcal{P}(t, \omega) = \mathcal{P}(t, \omega_{\cdot \wedge t})$ and $\mathbb{P}(\Omega_t^\omega) = 1$ whenever $\mathbb{P} \in \mathcal{P}(t, \omega)$. The graph $[[\mathcal{P}]]$ of \mathcal{P} , defined by $[[\mathcal{P}]] := \{(t, \omega, \mathbb{P}) : \mathbb{P} \in \mathcal{P}(t, \omega)\}$, is upper semi-analytic in $[0, T] \times \Omega \times \mathbb{M}_1$.

(iv) \mathcal{P} is stable under conditioning, that is, for every $(t, \omega) \in [0, T] \times \Omega$ and every $\mathbb{P} \in \mathcal{P}(t, \omega)$ together with an \mathbb{F} -stopping time τ taking values in $[t, T]$, there is a family of r.c.p.d. $(\mathbb{P}_w)_{w \in \Omega}$ such that $\mathbb{P}_w \in \mathcal{P}(\tau(w), w)$, for \mathbb{P} -a.e. $w \in \Omega$.

(v) \mathcal{P} is stable under concatenation, that is, for every $(t, \omega) \in [0, T] \times \Omega$ and $\mathbb{P} \in \mathcal{P}(t, \omega)$ together with a \mathbb{F} -stopping time τ taking values in $[t, T]$, let $(\mathbb{Q}_w)_{w \in \Omega}$ be a family of probability measures such that $\mathbb{Q}_w \in \mathcal{P}(\tau(w), w)$ for all $w \in \Omega$ and $w \mapsto \mathbb{Q}_w$ is \mathcal{F}_τ -measurable, then the concatenated probability measure $\mathbb{P} \otimes_\tau \mathbb{Q}$. $\in \mathcal{P}(t, \omega)$.

We notice that for $t = 0$, we have $\mathcal{P}_0 := \mathcal{P}(0, \omega)$ for any $\omega \in \Omega$.

2.2. *Spaces and norms.* We now give the spaces and norms which will be needed in the rest of the paper. Fix some $t \in [0, T]$ and some $\omega \in \Omega$. In what follows, $\mathbb{X} := (\mathcal{X}_s)_{t \leq s \leq T}$ will denote an arbitrary filtration on (Ω, \mathcal{F}_T) , and \mathbb{P} an arbitrary element in $\mathcal{P}(t, \omega)$. Denote also by $\mathbb{X}_{\mathbb{P}}$ the \mathbb{P} -augmented filtration associated to \mathbb{X} :

- For $p \geq 1$, $\mathbb{L}_{t,\omega}^p(\mathbb{X})$ (resp., $\mathbb{L}_{t,\omega}^p(\mathbb{X}, \mathbb{P})$) denotes the space of all \mathcal{X}_T -measurable scalar random variable ξ with

$$\|\xi\|_{\mathbb{L}_{t,\omega}^p}^p := \sup_{\mathbb{P} \in \mathcal{P}(t,\omega)} \mathbb{E}^{\mathbb{P}}[|\xi|^p] < +\infty \quad (\text{resp. } \|\xi\|_{\mathbb{L}_{t,\omega}^p(\mathbb{P})}^p := \mathbb{E}^{\mathbb{P}}[|\xi|^p] < +\infty).$$

- $\mathbb{H}_{t,\omega}^p(\mathbb{X})$ (resp., $\mathbb{H}_{t,\omega}^p(\mathbb{X}, \mathbb{P})$) denotes the space of all \mathbb{X} -predictable \mathbb{R}^d -valued processes Z , which are defined $\widehat{a}_s ds$ -a.e. on $[t, T]$, with

$$\begin{aligned} \|Z\|_{\mathbb{H}_{t,\omega}^p}^p &:= \sup_{\mathbb{P} \in \mathcal{P}(t,\omega)} \mathbb{E}^{\mathbb{P}} \left[\left(\int_t^T \|(\widehat{a}_s^{1/2})^\top Z_s\|^2 ds \right)^{\frac{p}{2}} \right] < +\infty \\ &\left(\text{resp., } \|Z\|_{\mathbb{H}_{t,\omega}^p(\mathbb{P})}^p := \mathbb{E}^{\mathbb{P}} \left[\left(\int_t^T \|(\widehat{a}_s^{1/2})^\top Z_s\|^2 ds \right)^{\frac{p}{2}} \right] < +\infty \right). \end{aligned}$$

- $\mathbb{M}_{t,\omega}^p(\mathbb{X}, \mathbb{P})$ denotes the space of all (\mathbb{X}, \mathbb{P}) -optional martingales M with \mathbb{P} -a.s. càdlàg paths on $[t, T]$, with $M_t = 0$, \mathbb{P} -a.s., and

$$\|M\|_{\mathbb{M}_{t,\omega}^p(\mathbb{P})}^p := \mathbb{E}^{\mathbb{P}}[|M|_T^{\frac{p}{2}}] < +\infty.$$

Furthermore, we will say that a family $(M^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}(t,\omega)}$ belongs to $\mathbb{M}_{t,\omega}^p((\mathbb{X}_{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}(t,\omega)})$ if, for any $\mathbb{P} \in \mathcal{P}(t, \omega)$, $M^{\mathbb{P}} \in \mathbb{M}_{t,\omega}^p(\mathbb{X}_{\mathbb{P}}, \mathbb{P})$ and

$$\sup_{\mathbb{P} \in \mathcal{P}(t,\omega)} \|M^{\mathbb{P}}\|_{\mathbb{M}_{t,\omega}^p(\mathbb{P})} < +\infty.$$

- $\mathbb{I}_{t,\omega}^p(\mathbb{X}, \mathbb{P})$ (resp., $\mathbb{I}_{t,\omega}^{o,p}(\mathbb{X}, \mathbb{P})$) denotes the space of all \mathbb{X} -predictable (resp. \mathbb{X} -optional) processes K with \mathbb{P} -a.s. càdlàg and nondecreasing paths on $[t, T]$, with $K_t = 0$, \mathbb{P} -a.s., and

$$\|K\|_{\mathbb{I}_{t,\omega}^p(\mathbb{P})}^p := \mathbb{E}^{\mathbb{P}}[K_T^p] < +\infty \quad (\text{resp. } \|K\|_{\mathbb{I}_{t,\omega}^{o,p}(\mathbb{P})}^p := \mathbb{E}^{\mathbb{P}}[K_T^p] < +\infty).$$

Similarly we will say that a family $(K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}(t, \omega)}$ belongs to $\mathbb{I}_{t, \omega}^p((\mathbb{X}_{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}(t, \omega)})$ (resp., $\mathbb{I}_{t, \omega}^{o, p}((\mathbb{X}_{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}(t, \omega)})$) if, for any $\mathbb{P} \in \mathcal{P}(t, \omega)$, $K^{\mathbb{P}} \in \mathbb{I}_{t, \omega}^p(\mathbb{X}_{\mathbb{P}}, \mathbb{P})$ (resp. $K^{\mathbb{P}} \in \mathbb{I}_{t, \omega}^{o, p}(\mathbb{X}_{\mathbb{P}}, \mathbb{P})$) and

$$\sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \|K^{\mathbb{P}}\|_{\mathbb{I}_{t, \omega}^p(\mathbb{P})} < +\infty \quad \left(\text{resp. } \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \|K^{\mathbb{P}}\|_{\mathbb{I}_{t, \omega}^{o, p}(\mathbb{P})} < +\infty\right).$$

• $\mathbb{D}_{t, \omega}^p(\mathbb{X})$ (resp., $\mathbb{D}_{t, \omega}^p(\mathbb{X}, \mathbb{P})$) denotes the space of all \mathbb{X} -progressively measurable \mathbb{R} -valued processes Y with $\mathcal{P}(t, \omega)$ -q.s. (resp., \mathbb{P} -a.s.) càdlàg paths on $[t, T]$, with

$$\|Y\|_{\mathbb{D}_{t, \omega}^p}^p := \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \mathbb{E}^{\mathbb{P}} \left[\sup_{t \leq s \leq T} |Y_s|^p \right] < +\infty$$

$$\left(\text{resp. } \|Y\|_{\mathbb{D}_{t, \omega}^p(\mathbb{P})}^p := \mathbb{E}^{\mathbb{P}} \left[\sup_{t \leq s \leq T} |Y_s|^p \right] < +\infty\right).$$

• For each $\xi \in \mathbb{L}_{t, \omega}^1(\mathbb{X})$ and $s \in [t, T]$, denote

$$\mathbb{E}_s^{\mathbb{P}, t, \omega, \mathbb{X}}[\xi] := \text{ess sup}_{\mathbb{P}' \in \mathcal{P}_{t, \omega}(s, \mathbb{P}, \mathbb{X})} \mathbb{E}^{\mathbb{P}'}[\xi | \mathcal{X}_s]$$

where $\mathcal{P}_{t, \omega}(s, \mathbb{P}, \mathbb{X}) := \{\mathbb{P}' \in \mathcal{P}(t, \omega), \mathbb{P}' = \mathbb{P} \text{ on } \mathcal{X}_s\}$.

Then we define for each $p \geq \kappa \geq 1$,

$$\mathbb{L}_{t, \omega}^{p, \kappa}(\mathbb{X}) := \{\xi \in L_{t, \omega}^p(\mathbb{X}), \|\xi\|_{\mathbb{L}_{t, \omega}^{p, \kappa}} < +\infty\},$$

where

$$\|\xi\|_{\mathbb{L}_{t, \omega}^{p, \kappa}}^p := \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \mathbb{E}^{\mathbb{P}} \left[\text{ess sup}_{t \leq s \leq T}^{\mathbb{P}} (\mathbb{E}_s^{\mathbb{P}, t, \omega, \mathbb{P}^+} [|\xi|^{\kappa}])^{\frac{p}{\kappa}} \right].$$

• Similarly, given a probability measure $\bar{\mathbb{P}}$ and a filtration $\bar{\mathbb{X}}$ on the enlarged canonical space $\bar{\Omega}$, we denote the corresponding spaces by $\mathbb{D}_{t, \omega}^p(\bar{\mathbb{X}}, \bar{\mathbb{P}})$, $\mathbb{H}_{t, \omega}^p(\bar{\mathbb{X}}, \bar{\mathbb{P}})$, $\mathbb{M}_{t, \omega}^p(\bar{\mathbb{X}}, \bar{\mathbb{P}}), \dots$ Furthermore, when $t = 0$, there is no longer any dependence on ω , since $\omega_0 = 0$, so that we simplify the notation by suppressing the ω -dependence and write $\mathbb{H}_0^p(\mathbb{X}), \mathbb{H}_0^p(\mathbb{X}, \mathbb{P}), \dots$ Similar notation are used on the enlarged canonical space.

2.3. *Control on a class of nonlinear stochastic kernels and the dynamic programming principle.* For every $(t, \omega) \in [0, T] \times \Omega$ and $\mathbb{P} \in \mathcal{P}(t, \omega)$, we consider the following BSDE:

$$(2.5) \quad \mathcal{Y}_s = \xi - \int_s^T f(r, X_{\cdot \wedge r}, \mathcal{Y}_r, (\hat{a}_r^{1/2})^\top \mathcal{Z}_r, \hat{a}_r, b_r^{\mathbb{P}}) dr$$

$$- \left(\int_s^T \mathcal{Z}_r \cdot dX_r^{c, \mathbb{P}} \right)^{\mathbb{P}} - \int_s^T d\mathcal{M}_r, \quad \mathbb{P}\text{-a.s.},$$

where $(\int_s^T \mathcal{Z}_r \cdot dX_r^{c,\mathbb{P}})^\mathbb{P}$ denotes the stochastic integrable of \mathcal{Z} w.r.t. $X^{c,\mathbb{P}}$ under the probability \mathbb{P} . Following El Karoui and Huang [34], we define a solution to BSDE (2.5) as a triple $(\mathcal{Y}_s^\mathbb{P}, \mathcal{Z}_s^\mathbb{P}, \mathcal{M}_s^\mathbb{P})_{s \in [t, T]} \in \mathbb{D}_{t, \omega}^p(\mathbb{F}_+^\mathbb{P}, \mathbb{P}) \times \mathbb{H}_{t, \omega}^p(\mathbb{F}_+^\mathbb{P}, \mathbb{P}) \times \mathbb{M}_{t, \omega}^p(\mathbb{F}_+^\mathbb{P}, \mathbb{P})$ satisfying the equality (2.5) under \mathbb{P} (well-posedness is a consequence of Lemma 2.2 below). We then define, for every $(t, \omega) \in [0, T] \times \Omega$,

$$(2.6) \quad \widehat{\mathcal{Y}}_t(\omega) := \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \mathbb{E}^\mathbb{P}[\mathcal{Y}_t^\mathbb{P}].$$

Our first main result is the following dynamic programming principle.

THEOREM 2.1. *Suppose that Assumption 2.1 holds true. Then for all $(t, \omega) \in [0, T] \times \Omega$, one has $\widehat{\mathcal{Y}}_t(\omega) = \widehat{\mathcal{Y}}_t(\omega_{t \wedge \cdot})$, and $(t, \omega) \mapsto \widehat{\mathcal{Y}}_t(\omega)$ is $\mathcal{B}([0, T]) \otimes \mathcal{F}_T$ -universally measurable. Moreover, for all $(t, \omega) \in [0, T] \times \Omega$ and \mathbb{F} -stopping time τ taking values in $[t, T]$, we have*

$$\widehat{\mathcal{Y}}_t(\omega) = \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \mathbb{E}^\mathbb{P}[\mathcal{Y}_t^\mathbb{P}(\tau, \widehat{\mathcal{Y}}_\tau)],$$

where $\mathcal{Y}_t^\mathbb{P}(\tau, \widehat{\mathcal{Y}}_\tau)$ is obtained from the solution to the following BSDE with terminal time τ and terminal condition $\widehat{\mathcal{Y}}_\tau$, verifying for any $t \in [0, \tau]$, \mathbb{P} -a.s.:

$$(2.7) \quad \begin{aligned} \mathcal{Y}_t &= \widehat{\mathcal{Y}}_\tau - \int_t^\tau f(s, X_{\cdot \wedge s}, \mathcal{Y}_s, (\widehat{a}_s^{1/2})^\top \mathcal{Z}_s, \widehat{a}_s, b_s^\mathbb{P}) ds \\ &\quad - \left(\int_t^\tau \mathcal{Z}_s \cdot dX_s^{c,\mathbb{P}} \right)^\mathbb{P} - \int_t^\tau d\mathcal{M}_s, \end{aligned}$$

where $(\int_t^\tau \mathcal{Z}_s \cdot dX_s^{c,\mathbb{P}})^\mathbb{P}$ denotes the stochastic integral of \mathcal{Z} w.r.t. $X^{c,\mathbb{P}}$ under probability \mathbb{P} .

REMARK 2.2. In some contexts, the sets $\mathcal{P}(t, \omega)$ are defined as the collections of probability measures induced by a family of controlled diffusion processes. For example, let \mathbb{C}_1 (resp., \mathbb{C}_2) denote the canonical space of all continuous paths ω^1 in $\mathbb{C}([0, T], \mathbb{R}^n)$ (resp., ω^2 in $\mathbb{C}([0, T], \mathbb{R}^m)$) such that $\omega_0^1 = 0$ (resp., $\omega_0^2 = 0$), with canonical process B , canonical filtration \mathbb{F}^1 , and let \mathbb{P}_0 be the corresponding Wiener measure. Let U be a Polish space, $(\mu, \sigma) : [0, T] \times \mathbb{C}_1 \times U \rightarrow \mathbb{R}^n \times \mathcal{M}_{n,m}$ be the coefficient functions. Then given $(t, \omega^1) \in [0, T] \times \mathbb{C}_1$, we denote by $\mathcal{J}(t, \omega^1)$ the collection of all terms

$$\alpha := (\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{P}^\alpha, \mathbb{F}^\alpha = (\mathcal{F}_t^\alpha)_{t \geq 0}, W^\alpha, (v_t^\alpha)_{t \geq 0}, X^\alpha),$$

where $(\Omega^\alpha, \mathcal{F}^\alpha, \mathbb{P}^\alpha, \mathbb{F}^\alpha)$ is a filtered probability space, W^α is a \mathbb{F}^α -Brownian motion, v^α is a U -valued \mathbb{F}^α -predictable process and X^α solves the SDE (under some appropriate additional conditions on μ and σ), with initial condition $X_s^\alpha = \omega_s^1$ for all $s \in [0, t]$,

$$\begin{aligned} X_s^\alpha &= \omega_t^1 + \int_t^s \mu(r, X_{r \wedge \cdot}^\alpha, v_r^\alpha) dr + \int_t^s \sigma(r, X_{r \wedge \cdot}^\alpha, v_r^\alpha) dW_r^\alpha, \\ & \quad s \in [t, T], \mathbb{P}^\alpha\text{-a.s.} \end{aligned}$$

In this case, one can let $d = m + n$ so that $\Omega = \mathbb{C}_1 \times \mathbb{C}_2$ and define $\mathcal{P}(t, \omega)$ for $\omega = (\omega^1, \omega^2)$ as the collection of all probability measures induced by $(X^\alpha, B^\alpha)_{\alpha \in \mathcal{J}(t, \omega^1)}$.

Then, with the choice of $\widehat{a}^{1/2}$ as in (2.3), one can recover σ from it directly, which may be useful for some applications. Moreover, notice that $\mathcal{P}(t, \omega)$ depends only on (t, ω^1) for $\omega = (\omega^1, \omega^2)$, then the value $\widehat{\mathcal{Y}}_t(\omega)$ in (2.6) depends also only on (t, ω^1) .

2.4. Proof of Theorem 2.1.

2.4.1. An equivalent formulation on enlarged canonical space. We would like formulate the BSDE (2.5) on the enlarged canonical space in an equivalent way. Remember that $\overline{\Omega} := \Omega \times \Omega'$ and for a probability measure \mathbb{P} on Ω , we define $\overline{\mathbb{P}} := \mathbb{P} \otimes \mathbb{P}_0$. Then a \mathbb{P} -null event on Ω becomes a $\overline{\mathbb{P}}$ -null event on $\overline{\Omega}$ if it is considered in the enlarged space. Let $\pi : \Omega \times \Omega' \rightarrow \Omega$ be the projection operator defined by $\pi(\omega, \omega') := \omega$, for any $(\omega, \omega') \in \overline{\Omega}$.

LEMMA 2.1. Let $A \subseteq \Omega$ be a subset in Ω . Then saying that A is a \mathbb{P} -null set is equivalent to saying that $\{\bar{\omega} : \pi(\bar{\omega}) \in A\}$ is a $\overline{\mathbb{P}} := \mathbb{P} \otimes \mathbb{P}_0$ -null set.

PROOF. For $A \subseteq \Omega$, denote $\overline{A} := \{\bar{\omega} : \pi(\bar{\omega}) \in A\} = A \times \Omega'$. Then by the definition of the product measure, it is clear that

$$\mathbb{P}(A) = 0 \iff \mathbb{P} \otimes \mathbb{P}_0(\overline{A}) = 0,$$

which concludes the proof. \square

We now consider two BSDEs on the enlarged canonical space, w.r.t. two different filtrations. The first one is the following BSDE on $(\overline{\Omega}, \overline{\mathcal{F}}_T^X, \overline{\mathbb{P}})$ w.r.t the filtration $\overline{\mathbb{F}}^{X, \overline{\mathbb{P}}}$:

$$(2.8) \quad \begin{aligned} \overline{\mathcal{Y}}_s &= \xi(X_\cdot) - \int_s^T f(r, X_{\cdot \wedge r}, \overline{\mathcal{Y}}_r, (\widehat{a}_r^{1/2})^\top \overline{\mathcal{Z}}_r, \widehat{a}_r, b_r^\mathbb{P}) dr \\ &\quad - \left(\int_s^T \overline{\mathcal{Z}}_r \cdot dX_r^{c, \mathbb{P}} \right)^{\overline{\mathbb{P}}} - \int_s^T d\overline{\mathcal{M}}_r, \end{aligned}$$

a solution being a triple $(\overline{\mathcal{Y}}_s^{\overline{\mathbb{P}}}, \overline{\mathcal{Z}}_s^{\overline{\mathbb{P}}}, \overline{\mathcal{M}}_s^{\overline{\mathbb{P}}})_{s \in [t, T]} \in \mathbb{D}_{t, \omega}^p(\overline{\mathbb{F}}_+^{X, \overline{\mathbb{P}}}, \overline{\mathbb{P}}) \times \mathbb{H}_{t, \omega}^p(\overline{\mathbb{F}}_+^{X, \overline{\mathbb{P}}}, \overline{\mathbb{P}}) \times \mathbb{M}_{t, \omega}^p(\overline{\mathbb{F}}_+^{X, \overline{\mathbb{P}}}, \overline{\mathbb{P}})$ satisfying (2.8). Notice that in the enlargement, the Brownian motion B is independent of X , so that the above BSDE (2.8) is equivalent to BSDE (2.5) (see Lemma 2.2 below for a precise statement and justification).

We then introduce a second BSDE on the enlarged space $(\bar{\Omega}, \bar{\mathcal{F}}_T, \bar{\mathbb{P}})$, w.r.t. the filtration $\bar{\mathbb{F}}$

$$(2.9) \quad \begin{aligned} \tilde{\mathcal{Y}}_s &= \xi(X.) - \int_s^T f(r, X_{\cdot \wedge r}, \tilde{\mathcal{Y}}_r, (\hat{a}_r^{1/2})^\top \tilde{\mathcal{Z}}_r, \hat{a}_r, b_r^\mathbb{P}) dr \\ &\quad - \left(\int_s^T \tilde{\mathcal{Z}}_r \cdot \hat{a}_r^{1/2} dW_r^\mathbb{P} \right)^{\bar{\mathbb{P}}} - \int_s^T d\tilde{\mathcal{M}}_r, \end{aligned}$$

a solution being a triple $(\tilde{\mathcal{Y}}_s^{\bar{\mathbb{P}}}, \tilde{\mathcal{Z}}_s^{\bar{\mathbb{P}}}, \tilde{\mathcal{M}}_s^{\bar{\mathbb{P}}})_{s \in [t, T]} \in \mathbb{D}_{t, \omega}^p(\bar{\mathbb{F}}_+^{\bar{\mathbb{P}}}, \bar{\mathbb{P}}) \times \mathbb{H}_{t, \omega}^p(\bar{\mathbb{F}}_+^{\bar{\mathbb{P}}}, \bar{\mathbb{P}}) \times \mathbb{M}_{t, \omega}^p(\bar{\mathbb{F}}_+^{\bar{\mathbb{P}}}, \bar{\mathbb{P}})$ satisfying (2.9).

LEMMA 2.2. *Let $(t, \omega) \in [0, T] \times \Omega$, $\mathbb{P} \in \mathcal{P}(t, \omega)$ and $\bar{\mathbb{P}} := \mathbb{P} \otimes \mathbb{P}_0$, then each of the three BSDEs (2.5), (2.8) and (2.9) has a unique solution, denoted respectively by $(\mathcal{Y}, \mathcal{Z}, \mathcal{M})$, $(\bar{\mathcal{Y}}, \bar{\mathcal{Z}}, \bar{\mathcal{M}})$ and $(\tilde{\mathcal{Y}}, \tilde{\mathcal{Z}}, \tilde{\mathcal{M}})$. Moreover, their solution coincide in the sense that there is some functional*

$$\Psi := (\Psi^Y, \Psi^Z, \Psi^M) : [t, T] \times \Omega \longrightarrow \mathbb{R} \times \mathbb{R}^d \times \mathbb{R},$$

such that Ψ^Y and Ψ^M are \mathbb{F}^+ -progressively measurable and \mathbb{P} -a.s. càdlàg, Ψ^Z is \mathbb{F} -predictable, and

$$\begin{aligned} \mathcal{Y}_s &= \Psi_s^Y, & \mathcal{Z}_r &= \Psi_r^Z, & \hat{a}_r & \text{dr-a.e. on } [t, s], \\ \mathcal{M}_s &= \Psi_s^M, & & & & \text{for all } s \in [t, T], \mathbb{P}\text{-a.s.}, \\ \bar{\mathcal{Y}}_s &= \tilde{\mathcal{Y}}_s = \Psi_s^Y(X.), & \bar{\mathcal{Z}}_r &= \tilde{\mathcal{Z}}_r = \Psi_r^Z(X.), & \hat{a}_r & \text{dr-a.e. on } [t, s], \\ \bar{\mathcal{M}}_s &= \tilde{\mathcal{M}}_s = \Psi_s^M(X.), \end{aligned}$$

for all $s \in [t, T]$, $\bar{\mathbb{P}}$ -a.s.

PROOF. (i) The existence and uniqueness of a solution to (2.9) is a classical result; we can, for example, refer to Theorem 4.1 of [13]. Then it is enough to show that the three BSDEs share the same solution in the sense given in the statement. Without loss of generality, we assume in the following $t = 0$.

(ii) We next show that (2.8) and (2.9) admit the same solution in $(\bar{\Omega}, \bar{\mathcal{F}}_T, \bar{\mathbb{P}})$. Notice that a solution to (2.8) is clearly a solution to (2.9) by (2.2). We then show that a solution to (2.9) is also a solution to (2.8).

Let $\zeta : \bar{\Omega} \longrightarrow \mathbb{R}$ be a $\bar{\mathcal{F}}_T^{X, \bar{\mathbb{P}}}$ -measurable random variable, which admits a unique martingale representation

$$(2.10) \quad \zeta = \mathbb{E}^{\bar{\mathbb{P}}}[\zeta] + \int_0^T \bar{\mathcal{Z}}_s^\zeta \cdot dX_s^{c, \bar{\mathbb{P}}} + \int_0^T \bar{\mathcal{M}}_s^\zeta,$$

w.r.t. the filtration $\bar{\mathbb{F}}_+^{X, \bar{\mathbb{P}}}$. Since B is independent of X in the enlarged space, and since X admits the same semimartingale triplet of characteristics in both space,

the above representation (2.10) w.r.t. $\overline{\mathbb{F}}_+^{X, \overline{\mathbb{P}}}$ is the same as the one w.r.t. $\overline{\mathbb{F}}_+^{\overline{\mathbb{P}}}$, which are all unique up to a $\overline{\mathbb{P}}$ -evanescent set. Remember now that the solution of BSDE (2.9) is actually obtained as an iteration of the above martingale representation (see, e.g., Section 2.4.2 below). Therefore, a solution to (2.9) is clearly a solution to (2.8).

(iii) We now show that a solution $(\overline{\mathcal{Y}}, \overline{\mathcal{Z}}, \overline{\mathcal{M}})$ to (2.8) induces a solution to (2.5). Notice that $\overline{\mathcal{Y}}$ and $\overline{\mathcal{M}}$ are $\overline{\mathbb{F}}_+^{X, \overline{\mathbb{P}}}$ -optional, and $\overline{\mathcal{Z}}$ is $\overline{\mathbb{F}}_+^{X, \overline{\mathbb{P}}}$ -predictable, then (see, e.g., Lemma 2.4 of [85] and Theorem IV.78 and Remark IV.74 of [25]) there exists a functional $(\overline{\Psi}^Y, \overline{\Psi}^Z, \overline{\Psi}^M) : [0, T] \times \Omega \rightarrow \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$ such that $\overline{\Psi}^Y$ and $\overline{\Psi}^M$ are $\overline{\mathbb{F}}_+^X$ -progressively measurable and $\overline{\mathbb{P}}$ -a.s. càdlàg, $\overline{\Psi}^Z$ is $\overline{\mathbb{F}}_+^X$ -predictable and $\overline{\mathcal{Y}}_t = \overline{\Psi}_t^Y, \overline{\mathcal{Z}}_t = \overline{\Psi}_t^Z$ and $\overline{\mathcal{M}}_t = \overline{\Psi}_t^M$, for all $t \in [0, T]$, $\overline{\mathbb{P}}$ -a.s. Define

$$(\overline{\Psi}^{Y,0}(\omega), \overline{\Psi}^{Z,0}(\omega), \overline{\Psi}^{M,0}(\omega)) := (\overline{\Psi}^Y(\omega, \mathbf{0}), \overline{\Psi}^Z(\omega, \mathbf{0}), \overline{\Psi}^M(\omega, \mathbf{0})),$$

where $\mathbf{0}$ denotes the path taking value 0 for all $t \in [0, T]$.

Since $(\overline{\Psi}^Y, \overline{\Psi}^Z, \overline{\Psi}^M)$ are $\overline{\mathbb{F}}_+^X$ -progressively measurable, the functions $(\overline{\Psi}^{Y,0}, \overline{\Psi}^{Z,0}, \overline{\Psi}^{M,0})$ are then $\overline{\mathbb{F}}$ -progressively measurable, and it is easy to see that they provide a version of a solution to (2.5) in $(\Omega, \mathcal{F}_T^{\overline{\mathbb{P}}}, \overline{\mathbb{P}})$.

(iv) Finally, let $(\mathcal{Y}, \mathcal{Z}, \mathcal{M})$ be a solution to (2.5), then there is a function $(\Psi^Y, \Psi^Z, \Psi^M) : [0, T] \times \Omega \rightarrow \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$ such that Ψ^Y and Ψ^M are \mathbb{F}^+ -progressively measurable and \mathbb{P} -a.s. càdlàg, Ψ^Z is \mathbb{F} -predictable, and $\mathcal{Y}_t = \Psi_t, \mathcal{Z}_t = \Psi_t^Z$ and $\mathcal{M}_t = \Psi_t^M$, for all $t \in [0, T]$, \mathbb{P} -a.s. Since $\overline{\mathbb{P}} := \mathbb{P} \otimes \mathbb{P}_0$, it is easy to see that (Ψ^Y, Ψ^Z, Ψ^M) is the required functional in the lemma. \square

The main interest of Lemma 2.2 is that it allows us, when studying the BSDE (2.5), to equivalently work with the BSDE (2.9), in which the Brownian motion $W^{\overline{\mathbb{P}}}$ appears explicitly. This will be particularly important for us when using linearization arguments. Indeed, in such type of arguments, one usually introduce a new probability measure equivalent to \mathbb{P} . But if we use formulation (2.5), then one must make the inverse of \hat{a} appear explicitly in the Radon–Nykodym density of the new probability measure. Since such an inverse is not always defined in our setting, we therefore take advantage of the enlarged space formulation to bypass this problem.

2.4.2. *An iterative construction of the solution to BSDE (2.5).* In preparation of the proof of the dynamic programming principle for control problem in Theorem 2.1, let us first recall the classical construction of the $\mathcal{Y}^{\overline{\mathbb{P}}}$ part of the solution to the BSDE (2.5) under some probability $\overline{\mathbb{P}} \in \mathcal{P}(t, \omega)$ using Picard’s iteration. Let us first define for any $m \geq 0$

$$\begin{aligned} \xi^m &:= (\xi \wedge m) \vee (-m), \\ f^m(t, \omega, y, z, a, b) &:= (f(t, \omega, y, z, a, b) \wedge m) \vee (-m). \end{aligned}$$

(i) First, let $\mathcal{Y}_s^{\mathbb{P},0,m} \equiv 0$ and $\mathcal{Z}_s^{\mathbb{P},0,m} \equiv 0$, for all $s \in [t, T]$.

(ii) Given a family of \mathbb{F}_+ -progressively measurable processes $(\mathcal{Y}_s^{\mathbb{P},n,m}, \mathcal{Z}_s^{\mathbb{P},n,m})_{s \in [t, T]}$, we let

$$(2.11) \quad \begin{aligned} \bar{\mathcal{Y}}_s^{\mathbb{P},n+1,m} &:= \mathbb{E}^{\mathbb{P}} \left[\xi^m - \int_s^T f^m(r, X_{\cdot \wedge r}, \mathcal{Y}_r^{\mathbb{P},n,m}, (\hat{a}_r^{1/2})^\top \mathcal{Z}_r^{\mathbb{P},n,m}, \right. \\ &\quad \left. \hat{a}_r, b_r^{\mathbb{P}}) dr \middle| \mathcal{F}_s \right], \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

(iii) Let $\mathcal{Y}^{\mathbb{P},n+1,m}$ be a right-continuous modification of $\bar{\mathcal{Y}}^{\mathbb{P},n+1,m}$ defined by

$$(2.12) \quad \mathcal{Y}_s^{\mathbb{P},n+1,m} := \limsup_{\mathbb{Q} \ni r \downarrow s} \bar{\mathcal{Y}}_r^{\mathbb{P},n+1,m}, \quad \mathbb{P}\text{-a.s.}$$

(iv) Notice that $\mathcal{Y}^{\mathbb{P},n+1,m}$ is a semimartingale under \mathbb{P} . Let $(\mathcal{Y}^{\mathbb{P},n+1,m}, X)^{\mathbb{P}}$ be the predictable quadratic covariation of the process $\mathcal{Y}^{\mathbb{P},n+1,m}$ and X under \mathbb{P} . Define

$$(2.13) \quad \mathcal{Z}_s^{\mathbb{P},n+1,m} := \hat{a}_s^\oplus \left(\limsup_{\mathbb{Q} \ni \varepsilon \downarrow 0} \frac{\langle \mathcal{Y}^{\mathbb{P},n+1,m}, X \rangle_s^{\mathbb{P}} - \langle \mathcal{Y}^{\mathbb{P},n+1,m}, X \rangle_{s-\varepsilon}^{\mathbb{P}}}{\varepsilon} \right).$$

(v) Notice that the sequence $(\mathcal{Y}^{\mathbb{P},n,m})_{n \geq 0}$ is a Cauchy sequence for the norm

$$\|(Y, Z)\|_\alpha^2 := \mathbb{E}^{\mathbb{P}} \left[\int_0^T e^{\alpha s} |Y_s|^2 ds \right]^2 + \mathbb{E}^{\mathbb{P}} \left[\int_0^T e^{\alpha s} \|(\hat{a}_s^{1/2})^\top Z_s\|^2 ds \right]^2,$$

for α large enough. Indeed, this is a consequence of the classical estimates for BSDEs, for which we refer to Section 4 of [13].⁵ Then by taking some suitable subsequence $(n_k^{\mathbb{P},m})_{k \geq 1}$, we can define

$$\mathcal{Y}_s^{\mathbb{P},m} := \limsup_{k \rightarrow \infty} \mathcal{Y}_s^{\mathbb{P},n_k^{\mathbb{P},m},m}.$$

(vi) Finally, we can again use the estimates given in [13] (see again Section 4) to show that the sequence $(\mathcal{Y}^{\mathbb{P},m})_{m \geq 0}$ is a Cauchy sequence in $\mathbb{D}_0^p(\bar{\mathbb{F}}_+^{\mathbb{P}}, \mathbb{P})$, so that by taking once more a suitable subsequence $(m_k^{\mathbb{P}})_{k \geq 1}$, we can define the solution to the BSDE as

$$(2.14) \quad \mathcal{Y}_s^{\mathbb{P}} := \limsup_{k \rightarrow \infty} \mathcal{Y}_s^{\mathbb{P},m_k^{\mathbb{P}}}.$$

⁵Notice here that the results of [13] apply for BSDEs of the form (2.9) in the enlarged space. However, by Lemma 2.2, this implies the same convergence result in the original space.

2.4.3. *On the measurability issues of the iteration.* Here, we show that the iteration in Section 2.4.2 can be taken in a measurable way w.r.t. the reference probability measure \mathbb{P} , which allows us to use the measurable selection theorem to derive the dynamic programming principle.

LEMMA 2.3. *Let \mathcal{P} be a measurable set in \mathbb{M}_1 , $(\mathbb{P}, \omega, t) \mapsto H_t^\mathbb{P}(\omega)$ be a measurable function such that for all $\mathbb{P} \in \mathcal{P}$, $H^\mathbb{P}$ is right-continuous, \mathbb{F}_+ -adapted and a $(\mathbb{P}, \mathbb{F}_+^\mathbb{P})$ -semimartingale. Then there is a measurable function $(\mathbb{P}, \omega, t) \mapsto \langle H \rangle_t^\mathbb{P}(\omega)$ such that for all $\mathbb{P} \in \mathcal{P}$, $\langle H \rangle^\mathbb{P}$ is right-continuous, \mathbb{F}_+ -adapted and $\mathbb{F}_+^\mathbb{P}$ -predictable, and*

$\langle H \rangle^\mathbb{P}$ is the predictable quadratic variation of the semimartingale $H^\mathbb{P}$ under \mathbb{P} .

PROOF. (i) For every $n \geq 1$, we define the following sequence of random times:

$$(2.15) \quad \begin{cases} \tau_0^{\mathbb{P},n}(\omega) := 0, & \omega \in \Omega, \\ \tau_{i+1}^{\mathbb{P},n}(\omega) := \inf\{t \geq \tau_i^{\mathbb{P},n}(\omega), |H_t^\mathbb{P}(\omega) - H_{\tau_i^{\mathbb{P},n}}^\mathbb{P}(\omega)| \geq 2^{-n}\} \wedge T, \\ \omega \in \Omega, i \geq 0. \end{cases}$$

We notice that the $\tau_i^{\mathbb{P},n}$ are all \mathbb{F}_+ -stopping times since the $H^\mathbb{P}$ are right-continuous and \mathbb{F}_+ -adapted. We then define

$$(2.16) \quad [H^\mathbb{P}]_t(\omega) := \limsup_{n \rightarrow +\infty} \sum_{i \geq 0} (H_{\tau_{i+1}^{\mathbb{P},n} \wedge t}^\mathbb{P}(\omega) - H_{\tau_i^{\mathbb{P},n} \wedge t}^\mathbb{P}(\omega))^2.$$

It is clear that $(\mathbb{P}, \omega, t) \mapsto [H^\mathbb{P}]_t(\omega)$ is a measurable function, and for all $\mathbb{P} \in \mathcal{P}$, $[H^\mathbb{P}]$ is nondecreasing, \mathbb{F}_+ -adapted and $\mathbb{F}_+^\mathbb{P}$ -optional. Then it follows by Karandikar [51] that $[H^\mathbb{P}]$ coincides with the quadratic variation of the semimartingale $H^\mathbb{P}$ under \mathbb{P} . Moreover, by taking its right limit over rational time instants, we can choose $[H^\mathbb{P}]$ to be right continuous.

(ii) Finally, using Proposition 5.1 of Neufeld and Nutz [66], we can then construct a process $\langle H \rangle_t^\mathbb{P}(\omega)$ satisfying the required conditions. \square

Notice that the construction above can also be carried out for the predictable quadratic covariation $\langle H^{\mathbb{P},1}, H^{\mathbb{P},2} \rangle^\mathbb{P}$, by defining it through the polarization identity

$$(2.17) \quad \langle H^{\mathbb{P},1}, H^{\mathbb{P},2} \rangle^\mathbb{P} := \frac{1}{4} (\langle H^{\mathbb{P},1} + H^{\mathbb{P},2} \rangle^\mathbb{P} - \langle H^{\mathbb{P},1} - H^{\mathbb{P},2} \rangle^\mathbb{P}),$$

for all measurable functions $H_t^{\mathbb{P},1}(\omega)$ and $H_t^{\mathbb{P},2}(\omega)$ satisfying the conditions in Lemma 2.3. We now show that the iteration in Section 2.4.2 can be taken in a measurable way w.r.t. \mathbb{P} , which provides a key step for the proof of Theorem 2.1.

LEMMA 2.4. *Let $m, n > 0$ be fixed, $(s, \omega, \mathbb{P}) \mapsto (\mathcal{Y}_s^{\mathbb{P},n,m}(\omega), \mathcal{Z}_s^{\mathbb{P},n,m}(\omega))$ be a measurable map such that for every $\mathbb{P} \in \mathcal{P}_t$, $\mathcal{Y}^{\mathbb{P},n,m}$ is right-continuous, \mathbb{F}_+ -adapted and $\mathbb{F}_+^{\mathbb{P}}$ -optional, $\mathcal{Z}^{\mathbb{P},n,m}$ is \mathbb{F}_+ -adapted and $\mathbb{F}_+^{\mathbb{P}}$ -predictable. Then we can choose a measurable map $(s, \omega, \mathbb{P}) \mapsto (\mathcal{Y}_s^{\mathbb{P},n,m}(\omega), \mathcal{Z}_s^{\mathbb{P},n,m}(\omega))$ such that for every $\mathbb{P} \in \mathcal{P}_t$, $\mathcal{Y}^{\mathbb{P},n+1,m}$ is right-continuous, \mathbb{F}_+ -adapted and $\mathbb{F}_+^{\mathbb{P}}$ -optional, and $\mathcal{Z}^{\mathbb{P},n+1,m}$ is \mathbb{F}_+ -adapted and $\mathbb{F}_+^{\mathbb{P}}$ -predictable.*

PROOF. (i) First, using Lemma 3.1 of Neufeld and Nutz [66], there is a version of $(\bar{\mathcal{Y}}^{\mathbb{P},n+1,m})$ defined by (2.11), such that $(\mathbb{P}, \omega) \mapsto \bar{\mathcal{Y}}_s^{\mathbb{P},n+1,m}$ is $\mathfrak{B} \otimes \mathcal{F}_s$ -measurable for every $s \in [t, T]$.

(ii) Next, we notice that the measurability is not lost by taking the limit along a countable sequence. Then with the above version of $(\bar{\mathcal{Y}}^{\mathbb{P},n+1,m})$, it is clear that the family $(\mathcal{Y}_s^{\mathbb{P},n+1,m}(\omega))$ defined by (2.12) is measurable in (s, ω, \mathbb{P}) , and for all $\mathbb{P} \in \mathcal{P}_t$, $\mathcal{Y}^{\mathbb{P},n+1,m}$ is \mathbb{F}_+ -adapted and $\mathbb{F}_+^{\mathbb{P}}$ -optional.

(iii) Then using Lemma 2.3 as well as the definition of the quadratic covariation in (2.17), it follows that there is a measurable function:

$$(s, \omega, \mathbb{P}) \mapsto \langle \mathcal{Y}^{\mathbb{P},n+1,m}, X \rangle_s^{\mathbb{P}}(\omega),$$

such that for every $\mathbb{P} \in \mathcal{P}_t$, $\langle \mathcal{Y}^{\mathbb{P},n+1,m}, X \rangle^{\mathbb{P}}$ is right-continuous, \mathbb{F}_+ -adapted and coincides with the predictable quadratic covariation of $\mathcal{Y}^{\mathbb{P},n+1,m}$ and X under \mathbb{P} .

(iv) Finally, with the above version of $(\langle \mathcal{Y}^{\mathbb{P},n+1,m}, X \rangle^{\mathbb{P}})$, it is clear that the family $(\mathcal{Z}_s^{\mathbb{P},n+1,m}(\omega))$ defined by (2.13) is measurable in (s, ω, \mathbb{P}) and for every $\mathbb{P} \in \mathcal{P}_t$, $\mathcal{Z}^{\mathbb{P},n+1,m}$ is \mathbb{F}_+ -adapted and $\mathbb{F}_+^{\mathbb{P}}$ -predictable. \square

LEMMA 2.5. *For every $\mathbb{P} \in \mathcal{P}_t$, there is a right-continuous, $\mathbb{F}_+^{\mathbb{P}}$ -martingale $\mathcal{M}^{\mathbb{P},n+1,m}$ orthogonal to X under \mathbb{P} , such that \mathbb{P} -a.s.*

$$\begin{aligned} \mathcal{Y}_t^{\mathbb{P},n+1,m} &= \xi^m - \int_t^T f^m(s, X_{\cdot \wedge s}, \mathcal{Y}_s^{\mathbb{P},n,m}, (\hat{a}_s^{1/2})^\top \mathcal{Z}_s^{\mathbb{P},n,m}, \hat{a}_s, b_s^{\mathbb{P}}) ds \\ (2.18) \quad & - \int_t^T d\mathcal{M}_s^{\mathbb{P},n+1,m} - \left(\int_t^T \mathcal{Z}_s^{\mathbb{P},n+1,m} \cdot dX_s^{c,\mathbb{P}} \right)^{\mathbb{P}}. \end{aligned}$$

PROOF. Using Doob’s upcrossing inequality, the limit $\lim_{r \downarrow s} \bar{\mathcal{Y}}_r^{\mathbb{P},n+1,m}$ exists \mathbb{P} -almost surely, for every $\mathbb{P} \in \mathcal{P}_t$. In other words, $\mathcal{Y}^{\mathbb{P},n+1,m}$ is version of the right-continuous modification of the semimartingale $\bar{\mathcal{Y}}^{\mathbb{P},n+1,m}$. Using the martingale representation, it follows that there is a right-continuous, $\mathbb{F}_+^{\mathbb{P}}$ -martingale $\mathcal{M}^{\mathbb{P},n+1,m}$ orthogonal to X under \mathbb{P} , and an $\mathbb{F}_+^{\mathbb{P}}$ -predictable process $\hat{\mathcal{Z}}^{\mathbb{P},n+1,m}$ such that

$$\begin{aligned} \mathcal{Y}_t^{\mathbb{P},n+1,m} &= \xi^m - \int_t^T f^m(s, X_{\cdot \wedge s}, \mathcal{Y}_s^{\mathbb{P},n,m}, (\hat{a}_s^{1/2})^\top \mathcal{Z}_s^{\mathbb{P},n,m}, \hat{a}_s, b_s^{\mathbb{P}}) ds \\ & - \int_t^T d\mathcal{M}_s^{\mathbb{P},n+1,m} - \left(\int_t^T \hat{\mathcal{Z}}_s^{\mathbb{P},n+1,m} \cdot dX_s^{c,\mathbb{P}} \right)^{\mathbb{P}}. \end{aligned}$$

In particular, $\widehat{\mathcal{Z}}^{\mathbb{P},n+1,m}$ satisfies

$$\langle \mathcal{Y}^{\mathbb{P},n+1,m}, X^{c,\mathbb{P}} \rangle_t = \int_0^t \widehat{a}_s \widehat{\mathcal{Z}}_s^{\mathbb{P},n+1,m} ds, \quad \mathbb{P}\text{-a.s.}$$

Besides, by the definition of $\mathcal{Z}^{\mathbb{P},n+1,m}$ in (2.13), one has

$$\int_0^t \widehat{a}_s \mathcal{Z}_s^{\mathbb{P},n+1,m} ds = \langle \mathcal{Y}^{\mathbb{P},n+1,m}, X^{c,\mathbb{P}} \rangle_t = \int_0^t \widehat{a}_s \widehat{\mathcal{Z}}_s^{\mathbb{P},n+1,m} ds, \quad \mathbb{P}\text{-a.s.}$$

It follows that

$$\begin{aligned} & \int_0^t \|(\widehat{a}_s^{1/2})^\top (\mathcal{Z}_s^{\mathbb{P},n+1,m} - \widehat{\mathcal{Z}}_s^{\mathbb{P},n+1,m})\|^2 ds \\ &= \int_0^t (\mathcal{Z}_s^{\mathbb{P},n+1,m} - \widehat{\mathcal{Z}}_s^{\mathbb{P},n+1,m})^\top \widehat{a}_s (\mathcal{Z}_s^{\mathbb{P},n+1,m} - \widehat{\mathcal{Z}}_s^{\mathbb{P},n+1,m}) ds = 0, \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

Hence, (2.18) holds true. \square

LEMMA 2.6. *There are families of subsequences $(n_k^{\mathbb{P},m}, k \geq 1)$ and $(m_i^{\mathbb{P}}, i \geq 1)$ such that the limit $\mathcal{Y}_s^{\mathbb{P}}(\omega) = \lim_{i \rightarrow \infty} \lim_{k \rightarrow \infty} \mathcal{Y}_s^{\mathbb{P},n_k^{\mathbb{P},m},m_i^{\mathbb{P}}}$ exists for all $s \in [t, T]$, \mathbb{P} -almost surely, for every $\mathbb{P} \in \mathcal{P}_t$, and $(s, \omega, \mathbb{P}) \mapsto \mathcal{Y}_s^{\mathbb{P}}(\omega)$ is a measurable function. Moreover, $\mathcal{Y}^{\mathbb{P}}$ provides a solution to the BSDE (2.5) for every $\mathbb{P} \in \mathcal{P}_t$.*

PROOF. By the conditions in (2.4), $(\mathcal{Y}^{\mathbb{P},n,m}, \mathcal{Z}^{\mathbb{P},n,m})_{n \geq 1}$ provides a Picard iteration under the (\mathbb{P}, β) -norm, for $\beta > 0$ large enough (see, e.g., Section 4 of [13]⁶), defined by

$$\|\varphi\|_{\mathbb{P},\beta}^2 := \mathbb{E}^{\mathbb{P}} \left[\sup_{t \leq s \leq T} e^{\beta s} |\varphi_s|^2 \right].$$

Hence, $\mathcal{Y}^{\mathbb{P},n,m}$ converges (under the (\mathbb{P}, β) -norm) to some process $\mathcal{Y}^{\mathbb{P},m}$ as $n \rightarrow \infty$, which solves the BSDE (2.5) with the truncated terminal condition ξ^m and truncated generator f^m . Moreover, by the estimates in Section 4 of [13] (see again Footnote 6), $(\mathcal{Y}^{\mathbb{P},m})_{m \geq 1}$ is a Cauchy sequence in $\mathbb{D}_{t,\omega}^{\mathbb{P}}(\overline{\mathbb{F}}_+, \mathbb{P})$. Then using Lemma 3.2 of [66], we can find two families of subsequences $(n_k^{\mathbb{P},m}, k \geq 1, \mathbb{P} \in \mathcal{P}_t)$ and $(m_i^{\mathbb{P}}, i \geq 1, \mathbb{P} \in \mathcal{P}_t)$ satisfying the required properties. \square

2.4.4. *End of the Proof of Theorem 2.1.* Now we can complete the proof of the dynamic programming in Theorem 2.1. Let us first provide a tower property for the BSDE (2.5).

⁶Again, we remind the reader that one should first apply the result of [13] to the corresponding Picard iteration of (2.9) in the enlarged space and then use Lemma 2.2 to go back to the original space.

LEMMA 2.7. *Let $(t, \omega) \in [0, T] \times \Omega$, $\mathbb{P} \in \mathcal{P}(t, \omega)$, τ be an \mathbb{F} -stopping time taking values in $[t, T]$ and $(\mathcal{Y}^\mathbb{P}, \mathcal{Z}^\mathbb{P}, \mathcal{M}^\mathbb{P})$ be a solution to the BSDE (2.5) under \mathbb{P} . Then one has*

$$\mathcal{Y}_t^\mathbb{P}(T, \xi) = \mathcal{Y}_t^\mathbb{P}(\tau, \mathcal{Y}_\tau^\mathbb{P}) = \mathcal{Y}_t^\mathbb{P}(\tau, \mathbb{E}^\mathbb{P}[\mathcal{Y}_\tau^\mathbb{P} | \mathcal{F}_\tau]), \quad \mathbb{P}\text{-a.s.}$$

PROOF. (i) Given a solution $(\mathcal{Y}^\mathbb{P}, \mathcal{Z}^\mathbb{P}, \mathcal{M}^\mathbb{P})$ to the BSDE (2.5) under \mathbb{P} w.r.t the filtration $\mathbb{F}_+^\mathbb{P} = (\mathcal{F}_{s+}^\mathbb{P})_{t \leq s \leq T}$, one then has

$$\begin{aligned} \mathcal{Y}_t^\mathbb{P} &= \mathcal{Y}_\tau^\mathbb{P} - \int_t^\tau f(s, X_{\cdot \wedge s}, \mathcal{Y}_s^\mathbb{P}, (\widehat{a}_s^{1/2})^\top \mathcal{Z}_s^\mathbb{P}, \widehat{a}_s, b_s^\mathbb{P}) ds \\ &\quad - \left(\int_t^\tau \mathcal{Z}_s^\mathbb{P} \cdot dX_s^{c, \mathbb{P}} \right)^\mathbb{P} - \int_t^\tau d\mathcal{M}_s^\mathbb{P}, \quad \mathbb{P}\text{-a.s.} \end{aligned}$$

By taking conditional expectation w.r.t. $\mathcal{F}_\tau^\mathbb{P}$ under \mathbb{P} , it follows that, \mathbb{P} -a.s.,

$$\begin{aligned} \mathcal{Y}_t^\mathbb{P} &= \mathbb{E}^\mathbb{P}[\mathcal{Y}_\tau^\mathbb{P} | \mathcal{F}_\tau^\mathbb{P}] + \int_t^\tau f(s, X_{\cdot \wedge s}, \mathcal{Y}_s^\mathbb{P}, (\widehat{a}_s^{1/2})^\top \mathcal{Z}_s^\mathbb{P}, \widehat{a}_s, b_s^\mathbb{P}) ds \\ &\quad - \left(\int_t^\tau \mathcal{Z}_s^\mathbb{P} \cdot dX_s^{c, \mathbb{P}} \right)^\mathbb{P} - \int_t^\tau d\widetilde{\mathcal{M}}_s^\mathbb{P}, \end{aligned}$$

where $\widetilde{\mathcal{M}}_\tau^\mathbb{P} := \mathbb{E}^\mathbb{P}[\mathcal{M}_\tau^\mathbb{P} | \mathcal{F}_\tau^\mathbb{P}]$, and $\widetilde{\mathcal{M}}_s^\mathbb{P} := \mathcal{M}_s^\mathbb{P}$ when $s < \tau$. By identification, we deduce that $\widetilde{\mathcal{M}}_\tau^\mathbb{P} = \mathcal{M}_\tau^\mathbb{P} + \mathbb{E}^\mathbb{P}[\mathcal{Y}_\tau^\mathbb{P} | \mathcal{F}_\tau^\mathbb{P}] - \mathcal{Y}_\tau^\mathbb{P}$. Moreover, it is clear that $\widetilde{\mathcal{M}}^\mathbb{P} \in \mathbb{M}_t^p(\mathbb{F}_+^\mathbb{P}, \mathbb{P})$ and $\widetilde{\mathcal{M}}^\mathbb{P}$ is orthogonal to the continuous martingale X under \mathbb{P} .

(ii) Let us now consider the BSDE with generator f and terminal condition $\mathbb{E}^\mathbb{P}[\mathcal{Y}_\tau^\mathbb{P} | \mathcal{F}_\tau^\mathbb{P}]$, on $[t, \tau]$. By uniqueness of the solution to BSDE, it follows that

$$\mathcal{Y}_t^\mathbb{P}(T, \xi) = \mathcal{Y}_t^\mathbb{P}(\tau, \mathcal{Y}_\tau^\mathbb{P}) = \mathcal{Y}_t^\mathbb{P}(\tau, \mathbb{E}^\mathbb{P}[\mathcal{Y}_\tau^\mathbb{P} | \mathcal{F}_\tau]), \quad \mathbb{P}\text{-a.s.} \quad \square$$

PROOF OF THEOREM 2.1. (i) First, by the item (iii) of Assumption 2.1, it is clear that $\widehat{\mathcal{Y}}_t(\omega) = \widehat{\mathcal{Y}}_t(\omega_{t \wedge \cdot})$. Moreover, since $(t, \omega, \mathbb{P}) \mapsto \mathcal{Y}_t^\mathbb{P}(\omega)$ is a Borel measurable map from $[0, T] \times \Omega \times \mathbb{M}_1$ to \mathbb{R} by Lemma 2.6, and the graph $[[\mathcal{P}]]$ is also a Borel measurable in $[0, T] \times \Omega \times \mathbb{M}_1$ by Assumption 2.1, it follows by the measurable selection theorem that $(t, \omega) \mapsto \widehat{\mathcal{Y}}_t(\omega)$ is $\mathcal{B}([0, T]) \otimes \mathcal{F}_T$ -universally measurable (or more precisely upper semi-analytic; see, for example, Proposition 7.47 of Bertsekas and Shreve [6] or Theorem III.82 (page 161) of Dellacherie and Meyer [25]).

(ii) Now, using the measurable selection argument, the DPP is a direct consequence of the comparison principle and the stability of BSDE (2.5). First, for every $\mathbb{P} \in \mathcal{P}_t$, we have

$$\mathcal{Y}_t^\mathbb{P}(T, \xi) = \mathcal{Y}_t^\mathbb{P}(\tau, \mathcal{Y}_\tau^\mathbb{P}) = \mathcal{Y}_t^\mathbb{P}(\tau, \mathbb{E}^\mathbb{P}[\mathcal{Y}_\tau^\mathbb{P} | \mathcal{F}_\tau]), \quad \mathbb{P}\text{-a.s.}$$

It follows by the comparison principle of the BSDE (2.5) (see Lemma A.3 in the Appendix together with Lemma 2.2) that

$$\begin{aligned} \widehat{\mathcal{Y}}_t(\omega) &:= \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \mathbb{E}^{\mathbb{P}}[\mathcal{Y}_t^{\mathbb{P}}(T, \xi)] = \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \mathbb{E}^{\mathbb{P}}[\mathcal{Y}_t^{\mathbb{P}}(\tau, \mathbb{E}^{\mathbb{P}}[\mathcal{Y}_\tau^{\mathbb{P}} | \mathcal{F}_\tau])] \\ &\leq \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \mathcal{Y}_t^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_\tau). \end{aligned}$$

Next, for every $\mathbb{P} \in \mathcal{P}(t, \omega)$ and $\varepsilon > 0$, using the measurable selection theorem (see, e.g., Proposition 7.50 of [6] or Theorem III.82 in [25]), one can choose a family of probability measures $(\mathbb{Q}_w^\varepsilon)_{w \in \Omega}$ such that $w \mapsto \mathbb{Q}_w^\varepsilon$ is \mathcal{F}_τ -measurable, and for \mathbb{P} -a.e. $w \in \Omega$,

$$\mathbb{Q}_w^\varepsilon \in \mathcal{P}(\tau(w), w) \text{ and } \mathbb{E}^{\mathbb{Q}_w^\varepsilon}[\mathcal{Y}_{\tau(w)}^{\mathbb{Q}_w^\varepsilon}(T, \xi)] \geq \widehat{\mathcal{Y}}_{\tau(w)}(w) - \varepsilon.$$

We can then define the concatenated probability $\mathbb{P}^\varepsilon := \mathbb{P} \otimes_\tau \mathbb{Q}^\varepsilon$ so that, by Assumption 2.1(v), $\mathbb{P}^\varepsilon \in \mathcal{P}(t, \omega)$. Finally, using the stability of the solution to BSDE (2.5) in Lemma A.1 (together with Lemma 2.2), it follows that

$$\begin{aligned} \widehat{\mathcal{Y}}_t(\omega) &\geq \mathbb{E}^{\mathbb{P}^\varepsilon}[\mathcal{Y}_t^{\mathbb{P}^\varepsilon}] = \mathbb{E}^{\mathbb{P}^\varepsilon}[\mathcal{Y}_t^{\mathbb{P}^\varepsilon}(\tau, \mathcal{Y}_\tau^{\mathbb{P}^\varepsilon})] = \mathbb{E}^{\mathbb{P}}[\mathcal{Y}_t^{\mathbb{P}}(\tau, \mathbb{E}^{\mathbb{P}^\varepsilon}[\mathcal{Y}_\tau^{\mathbb{P}^\varepsilon} | \mathcal{F}_\tau])] \\ &\geq \mathbb{E}^{\mathbb{P}}[\mathcal{Y}_t^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_\tau)] - C\varepsilon, \end{aligned}$$

for some constant $C > 0$ independent of ε , and hence the other inequality of the DPP holds true by the arbitrariness of $\varepsilon > 0$ as well as that of $\mathbb{P} \in \mathcal{P}(t, \omega)$. \square

2.4.5. *Further discussions.* Notice that the essential arguments to prove the measurability of $\widehat{\mathcal{Y}}_t(\omega)$ is to construct the solution of the BSDE in a measurable way with respect to different probabilities. Then the dynamic programming principle follows directly from the measurable selection theorem together with the comparison and stability of the BSDE. This general approach is not limited to BSDEs with Lipschitz generator. Indeed, the solution of any BSDEs that can be approximated by a countable sequence of Lipschitz BSDEs inherits directly the measurability property. More precisely, we have the following proposition which also applies to specific super-solutions (see Section 2.3 in [37] for a precise definition) of the BSDEs.

PROPOSITION 2.1. *Let $\mathcal{Y}^{\mathbb{P}}$ be the first component of the (minimal) super-solution of a BSDE with possibly non-Lipschitz generator. We have:*

(i) *If there is a family $(\mathcal{Y}^{\mathbb{P}, n})$, which corresponds to the first component of a family of Lipschitz BSDEs, and a family of subsequence $(n_k^{\mathbb{P}})_{k \geq 1}$ such that, $\mathbb{P} \mapsto n_k^{\mathbb{P}}$ is (Borel) measurable, and $\mathcal{Y}^{\mathbb{P}} = \lim_{k \rightarrow \infty} \mathcal{Y}^{\mathbb{P}, n_k^{\mathbb{P}}}$. Then $(s, \omega, \mathbb{P}) \mapsto \mathbb{E}^{\mathbb{P}}[\mathcal{Y}_s^{\mathbb{P}}]$ is a measurable map, and $(t, \omega) \mapsto \widehat{\mathcal{Y}}_t(\omega)$ is $\mathcal{B}([0, T]) \otimes \mathcal{F}_T$ -universally measurable.*

(ii) *Furthermore, if the (possibly non-Lipschitz) BSDE for $\mathcal{Y}^{\mathbb{P}}$ admits the comparison principle and the stability result w.r.t. the terminal conditions, then for all $(t, \omega) \in [0, T] \times \Omega$ and \mathbb{F} -stopping times τ taking value in $[t, T]$, we have*

$$\widehat{\mathcal{Y}}_t(\omega) = \sup_{\mathbb{P} \in \mathcal{P}(t, \omega)} \mathbb{E}^{\mathbb{P}}[\mathcal{Y}_t^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau})].$$

In particular, this result can be applied to BSDEs with linear growth [59], to BSDEs with general growth in y [73], to quadratic BSDEs [3, 54, 55], to BSDEs with unbounded horizon [23], to reflected BSDEs [36], to constrained BSDEs [19, 75] (for the point (i) only), etc.

REMARK 2.3. In Assumption 2.1, the terminal condition $\xi : \Omega \rightarrow \mathbb{R}$ is assumed to be Borel measurable, which is more restrictive comparing to the results in the context of controlled diffusion/jump process problems (where ξ is only assumed to be upper semi-analytic; see, e.g., [65] or [40]). This Borel measurability condition is however crucial in our BSDE context. For example, when $f(t, \omega, y, z, a, b) = |z|$, we know the solution of the BSDE (2.5) is given by $\inf_{\tilde{\mathbb{P}} \in \tilde{\mathcal{P}}} \mathbb{E}^{\tilde{\mathbb{P}}}[\xi]$ for some family $\tilde{\mathcal{P}}$ of probability measure equivalent to \mathbb{P} . However, as is well known, the upper semi-analytic property is stable by taking supremum but not by taking infimum.

3. Path regularization of the value function. In this section, we will characterize a càdlàg modification of the value function $\widehat{\mathcal{Y}}$ as a semimartingale under any $\mathbb{P} \in \mathcal{P}_0$ and give its decomposition. In the next section, we will show that this càdlàg modification of the value function $\widehat{\mathcal{Y}}$ is the solution to some second-order BSDE defined there (see Definition 4.1). First of all, we recall from Theorem 2.1 that we have for any \mathbb{F} -stopping times $\tau \geq \sigma$,

$$(3.1) \quad \widehat{\mathcal{Y}}_{\sigma(\omega)}(\omega) = \sup_{\mathbb{P} \in \mathcal{P}(\sigma(\omega), \omega)} \mathbb{E}^{\mathbb{P}}[\mathcal{Y}_{\sigma(\omega)}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau})].$$

Moreover, we also have

$$(3.2) \quad \widehat{\mathcal{Y}}_{\sigma(\omega)}(\omega) = \sup_{\mathbb{P} \in \mathcal{P}(\sigma(\omega), \omega)} \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0}[\widetilde{\mathcal{Y}}_{\sigma(\omega)}^{\mathbb{P} \otimes \mathbb{P}_0}(\tau, \widehat{\mathcal{Y}}_{\tau})],$$

where $\widetilde{\mathcal{Y}}_{\sigma(\omega)}^{\mathbb{P} \otimes \mathbb{P}_0}(\tau, \widehat{\mathcal{Y}}_{\tau})$ is the equivalent of $\mathcal{Y}_{\sigma(\omega)}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau})$ but defined on the enlarged space, recall (2.9) and Lemma 2.2.

Let us start with the following technical lemma. Formally, it can be obtained by simply taking conditional expectations of the corresponding BSDEs. However, this raises subtle problems about negligible sets and conditional probability measures. We therefore refer the reader to [16] for the precise details.

LEMMA 3.1. For any $\mathbb{P} \in \mathcal{P}_0$, for any \mathbb{F} -stopping times $0 \leq \sigma \leq \tau \leq T$, we have

$$\begin{aligned} & \mathbb{E}^{\mathbb{P}_\omega^{\sigma(\omega)} \otimes \mathbb{P}_0} [\widehat{\mathcal{Y}}_{\sigma(\omega)}^{\mathbb{P}_\omega^{\sigma(\omega)} \otimes \mathbb{P}_0}(\tau, \widehat{\mathcal{Y}}_\tau)] \\ &= \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} [\widehat{\mathcal{Y}}_{\sigma}^{\mathbb{P} \otimes \mathbb{P}_0}(\tau, \widehat{\mathcal{Y}}_\tau) | \overline{\mathcal{F}}_\sigma](\omega, \omega'), \quad \mathbb{P} \otimes \mathbb{P}_0\text{-a.e. on } \overline{\Omega}, \\ & \mathbb{E}^{\mathbb{P}_\omega^{\sigma(\omega)}} [\mathcal{Y}_{\sigma(\omega)}^{\mathbb{P}_\omega^{\sigma(\omega)}}(\tau, \widehat{\mathcal{Y}}_\tau)] \\ &= \mathbb{E}^{\mathbb{P}} [\mathcal{Y}_{\sigma}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_\tau) | \mathcal{F}_\sigma](\omega), \quad \text{for } \mathbb{P}\text{-a.e. } \omega \in \Omega. \end{aligned}$$

Let us next remark the following immediate consequences of (3.1) and (3.2):

$$(3.3) \quad \widehat{\mathcal{Y}}_{\sigma(\omega)} \geq \mathbb{E}^{\mathbb{P}} [\mathcal{Y}_{\sigma(\omega)}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_\tau)], \quad \text{for any } \mathbb{P} \in \mathcal{P}(\sigma(\omega), \omega),$$

$$(3.4) \quad \widehat{\mathcal{Y}}_{\sigma(\omega)} \geq \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} [\widehat{\mathcal{Y}}_{\sigma}^{\mathbb{P} \otimes \mathbb{P}_0}(\tau, \widehat{\mathcal{Y}}_\tau)], \quad \text{for any } \mathbb{P} \in \mathcal{P}(\sigma(\omega), \omega).$$

With these inequalities, we can prove a down-crossing inequality for $\widehat{\mathcal{Y}}$, which ensures that $\widehat{\mathcal{Y}}$ admits right- and left-limits outside a \mathcal{P}_0 -polar set. Recall that

$$\widehat{f}_s^{\mathbb{P}}(y, z) := f(s, X_{\cdot \wedge s}, y, z, \widehat{a}_s, b_s^{\mathbb{P}}), \quad \widehat{f}_s^{\mathbb{P}, 0} := f(s, X_{\cdot \wedge s}, 0, 0, \widehat{a}_s, b_s^{\mathbb{P}}).$$

Let $J := (\tau_n)_{n \in \mathbb{N}}$ be a countable family of \mathbb{F} -stopping times taking values in $[0, T]$ such that for any $(i, j) \in \mathbb{N}^2$, one has either $\tau_i \leq \tau_j$, or $\tau_i \geq \tau_j$, for every $\omega \in \Omega$. Let $a > b$ and $J_n \subset J$ be a finite subset ($J_n = \{0 \leq \tau_1 \leq \dots \leq \tau_n \leq T\}$). We denote by $D_a^b(\widehat{\mathcal{Y}}, J_n)$ the number of down-crossings of the process $(\widehat{\mathcal{Y}}_{\tau_k})_{1 \leq k \leq n}$ from b to a . We then define

$$D_a^b(\widehat{\mathcal{Y}}, J) := \sup\{D_a^b(\widehat{\mathcal{Y}}, J_n) : J_n \subset J, \text{ and } J_n \text{ is a finite set}\}.$$

The following lemma follows very closely the related result proved in Lemma A.1 of [12]. However, since $\widehat{\mathcal{Y}}$ is not exactly an $\mathcal{E}^{\widehat{f}^{\mathbb{P}}}$ -supermartingale in their terminology, we give a short proof.

LEMMA 3.2. Fix some $\mathbb{P} \in \mathcal{P}_0$ and let Assumption 2.1 hold. Denote by L the Lipschitz constant of the generator f . Then, for all $a < b$, there exists a probability measure $\overline{\mathbb{Q}}$, equivalent to $\mathbb{P} \otimes \mathbb{P}_0$, such that

$$\begin{aligned} & \mathbb{E}^{\overline{\mathbb{Q}}} [D_a^b(\widehat{\mathcal{Y}}, J)] \\ & \leq \frac{e^{LT}}{b-a} \mathbb{E}^{\overline{\mathbb{Q}}} \left[e^{LT} (\widehat{\mathcal{Y}}_0 \wedge b - a) - e^{-LT} (\widehat{\mathcal{Y}}_T \wedge b - a)^+ + e^{LT} (\widehat{\mathcal{Y}}_T \wedge b - a)^- \right. \\ & \quad \left. + e^{LT} \int_0^T |\widehat{f}_s^{\mathbb{P}}(a, 0)| ds \right]. \end{aligned}$$

Moreover, outside a \mathcal{P}_0 -polar set, we have

$$\begin{aligned} \overline{\lim}_{r \in \mathbb{Q} \cap (t, T], r \downarrow t} \widehat{\mathcal{Y}}_r(\omega) &= \lim_{r \in \mathbb{Q} \cap (t, T], r \downarrow t} \widehat{\mathcal{Y}}_r(\omega), \quad \text{and} \\ \overline{\lim}_{r \in \mathbb{Q} \cap (t, T], r \uparrow t} \widehat{\mathcal{Y}}_r(\omega) &= \lim_{r \in \mathbb{Q} \cap (t, T], r \uparrow t} \widehat{\mathcal{Y}}_r(\omega). \end{aligned}$$

PROOF. Without loss of generality, we can always suppose that 0 and T belong to J and $b > a = 0$. Let $J_n = \{\tau_0, \tau_1, \dots, \tau_n\}$ with $0 = \tau_0 < \tau_1 < \dots < \tau_n = T$. We then consider for any $i = 1, \dots, n$, and any $\omega \in \Omega$, the following BSDE in the enlarged space under $\overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)} := \mathbb{P}_\omega^{\tau_{i-1}(\omega)} \otimes \mathbb{P}_0$ on $[\tau_{i-1}, \tau_i]$:

$$\begin{aligned} \widehat{\mathcal{Y}}_t^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)} &:= \widehat{\mathcal{Y}}_{\tau_i} - \int_t^{\tau_i} (\widehat{f}_s^{\mathbb{P}_\omega^{\tau_{i-1}(\omega)}}, 0 + \lambda_s^i \widehat{\mathcal{Y}}_s^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)} + \eta_s^i \cdot (\widehat{a}_s^{1/2})^\top \widehat{\mathcal{Z}}_s^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)}) ds \\ &\quad - \int_t^{\tau_i} \widehat{\mathcal{Z}}_s^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)} \cdot \widehat{a}_s^{1/2} dW_s^{\mathbb{P}_\omega^{\tau_{i-1}(\omega)}} - \int_t^{\tau_i} d\widehat{\mathcal{M}}_s^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)}, \\ &\quad \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)}\text{-a.s.}, \end{aligned}$$

where λ^i and η^i are the two bounded processes appearing in the linearization of \widehat{f} [recall Assumption 2.1(i)]. We consider then the linear BSDE, also on the enlarged space:

$$\begin{aligned} \overline{\mathcal{Y}}_t^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)} &:= \widehat{\mathcal{Y}}_{\tau_i} - \int_t^{\tau_i} ((\widehat{f}_s^{\mathbb{P}_\omega^{\tau_{i-1}(\omega)}}, 0) + \lambda_s^i \overline{\mathcal{Y}}_s^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)} + \eta_s^i \cdot (\widehat{a}_s^{1/2})^\top \overline{\mathcal{Z}}_s^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)}) ds \\ &\quad - \int_t^{\tau_i} \overline{\mathcal{Z}}_s^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)} \cdot \widehat{a}_s^{1/2} dW_s^{\mathbb{P}_\omega^{\tau_{i-1}(\omega)}} - \int_t^{\tau_i} d\overline{\mathcal{M}}_s^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)}, \\ &\quad \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)}\text{-a.s.} \end{aligned}$$

It is immediate that

$$\begin{aligned} \overline{\mathcal{Y}}_{\tau_{i-1}}^i, \overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)} &= \mathbb{E}^{\overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)}} \left[L_{\tau_i} \left(\widehat{\mathcal{Y}}_{\tau_i} e^{\int_{\tau_{i-1}}^{\tau_i} \lambda_s^i ds} \right. \right. \\ &\quad \left. \left. - \int_{\tau_{i-1}}^{\tau_i} e^{\int_{\tau_{i-1}}^s \lambda_r^i dr} \widehat{f}_s^{\mathbb{P}_\omega^{\tau_{i-1}(\omega)}}, 0 \Big| \mathcal{F}_{\tau_{i-1}}^+ \right) \right], \end{aligned}$$

where

$$L_t := \mathcal{E} \left(\int_{\tau_{i-1}}^t \eta_s^i \cdot dW_s^{\overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)}} \right), \quad t \in [\tau_{i-1}, \tau_i].$$

By Assumption 2.1(iv), for \mathbb{P} -a.e. $\omega \in \Omega$, $\overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)} \in \mathcal{P}(\tau_{i-1}(\omega), \omega)$. Hence, by the comparison principle for BSDEs, recalled in Lemma A.3 below, and (3.2), it is clear that

$$\begin{aligned} \mathbb{E}^{\overline{\mathbb{P}}_\omega^{\tau_{i-1}(\omega)}} \left[L_{\tau_i} \left(\widehat{\mathcal{Y}}_{\tau_i} e^{\int_{\tau_{i-1}}^{\tau_i} \lambda_s^i ds} - \int_{\tau_{i-1}}^{\tau_i} e^{\int_{\tau_{i-1}}^s \lambda_r^i dr} \widehat{f}_s^{\mathbb{P}_\omega^{\tau_{i-1}(\omega)}}, 0 \Big| \mathcal{F}_{\tau_{i-1}}^+ \right) \right] \\ \leq \widehat{\mathcal{Y}}_{\tau_{i-1}}(\omega). \end{aligned}$$

But, by the definition of the r.p.c.d., this implies that

$$\mathbb{E}^{\overline{\mathbb{Q}}}\left[\widehat{\mathcal{Y}}_{\tau_i} e^{\int_{\tau_{i-1}}^{\tau_i} \lambda_s^i ds} - \int_{\tau_{i-1}}^{\tau_i} e^{\int_{\tau_{i-1}}^s \lambda_r^i dr} \left| \widehat{f}_s^{\mathbb{P}, \tau_{i-1}(\cdot), 0} \right| ds \middle| \overline{\mathcal{F}}_{\tau_{i-1}}\right] \leq \widehat{\mathcal{Y}}_{\tau_{i-1}},$$

$\mathbb{P} \otimes \mathbb{P}_0$ -a.s.,

where the probability measure $\overline{\mathbb{Q}}$ is equivalent to $\mathbb{P} \otimes \mathbb{P}_0$ and defined by

$$\frac{d\overline{\mathbb{Q}}}{d\mathbb{P} \otimes \mathbb{P}_0} := \mathcal{E}\left(\int_{\tau_{i-1}}^t \eta_s^i \cdot dW_s^{\mathbb{P}}\right), \quad t \in [\tau_{i-1}, \tau_i].$$

Let $\lambda_s := \sum_{i=1}^n \lambda_s^i \mathbf{1}_{[\tau_{i-1}, \tau_i)}(s)$, then one has that the discrete process $(V_{\tau_i})_{0 \leq i \leq n}$ defined by

$$V_{\tau_i} := \widehat{\mathcal{Y}}_{\tau_i} e^{\int_0^{\tau_i} \lambda_s ds} - \int_0^{\tau_i} e^{\int_0^s \lambda_r dr} \left| \widehat{f}_s^{\mathbb{P}, 0} \right| ds,$$

is a $\overline{\mathbb{Q}}$ -supermartingale relative to $\overline{\mathbb{F}}$. Then the control on the down-crossings can be obtained exactly as in the proof of Lemma A.1 in [12]. Indeed, it is enough to observe that the original down-crossing inequality for supermartingales (see, e.g., [28], page 446) does not require the filtration to satisfy the usual assumptions. We now prove the second part of the lemma. We define the set

$$\Sigma := \{\omega \in \Omega \text{ s.t. } \widehat{\mathcal{Y}}(\omega) \text{ has no right- or left-limits along the rationals}\}.$$

We claim that Σ is a \mathcal{P}_0 -polar set. Indeed, suppose that there exists $\mathbb{P} \in \mathcal{P}_0$ satisfying $\mathbb{P}(\Sigma) > 0$. Then, Σ is nonempty and for any $\omega \in \Sigma$, the path $\widehat{\mathcal{Y}}(\omega)$ has, for example, no right-limit along the rationals at some point $t \in [0, T]$. We can therefore find two rational numbers a, b such that

$$\liminf_{r \in \mathbb{Q} \cap (t, T], r \downarrow t} \widehat{\mathcal{Y}}_r(\omega) < a < b < \limsup_{r \in \mathbb{Q} \cap (t, T], r \downarrow t} \widehat{\mathcal{Y}}_r(\omega),$$

and the number of down-crossings $D_a^b(\widehat{\mathcal{Y}}, J)(\omega)$ of the path $\widehat{\mathcal{Y}}(\omega)$ on the interval $[a, b]$ is equal to $+\infty$. However, the down-crossing inequality proved above shows that $D_a^b(\widehat{\mathcal{Y}}, J)$ is $\overline{\mathbb{Q}}$ -a.s., and thus \mathbb{P} -a.s. (see Lemma 2.1), finite, for any pair (a, b) . This implies a contradiction since we assumed that $\mathbb{P}(\Sigma) > 0$. Therefore, outside the \mathcal{P}_0 -polar set Σ , $\widehat{\mathcal{Y}}$ admits both right- and left-limits along the rationals. \square

We next define, for all $(t, \omega) \in [0, T) \times \Omega$,

$$(3.5) \quad \widehat{\mathcal{Y}}_t^+ := \limsup_{r \in \mathbb{Q} \cap (t, T], r \downarrow t} \widehat{\mathcal{Y}}_r,$$

and $\widehat{\mathcal{Y}}_T^+ := \mathcal{Y}_T$. By Lemma 3.2,

$$\widehat{\mathcal{Y}}_t^+ := \lim_{r \in \mathbb{Q} \cap (t, T], r \downarrow t} \widehat{\mathcal{Y}}_r, \quad \text{outside a } \mathcal{P}_0\text{-polar set,}$$

and we deduce that $\widehat{\mathcal{Y}}^+$ is càd outside a \mathcal{P}_0 -polar set. Hence, since for any $t \in [0, T]$, $\widehat{\mathcal{Y}}_t^+$ is by definition $\mathcal{F}_t^{U^+}$ -measurable, we deduce that $\widehat{\mathcal{Y}}^+$ is actually $\mathbb{F}^{\mathcal{P}_0^+}$ -optional. Our next result extends (3.3) to $\widehat{\mathcal{Y}}^+$.

LEMMA 3.3. For any $0 \leq s \leq t \leq T$, for any $\mathbb{P} \in \mathcal{P}_0$, we have

$$\widehat{\mathcal{Y}}_s^+ \geq \mathcal{Y}_s^\mathbb{P}(t, \widehat{\mathcal{Y}}_t^+), \quad \mathbb{P}\text{-a.s.}$$

PROOF. Fix some $(s, t, \omega) \in [0, T] \times [s, T] \times \Omega$ and some $\mathbb{P} \in \mathcal{P}_0$. Let $r_n^1 \in \mathbb{Q} \cap (s, T]$, $r_n^1 \downarrow s$ and $r_n^2 \in \mathbb{Q} \cap (t, T]$, $r_n^2 \downarrow t$. By (3.3), we have for any $m, n \geq 1$ and $\widetilde{\mathbb{P}} \in \mathcal{P}(r_n^1, \omega)$:

$$\widehat{\mathcal{Y}}_{r_n^1}(\omega) \geq \mathbb{E}^{\widetilde{\mathbb{P}}}[\mathcal{Y}_{r_n^1}^{\widetilde{\mathbb{P}}}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2})].$$

In particular, thanks to Assumption 2.1(iv), for \mathbb{P} -a.e. $\omega \in \Omega$, we have

$$(3.6) \quad \widehat{\mathcal{Y}}_{r_n^1}(\omega) \geq \mathbb{E}^{\mathbb{P}'_{\omega}}[\mathcal{Y}_{r_n^1}^{\mathbb{P}'_{\omega}}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2})] = \mathbb{E}^\mathbb{P}[\mathcal{Y}_{r_n^1}^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}) | \mathcal{F}_{r_n^1}](\omega),$$

where we have used Lemma 3.1. By definition, we have

$$\lim_{n \rightarrow +\infty} \widehat{\mathcal{Y}}_{r_n^1} = \widehat{\mathcal{Y}}_s^+, \quad \mathbb{P}\text{-a.s.}$$

Next, we want to show that

$$\mathbb{E}^\mathbb{P}[\mathcal{Y}_{r_n^1}^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}) | \mathcal{F}_{r_n^1}] \xrightarrow{n \rightarrow +\infty} \mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}), \quad \text{for the norm } \|\cdot\|_{\mathbb{L}_{s,\omega}^1}.$$

Indeed, we have

$$\begin{aligned} & \mathbb{E}^\mathbb{P}[|\mathbb{E}^\mathbb{P}[\mathcal{Y}_{r_n^1}^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}) | \mathcal{F}_{r_n^1}] - \mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2})|] \\ &= \mathbb{E}^\mathbb{P}[|\mathbb{E}^\mathbb{P}[\mathcal{Y}_{r_n^1}^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}) - \mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}) | \mathcal{F}_{r_n^1}]|] \\ &\leq \mathbb{E}^\mathbb{P}[\mathbb{E}^\mathbb{P}[|\mathcal{Y}_{r_n^1}^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}) - \mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2})| | \mathcal{F}_{r_n^1}]] \\ &= \mathbb{E}^\mathbb{P}[|\mathcal{Y}_{r_n^1}^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}) - \mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2})|]. \end{aligned}$$

Then, since $\mathcal{Y}_{r_n^1}^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2})$ is càdlàg, we know that $\mathcal{Y}_{r_n^1}^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2})$ converges, \mathbb{P} -a.s., to $\mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2})$, as n goes to $+\infty$. Moreover, by the estimates of Lemma A.1 (together with Lemma 2.2), the quantity in the expectation above is uniformly bounded in $\mathbb{L}^p(\mathcal{F}_T^{\mathbb{P}^+}, \mathbb{P})$ and, therefore, forms a uniformly integrable family by de la Vallée–Poussin criterion (since $p > 1$). Therefore, the desired convergence is a simple consequence of the dominated convergence theorem. Hence, taking a subsequence if necessary, we have that the right-hand side of (3.6) goes \mathbb{P} -a.s. to $\mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2})$ as n goes to $+\infty$, so that we have

$$\widehat{\mathcal{Y}}_s^+ \geq \mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}), \quad \mathbb{P}\text{-a.s.}$$

Next, we have by the dynamic programming for BSDEs

$$\begin{aligned} & \mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}) - \mathcal{Y}_s^\mathbb{P}(t, \widehat{\mathcal{Y}}_t^+) \\ &= \mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_{r_m^2}) - \mathcal{Y}_s^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_t^+) + \mathcal{Y}_s^\mathbb{P}(t, \mathcal{Y}_t^\mathbb{P}(r_m^2, \widehat{\mathcal{Y}}_t^+)) - \mathcal{Y}_s^\mathbb{P}(t, \widehat{\mathcal{Y}}_t^+). \end{aligned}$$

The first difference on the right-hand side converges to 0, \mathbb{P} -a.s., once more thanks to the estimates of Lemma A.1 (together with Lemma 2.2) and the definition of $\widehat{\mathcal{Y}}^+$. As for the second difference, the same estimates show that it is controlled by

$$\mathbb{E}^{\mathbb{P}}[|\widehat{\mathcal{Y}}_t^+ - \mathcal{Y}_t^{\mathbb{P}}(r_m^2, \widehat{\mathcal{Y}}_t^+)|^{\tilde{p}} | \mathcal{F}_s],$$

for some $1 < \tilde{p} < p$. This term goes \mathbb{P} -a.s. (at least along a subsequence) to 0 as m goes to $+\infty$ as well by Lemma A.2 (together with Lemma 2.2), which completes the proof. \square

The next lemma follows the classical proof of the optional sampling theorem for càdlàg supermartingales and extends the previous result to stopping times.

LEMMA 3.4. *For any \mathbb{F} -stopping times $0 \leq \sigma \leq \tau \leq T$, for any $\mathbb{P} \in \mathcal{P}_0$, we have*

$$\widehat{\mathcal{Y}}_{\sigma}^+ \geq \mathcal{Y}_{\sigma}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau}^+), \quad \mathbb{P}\text{-a.s.}$$

In particular $\widehat{\mathcal{Y}}^+$ is càdlàg, \mathcal{P}_0 -q.s.

PROOF. Assume first that σ takes a finite number of values $\{t_1, \dots, t_n\}$ and that τ is deterministic. Then we have for any $\mathbb{P} \in \mathcal{P}_0$

$$\widehat{\mathcal{Y}}_{\sigma}^+ = \sum_{i=1}^n \widehat{\mathcal{Y}}_{t_i}^+ \mathbf{1}_{\{\sigma=t_i\}} \geq \sum_{i=1}^n \mathcal{Y}_{t_i}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau}^+) \mathbf{1}_{\{\sigma=t_i\}} = \mathcal{Y}_{\sigma}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau}^+), \quad \mathbb{P}\text{-a.s.}$$

Assume next that both τ and σ take a finite number of values $\{t_1, \dots, t_n\}$. We have similarly

$$\mathcal{Y}_{\sigma}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau}^+) = \sum_{i=1}^n \mathcal{Y}_{\sigma}^{\mathbb{P}}(t_i, \widehat{\mathcal{Y}}_{t_i}^+) \mathbf{1}_{\{\tau=t_i\}} \leq \sum_{i=1}^n \widehat{\mathcal{Y}}_{\sigma}^+ \mathbf{1}_{\{\tau=t_i\}} = \widehat{\mathcal{Y}}_{\sigma}^+, \quad \mathbb{P}\text{-a.s.}$$

Then, if σ is general, we can always approach it from above by a decreasing sequence of \mathbb{F}^+ -stopping times $(\sigma^n)_{n \geq 1}$ taking only a finite number of values. The above results imply directly that

$$\widehat{\mathcal{Y}}_{\sigma^n \wedge \tau}^+ \geq \mathcal{Y}_{\sigma^n \wedge \tau}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau}^+), \quad \mathbb{P}\text{-a.s.}$$

Then we can use the right-continuity of $\widehat{\mathcal{Y}}^+$ and $\mathcal{Y}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau}^+)$ to let n go to $+\infty$ and obtain

$$\widehat{\mathcal{Y}}_{\sigma}^+ \geq \mathcal{Y}_{\sigma}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau}^+), \quad \mathbb{P}\text{-a.s.}$$

Finally, let us take a general stopping time τ . We once more approach it by a decreasing sequence of \mathbb{F}^+ -stopping times $(\tau^n)_{n \geq 1}$ taking only a finite number of values. We thus have

$$\widehat{\mathcal{Y}}_{\sigma}^+ \geq \mathcal{Y}_{\sigma}^{\mathbb{P}}(\tau^n, \widehat{\mathcal{Y}}_{\tau^n}^+), \quad \mathbb{P}\text{-a.s.}$$

The term on the right-hand side converges (along a subsequence if necessary) \mathbb{P} -a.s. to $\mathcal{Y}_\sigma^\mathbb{P}(\tau, \widehat{\mathcal{Y}}_\tau^+)$ by Lemma A.2 (together with Lemma 2.2).

It remains to justify that $\widehat{\mathcal{Y}}^+$ admits left-limits outside a \mathcal{P}_0 -polar set. Fix some $\mathbb{P} \in \mathcal{P}_0$. Following the same arguments as in the proof of Lemma 3.2, we can show that for some probability measure $\overline{\mathbb{Q}}$ equivalent to $\mathbb{P} \otimes \mathbb{P}_0$ and some bounded process λ ,

$$V_t := \widehat{\mathcal{Y}}_t e^{\int_0^t \lambda_s ds} - \int_0^t e^{\int_0^s \lambda_u du} |\widehat{f}_s^{\mathbb{P},0}| ds,$$

is a right-continuous $(\overline{\mathbb{Q}}, \overline{\mathbb{F}}_+)$ -supermartingale, which is in addition uniformly integrable under $\overline{\mathbb{Q}}$ since $\widehat{\mathcal{Y}}$ and $\widehat{f}^{\mathbb{P},0}$ are uniformly bounded in $\mathbb{L}^p(\overline{\mathcal{F}}_T, \mathbb{P} \otimes \mathbb{P}_0)$, and thus in $\mathbb{L}^{\tilde{p}}(\overline{\mathcal{F}}_T, \overline{\mathbb{Q}})$ for some $1 < \tilde{p} < p$. Therefore, for any increasing sequence of $\overline{\mathbb{F}}^+$ -stopping times $(\rho_n)_{n \geq 0}$ taking values in $[0, T]$, the sequence $(\mathbb{E}^{\overline{\mathbb{Q}}}[V_{\rho_n}])_{n \geq 0}$ is nonincreasing and admits a limit. By Theorem VI-48 and Remark VI-50(f) of [25], we deduce that V , and thus $\widehat{\mathcal{Y}}^+$, admit left-limits outside a $\overline{\mathbb{Q}}$ -negligible (and thus \mathbb{P} -negligible by Lemma 2.1) set. Moreover, the above implies that the set

$$\{\omega \in \Omega : \widehat{\mathcal{Y}}^+(\omega) \text{ admits left-limits}\},$$

is \mathcal{P}_0 -polar, which completes the proof. \square

Our next result shows that $\widehat{\mathcal{Y}}^+$ satisfies the representation formula (4.3). Part of it requires the following stronger integrability assumption.

ASSUMPTION 3.1. There is some $\kappa \in (1, p]$ such that, $\xi \in \mathbb{L}_0^{p,\kappa}(\mathbb{F})$,

$$\phi_f^{p,\kappa} := \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^\mathbb{P} \left[\text{ess sup}_{0 \leq t \leq T}^\mathbb{P} \left(\text{ess sup}_{\mathbb{P}' \in \mathcal{P}_0(t, \mathbb{P}, \mathbb{F}_+)}^\mathbb{P} \mathbb{E}^{\mathbb{P}'} \left[\int_0^T |\widehat{f}_s^{\mathbb{P}',0}|^\kappa ds \mid \mathcal{F}_t^+ \right] \right)^\frac{p}{\kappa} \right] < +\infty.$$

LEMMA 3.5. For any \mathbb{F} -stopping times $0 \leq \sigma \leq \tau \leq T$, for any $0 \leq t \leq T$, and for any $\mathbb{P} \in \mathcal{P}_0$, we have

$$\begin{aligned} \widehat{\mathcal{Y}}_\sigma &= \text{ess sup}_{\mathbb{P}' \in \mathcal{P}_0(\sigma, \mathbb{P}, \mathbb{F})}^\mathbb{P} \mathbb{E}^{\mathbb{P}'}[\mathcal{Y}_\sigma^{\mathbb{P}'}(\tau, \mathcal{Y}_\tau^{\mathbb{P}'}) \mid \mathcal{F}_\sigma], & \mathbb{P}\text{-a.s.}, & \text{ and} \\ \widehat{\mathcal{Y}}_t^+ &= \text{ess sup}_{\mathbb{P}' \in \mathcal{P}_0(t, \mathbb{P}, \mathbb{F}_+)}^\mathbb{P} \mathcal{Y}_t^{\mathbb{P}'}(T, \xi), & \mathbb{P}\text{-a.s.}, \end{aligned}$$

where $\mathcal{P}_0(\sigma, \mathbb{P}, \mathbb{F})$ is defined in Section 2.2. In particular, if Assumption 3.1 holds, one has $\widehat{\mathcal{Y}}^+ \in \mathbb{D}_0^p(\mathbb{F}^{\mathcal{P}_0^+})$.

PROOF. We start with the first equality. By definition and Lemma 3.1, for any $\mathbb{P}' \in \mathcal{P}_0(\sigma, \mathbb{P}, \mathbb{F})$ we have

$$\widehat{\mathcal{Y}}_\sigma \geq \mathbb{E}^{\mathbb{P}'}[\mathcal{Y}_\sigma^{\mathbb{P}'}(\tau, \mathcal{Y}_\tau^{\mathbb{P}'}) \mid \mathcal{F}_\sigma], \quad \mathbb{P}'\text{-a.s.}$$

But since both sides of the inequality are \mathcal{F}_σ^U -measurable and \mathbb{P}' coincides with \mathbb{P} on \mathcal{F}_σ (and thus on \mathcal{F}_σ^U , by uniqueness of universal completion) the above also holds \mathbb{P} -a.s. We deduce

$$\widehat{\mathcal{Y}}_\sigma \geq \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(\sigma, \mathbb{P}, \mathbb{F})} \mathbb{E}^{\mathbb{P}'} [\mathcal{Y}_\sigma^{\mathbb{P}'}(\tau, \mathcal{Y}_\tau^{\mathbb{P}'}) | \mathcal{F}_\sigma], \quad \mathbb{P}\text{-a.s.}$$

Next, notice that by Lemmas 2.6 and 2.7, $(t, \omega, \mathbb{Q}) \mapsto \mathbb{E}^{\mathbb{Q}}[\mathcal{Y}_t^{\mathbb{Q}}(T, \xi)] = \mathbb{E}^{\mathbb{Q}}[\mathcal{Y}_t^{\mathbb{Q}}(\tau, \mathcal{Y}_\tau^{\mathbb{Q}})]$ is Borel measurable. As in the proof of Theorem 2.1, it follows by the measurable selection theorem (see, e.g., Proposition 7.47 of [6]) that for every $\varepsilon > 0$, there is a family of probability measures $(\mathbb{Q}_w^\varepsilon)_{w \in \Omega}$ such that $w \mapsto \mathbb{Q}_w^\varepsilon$ is \mathcal{F}_σ measurable and for \mathbb{P} -a.e. $w \in \Omega$,

$$\widehat{\mathcal{Y}}_{\sigma(w)}(w) \leq \mathbb{E}^{\mathbb{Q}_w^\varepsilon}[\mathcal{Y}_{\sigma(w)}^{\mathbb{Q}_w^\varepsilon}] + \varepsilon, \quad \mathbb{P}\text{-a.s.}$$

Let us now define the concatenated probability $\mathbb{P}^\varepsilon := \mathbb{P} \otimes_\sigma \mathbb{Q}^\varepsilon$ so that $\mathbb{P}^\varepsilon \in \mathcal{P}_0(\sigma, \mathbb{P}, \mathbb{F})$, it follows then by Lemma 3.1 that

$$\mathcal{Y}_\sigma \leq \mathbb{E}^{\mathbb{P}^\varepsilon}[\mathcal{Y}_\sigma^{\mathbb{P}^\varepsilon}(\tau, \mathcal{Y}_\tau^{\mathbb{P}^\varepsilon}) | \mathcal{F}_\sigma] + \varepsilon \leq \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(\sigma, \mathbb{P}, \mathbb{F})} \mathbb{E}^{\mathbb{P}'}[\mathcal{Y}_\sigma^{\mathbb{P}'}(\tau, \mathcal{Y}_\tau^{\mathbb{P}'}) | \mathcal{F}_\sigma] + \varepsilon, \quad \mathbb{P}\text{-a.s.}$$

We hence complete the proof of the first equality by arbitrariness of $\varepsilon > 0$.

Let us now prove the second equality. Let $r_n^1 \in \mathbb{Q} \cap (t, T]$, $r_n^1 \downarrow t$. By the first part of the proof, we have

$$\widehat{\mathcal{Y}}_{r_n^1} = \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(r_n^1, \mathbb{P}, \mathbb{F})} \mathbb{E}^{\mathbb{P}'}[\mathcal{Y}_{r_n^1}^{\mathbb{P}'}(T, \xi) | \mathcal{F}_{r_n^1}], \quad \mathbb{P}\text{-a.s.}$$

Since for every $n \in \mathbb{N}$, $\mathcal{P}_0(r_n^1, \mathbb{P}, \mathbb{F}) \subset \mathcal{P}_0(t, \mathbb{P}, \mathbb{F}_+)$, we deduce as above that for any $\mathbb{P}' \in \mathcal{P}_0(t, \mathbb{P}, \mathbb{F}_+)$ and for n large enough

$$\widehat{\mathcal{Y}}_{r_n^1} \geq \mathbb{E}^{\mathbb{P}'}[\mathcal{Y}_{r_n^1}^{\mathbb{P}'}(T, \xi) | \mathcal{F}_{r_n^1}], \quad \mathbb{P}\text{-a.s.}$$

Arguing exactly as in the proof of Lemma 3.3, we can let n go to $+\infty$ to obtain

$$\widehat{\mathcal{Y}}_t^+ \geq \mathcal{Y}_t^{\mathbb{P}'}(T, \xi), \quad \mathbb{P}\text{-a.s.},$$

which implies by arbitrariness of \mathbb{P}'

$$\widehat{\mathcal{Y}}_t^+ \geq \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(t, \mathbb{P}, \mathbb{F}_+)} \mathcal{Y}_t^{\mathbb{P}'}(T, \xi), \quad \mathbb{P}\text{-a.s.}$$

We claim next that for any $n \in \mathbb{N}$, the following family is upward directed:

$$\{\mathbb{E}^{\mathbb{P}'}[\mathcal{Y}_{r_n^1}^{\mathbb{P}'}(T, \mathcal{Y}_T^{\mathbb{P}'}) | \mathcal{F}_{r_n^1}], \mathbb{P}' \in \mathcal{P}_0(r_n^1, \mathbb{P}, \mathbb{F})\}.$$

Indeed, this can be proved exactly as in Step 2 of the proof of Theorem 4.2. According to [67], we then know that there exists some sequence $(\mathbb{P}_n^m)_{m \geq 0} \subset$

$\mathcal{P}_0(r_n^1, \mathbb{P}, \mathbb{F})$ such that

$$\widehat{\mathcal{Y}}_{r_n^1} = \lim_{m \rightarrow +\infty} \uparrow \mathbb{E}^{\mathbb{P}^m} [\mathcal{Y}_{r_n^1}^{\mathbb{P}^m}(T, \xi) | \mathcal{F}_{r_n^1}], \quad \mathbb{P}\text{-a.s.}$$

By dominated convergence (recall that the $\mathcal{Y}^{\mathbb{P}}$ are in $\mathbb{D}_0^p(\mathbb{F}^{\mathbb{P}^+}, \mathbb{P})$, with a norm independent of \mathbb{P} , by Lemma A.1), the above convergence also holds for the $\mathbb{L}_0^{\tilde{p}}(\mathbb{P})$ -norm, for any $1 < \tilde{p} < p$. By the stability result of Lemma A.1 (together with Lemma 2.2) and the monotone convergence theorem, we deduce that

$$\begin{aligned} \mathcal{Y}_t^{\mathbb{P}}(r_n^1, \widehat{\mathcal{Y}}_{r_n^1}^+) &= \mathcal{Y}_t^{\mathbb{P}}(r_n^1, \lim_{m \rightarrow +\infty} \uparrow \mathbb{E}^{\mathbb{P}^m} [\mathcal{Y}_{r_n^1}^{\mathbb{P}^m}(T, \xi) | \mathcal{F}_{r_n^1}]), \quad \mathbb{P}\text{-a.s.} \\ &= \lim_{m \rightarrow +\infty} \mathcal{Y}_t^{\mathbb{P}}(r_n^1, \mathbb{E}^{\mathbb{P}^m} [\mathcal{Y}_{r_n^1}^{\mathbb{P}^m}(T, \xi) | \mathcal{F}_{r_n^1}]), \quad \mathbb{P}\text{-a.s.} \\ &= \lim_{m \rightarrow +\infty} \mathcal{Y}_t^{\mathbb{P}^m}(r_n^1, \mathbb{E}^{\mathbb{P}^m} [\mathcal{Y}_{r_n^1}^{\mathbb{P}^m}(T, \xi) | \mathcal{F}_{r_n^1}]), \quad \mathbb{P}\text{-a.s.} \\ &= \lim_{m \rightarrow +\infty} \mathcal{Y}_t^{\mathbb{P}^m}(r_n^1, \mathcal{Y}_{r_n^1}^{\mathbb{P}^m}(T, \xi)), \quad \mathbb{P}\text{-a.s.} \\ &= \lim_{m \rightarrow +\infty} \mathcal{Y}_t^{\mathbb{P}^m}(T, \xi), \quad \mathbb{P}\text{-a.s.} \\ &\leq \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(t, \mathbb{P}, \mathbb{F}^+)}^{\mathbb{P}} \mathcal{Y}_t^{\mathbb{P}'}(T, \xi), \quad \mathbb{P}\text{-a.s.,} \end{aligned}$$

where we have used in the third equality the fact that \mathbb{P}^m coincides with \mathbb{P} on $\mathcal{F}_{r_n^1}$ and that $\mathcal{Y}_t^{\mathbb{P}}$ is \mathcal{F}_t^+ -measurable, Lemma 2.7 in the fourth equality and the dynamic programming principle for BSDEs in the fifth equality.

Finally, it remains to let n go to $+\infty$ and to use Lemma A.2 (together with Lemma 2.2) to obtain the desired equality, from which we deduce exactly as in the proof of Theorem 4.4 that $\widehat{\mathcal{Y}}^+ \in \mathbb{D}_0^p(\mathbb{F}^{\mathcal{P}_0^+})$. \square

The next result shows that $\widehat{\mathcal{Y}}^+$ is actually a semimartingale under any $\mathbb{P} \in \mathcal{P}_0$, and gives its decomposition.

LEMMA 3.6. *Let Assumptions 2.1 and 3.1 hold. For any $\mathbb{P} \in \mathcal{P}_0$, there is $(Z^{\mathbb{P}}, M^{\mathbb{P}}, K^{\mathbb{P}}) \in \mathbb{H}_0^p(\mathbb{F}^{\mathbb{P}^+}, \mathbb{P}) \times \mathbb{M}_0^p(\mathbb{F}^{\mathbb{P}^+}, \mathbb{P}) \times \mathbb{I}_0^p(\mathbb{F}^{\mathbb{P}^+}, \mathbb{P})$ such that*

$$\begin{aligned} \widehat{\mathcal{Y}}_t^+ &= \xi - \int_t^T \widehat{f}_s^{\mathbb{P}}(\widehat{\mathcal{Y}}_s^+, (\widehat{a}_s^{1/2})^\top Z_s^{\mathbb{P}}) ds - \int_t^T Z_s^{\mathbb{P}} \cdot dX_s^{c, \mathbb{P}} \\ &\quad - \int_t^T dM_s^{\mathbb{P}} + \int_t^T dK_s^{\mathbb{P}}, \quad t \in [0, T], \mathbb{P}\text{-a.s.} \end{aligned}$$

Moreover, there is some $\mathbb{F}^{\mathcal{P}_0}$ -predictable process Z which aggregates the family $(Z^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}$.

PROOF. Fix some $\mathbb{P} \in \mathcal{P}_0$. Consider the following reflected BSDE on the enlarged space. For $0 \leq t \leq T$, $\mathbb{P} \otimes \mathbb{P}_0$ -a.s.,

$$\left\{ \begin{aligned} \bar{y}_t^\mathbb{P} &= \xi - \int_t^T \widehat{f}_s^\mathbb{P}(\bar{y}_s^\mathbb{P}, (\widehat{a}_s^{1/2})^\top \bar{z}_s^\mathbb{P}) ds - \int_t^T \bar{z}_s^\mathbb{P} \cdot \widehat{a}_s^{1/2} dW_s^\mathbb{P} \\ &\quad - \int_t^T d\bar{m}_s^\mathbb{P} + \bar{k}_T^\mathbb{P} - \bar{k}_t^\mathbb{P}, \\ \bar{y}_t^\mathbb{P} &\geq \widehat{\mathcal{Y}}_t^+, \\ \int_0^T (\bar{y}_{t-}^\mathbb{P} - \widehat{\mathcal{Y}}_{t-}^+) d\bar{k}_t^\mathbb{P} &= 0. \end{aligned} \right.$$

By Theorem 3.1 in [13], this reflected BSDE is well-posed and $\bar{y}^\mathbb{P}$ is càdlàg. By abuse of notation, we denote $\widehat{\mathcal{Y}}^+(\bar{\omega}) := \widehat{\mathcal{Y}}^+(\pi(\bar{\omega}))$. We claim that $\bar{y}^\mathbb{P} = \widehat{\mathcal{Y}}^+$, $\mathbb{P} \otimes \mathbb{P}_0$ -a.s. Indeed, we argue by contradiction, and assume without loss of generality that $\bar{y}_0^\mathbb{P} > \widehat{\mathcal{Y}}_0^+$. For each $\varepsilon > 0$, denote $\tau_\varepsilon := \inf\{t : \bar{y}_t^\mathbb{P} \leq \widehat{\mathcal{Y}}_t^+ + \varepsilon\}$. Then τ_ε is an $\overline{\mathbb{F}}_+$ -stopping time and $\bar{y}_{t-}^\mathbb{P} \geq \widehat{\mathcal{Y}}_{t-}^+ + \varepsilon > \widehat{\mathcal{Y}}_{t-}^+$ for all $t \leq \tau_\varepsilon$. Thus $\bar{k}_t^\mathbb{P} = 0$, $\mathbb{P} \otimes \mathbb{P}_0$ -a.s., for $0 \leq t \leq \tau_\varepsilon$, and thus

$$\begin{aligned} \bar{y}_t^\mathbb{P} &= \bar{y}_{\tau_\varepsilon}^\mathbb{P} - \int_t^{\tau_\varepsilon} \widehat{f}_s^\mathbb{P}(\bar{y}_s^\mathbb{P}, (\widehat{a}_s^{1/2})^\top \bar{z}_s^\mathbb{P}) ds \\ &\quad - \int_t^{\tau_\varepsilon} \bar{z}_s^\mathbb{P} \cdot \widehat{a}_s^{1/2} dW_s^\mathbb{P} - \int_t^{\tau_\varepsilon} d\bar{m}_s^\mathbb{P}, \quad \mathbb{P} \otimes \mathbb{P}_0\text{-a.s.} \end{aligned}$$

The same linearization argument that we used in the proof of Lemma A.1 implies that

$$\bar{y}_0^\mathbb{P} \leq \mathcal{Y}_0^{\mathbb{P} \otimes \mathbb{P}_0}(\tau_\varepsilon, \widehat{\mathcal{Y}}_{\tau_\varepsilon}^+) + C \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0}[\bar{y}_{\tau_\varepsilon}^\mathbb{P} - \widehat{\mathcal{Y}}_{\tau_\varepsilon}^+] \leq \mathcal{Y}_0^{\mathbb{P} \otimes \mathbb{P}_0}(\tau_\varepsilon, \widehat{\mathcal{Y}}_{\tau_\varepsilon}^+) + C\varepsilon,$$

for some $C > 0$. However, by Lemma 3.4, we know that $\mathcal{Y}_0^{\mathbb{P} \otimes \mathbb{P}_0}(\tau_\varepsilon, \widehat{\mathcal{Y}}_{\tau_\varepsilon}^+) \leq \widehat{\mathcal{Y}}_0^+$, which contradicts the fact that $\bar{y}_0^\mathbb{P} > \widehat{\mathcal{Y}}_0^+$.

Then, by exactly the same arguments as in Lemma 2.2, we can go from the enlarged space to Ω and obtain for some $(Z^\mathbb{P})_{\mathbb{P} \in \mathcal{P}_0} \subset \mathbb{H}_0^p(\overline{\mathbb{F}}_+, \mathbb{P})$, and $(M^\mathbb{P}, K^\mathbb{P})_{\mathbb{P} \in \mathcal{P}_0} \subset \mathbb{M}_0^p(\overline{\mathbb{F}}^+, \mathbb{P}) \times \mathbb{I}_0^p(\overline{\mathbb{F}}^+, \mathbb{P})$

$$\begin{aligned} \widehat{\mathcal{Y}}_t^+ &= \xi - \int_t^T \widehat{f}_s^\mathbb{P}(\widehat{\mathcal{Y}}_s^+, (\widehat{a}_s^{1/2})^\top Z_s^\mathbb{P}) ds - \int_t^T Z_s^\mathbb{P} \cdot dX_s^{c, \mathbb{P}} \\ &\quad - \int_t^T dM_s^\mathbb{P} + \int_t^T dK_s^\mathbb{P}, \quad t \in [0, T], \mathbb{P}\text{-a.s.} \end{aligned}$$

Then, by Karandikar [51], since $\widehat{\mathcal{Y}}^+$ is a càdlàg semimartingale, we can define a universal process denoted by $\langle \widehat{\mathcal{Y}}^+, X \rangle$ which coincides with the quadratic covariation of $\widehat{\mathcal{Y}}^+$ and X under each probability $\mathbb{P} \in \mathcal{P}_0$. In particular, the process $\langle \widehat{\mathcal{Y}}^+, X \rangle$ is \mathcal{P}_0 -quasi-surely continuous, and hence is $\mathbb{F}^{\mathcal{P}_0+}$ -predictable (or equivalently $\mathbb{F}^{\mathcal{P}_0}$ -predictable). Similar to the proof of Theorem 2.4 of [69], we can then

define a universal $\mathbb{F}^{\mathcal{P}_0}$ -predictable process Z by

$$Z_t := \widehat{a}_t^{\oplus} \frac{d\langle \widehat{\mathcal{Y}}^+, X \rangle_t}{dt},$$

where we recall that \widehat{a}_t^{\oplus} represents the Moore–Penrose pseudo-inverse of \widehat{a}_t . In particular, Z aggregates the family $\{Z^{\mathbb{P}}, \mathbb{P} \in \mathcal{P}_0\}$. \square

We end this section with a remark, which explains that in some cases, the path regularization that we used is actually unnecessary, and we can obtain a semimartingale decomposition for \widehat{Y} directly.

REMARK 3.1. Assume that for any $(t, \omega) \in [0, T] \times \Omega$, all the probability measures in $\mathcal{P}(t, \omega)$ satisfy the Blumenthal 0–1 law. This would be the case for instance if we were working with the set $\overline{\mathcal{P}}_S$ defined and used in [84]. Then, for any $\mathbb{P} \in \mathcal{P}(t, \omega)$, the filtration $\mathbb{F}^{\mathbb{P}}$ is right-continuous and, therefore, satisfies the usual conditions. Assume that the process \widehat{Y} is $\mathbb{F}^{\mathbb{P}}$ -optional, then it is also $\text{l\`a}d\text{l\`a}g$, and by Lemma 3.5, it verifies

$$\sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}}[\text{esssup}_{t \in [0, T]}^{\mathbb{P}} |\widehat{Y}_t|^p] < +\infty.$$

Moreover, by the Blumenthal 0–1 law, (3.1) and (3.2) rewrite

$$\widehat{\mathcal{Y}}_{\sigma(\omega)}(\omega) = \sup_{\mathbb{P} \in \mathcal{P}(\sigma(\omega), \omega)} \mathcal{Y}_{\sigma(\omega)}^{\mathbb{P}}(\tau, \widehat{\mathcal{Y}}_{\tau}) = \sup_{\mathbb{P} \in \mathcal{P}(\sigma(\omega), \omega)} \widetilde{\mathcal{Y}}_{\sigma(\omega)}^{\mathbb{P} \otimes \mathbb{P}_0}(\tau, \widehat{\mathcal{Y}}_{\tau}).$$

Hence, $\widehat{\mathcal{Y}}$ is a $\mathcal{E}^{\widehat{\mathbb{F}}^{\mathbb{P}}}$ -supermartingale in the terminology of [12]. We can then apply Theorem 3.1 of [12] to obtain directly the semimartingale decomposition of Lemma 3.6. The aggregation of the family $(Z^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}$ can still be done, but requires to use Karandikar’s approach [51], combined with the Itô formula for $\text{l\`a}d\text{l\`a}g$ processes of [58], page 538. Then one can also generalize the results on 2BSDEs of the section below. This however requires that in the definition of a 2BSDE (see Definition 4.1), the processes Y and K are only $\text{l\`a}d\text{l\`a}g$, instead of $\text{c\`a}d\text{l\`a}g$. With this change, all our results still go through.

4. Application to 2BSDEs.

4.1. *Definition.* We shall consider the following 2BSDE, which verifies \mathcal{P}_0 -q.s.:

$$(4.1) \quad \begin{aligned} Y_t = \xi - \int_t^T \widehat{f}_s^{\mathbb{P}}(Y_s, (\widehat{a}_s^{1/2})^{\top} Z_s) ds - \left(\int_t^T Z_s \cdot dX_s^{c, \mathbb{P}} \right)^{\mathbb{P}} \\ - \int_t^T dM_s^{\mathbb{P}} + K_T^{\mathbb{P}} - K_t^{\mathbb{P}}, \quad 0 \leq t \leq T. \end{aligned}$$

DEFINITION 4.1. We will say that the quadruple $(Y, Z, (M^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}, (K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}) \in \mathbb{D}_0^P(\mathbb{F}^{\mathcal{P}_0+}) \times \mathbb{H}_0^P(\mathbb{F}^{\mathcal{P}_0}) \times \mathbb{M}_0^P((\mathbb{F}^{\mathcal{P}_0+})_{\mathbb{P} \in \mathcal{P}_0}) \times \mathbb{I}_0^P((\mathbb{F}_+^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0})$ is a solution to the 2BSDE (4.1) if (4.1) holds \mathcal{P}_0 -q.s. and if the family $\{K^{\mathbb{P}}, \mathbb{P} \in \mathcal{P}_0\}$ satisfies the minimality condition

$$(4.2) \quad K_t^{\mathbb{P}} = \operatorname{ess\,inf}_{\mathbb{P}' \in \mathcal{P}_0(t, \mathbb{P}, \mathbb{F}_+)} \mathbb{E}^{\mathbb{P}'} [K_T^{\mathbb{P}'} | \mathcal{F}_t^{\mathbb{P}'+}], \quad 0 \leq t \leq T, \mathbb{P}\text{-a.s.}, \forall \mathbb{P} \in \mathcal{P}_0.$$

REMARK 4.1. If we assume that $b^{\mathbb{P}} = 0$, \mathbb{P} -a.s. for any $\mathbb{P} \in \mathcal{P}_0$, then we have that $X^{c, \mathbb{P}} = X$, \mathbb{P} -a.s. for any $\mathbb{P} \in \mathcal{P}_0$. Then we can use the general result given by Nutz [68]⁷ to obtain the existence of a \mathcal{P}_0 -q.s. càdlàg $\mathbb{F}^{\mathcal{P}_0+}$ -progressively measurable process, which we denote by $\int_0^\cdot Z_s \cdot dX_s$, such that

$$\int_0^\cdot Z_s \cdot dX_s = \left(\int_0^\cdot Z_s \cdot dX_s \right)^{\mathbb{P}}, \quad \mathbb{P}\text{-a.s.}$$

Hence, we can then also find an $\mathbb{F}^{\mathcal{P}_0+}$ -progressively measurable process N which aggregates the process $M^{\mathbb{P}} - K^{\mathbb{P}}$, and which is therefore a $(\mathbb{F}_+^{\mathbb{P}}, \mathbb{P})$ -supermartingale for any $\mathbb{P} \in \mathcal{P}_0$. However, the Doob–Meyer decomposition of N into a sum of a martingale and a nondecreasing process generally depends on \mathbb{P} . If furthermore the set \mathcal{P}_0 only contains elements satisfying the predictable martingale representation property, for instance the set $\overline{\mathcal{P}}_S$ used in [84], then we have that $M^{\mathbb{P}} = 0$, \mathbb{P} -a.s., for any $\mathbb{P} \in \mathcal{P}_0$, so that the above reasoning allows to aggregate the nondecreasing processes $K^{\mathbb{P}}$.

We first state the main result of this part.

THEOREM 4.1. *Let $\xi \in \mathbb{L}_0^{P, \kappa}$. Under Assumptions 2.1 and 3.1, there exists a unique solution $(Y, Z, (M^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}, (K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0})$ to the 2BSDE (4.1).*

4.2. *Uniqueness, stochastic control representation and comparison.* We start by proving a representation of a solution to 2BSDEs, which provides incidentally its uniqueness.

⁷Notice that this result only holds under some particular set-theoretic axioms. For instance, one can assume the usual Zermelo–Fraenkel set theory, plus the axiom of choice (ZFC for short), and either add the continuum hypothesis or Martin’s axiom (which is compatible with the negation of the continuum hypothesis). Actually, the required axioms must imply the existence of the so-called medial limits in the sense of Mokobodzki. As far as we know, the weakest set of axioms known to be sufficient for the existence of medial limits (see [45], 538S and [57]) is ZFC plus the statement that the reals are not a union of fewer than continuum many meager sets. Moreover, ZFC alone is not sufficient, in the sense that by Corollary 3.3 of [57], If ZFC is consistent, then so is ZFC + “there exist no medial limits”.

THEOREM 4.2. *Let Assumptions 2.1 and 3.1 hold, and let $\xi \in \mathbb{L}_0^{p,\kappa}$. Consider a solution $(Y, Z, (M^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}, (K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0})$ to the 2BSDE (4.1). For any $\mathbb{P} \in \mathcal{P}_0$, let $(\mathcal{Y}^{\mathbb{P}}, \mathcal{Z}^{\mathbb{P}}, \mathcal{M}^{\mathbb{P}}) \in \mathbb{D}_0^p(\mathbb{F}_+^{\mathbb{P}}, \mathbb{P}) \times \mathbb{H}_0^p(\mathbb{F}_+^{\mathbb{P}}, \mathbb{P}) \times \mathbb{M}_0^p(\mathbb{F}_+^{\mathbb{P}}, \mathbb{P})$ be the solutions of the corresponding BSDEs (2.7). Then, for any $\mathbb{P} \in \mathcal{P}_0$ and $0 \leq t_1 \leq t_2 \leq T$,*

$$(4.3) \quad Y_{t_1} = \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)}^{\mathbb{P}} \mathcal{Y}_{t_1}^{\mathbb{P}'}(t_2, Y_{t_2}), \quad \mathbb{P}\text{-a.s.}$$

Thus, the 2BSDE (4.1) has at most one solution in $\mathbb{D}_0^p(\mathbb{F}^{\mathcal{P}_0+}) \times \mathbb{H}_0^p(\mathbb{F}^{\mathcal{P}_0}) \times \mathbb{M}_0^p((\mathbb{F}_+^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}) \times \mathbb{I}_0^p((\mathbb{F}_+^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0})$.

PROOF. We start by proving the representation (4.3) in three steps.

(i) Fix some $\mathbb{P} \in \mathcal{P}_0$ and then some $\mathbb{P}' \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+^{\mathbb{P}})$. Since (4.1) holds \mathbb{P}' -a.s., we can see Y as a super-solution of the BSDE on $[t_1, t_2]$, under \mathbb{P}' , with generator $\widehat{f}^{\mathbb{P}'}$ and terminal condition Y_{t_2} . By the comparison principle of Lemma A.3 (together with Lemma 2.2), we deduce immediately that $Y_{t_1} \geq \mathcal{Y}_{t_1}^{\mathbb{P}'}(t_2, Y_{t_2})$, \mathbb{P}' -a.s. Then, since $\mathcal{Y}_{t_1}^{\mathbb{P}'}(t_2, Y_{t_2})$ (or a \mathbb{P}' -version of it) is $\mathcal{F}_{t_1}^+$ -measurable and since Y_{t_1} is $\mathcal{F}_{t_1}^{\mathcal{P}_0+}$ -measurable, we deduce that the inequality also holds \mathbb{P} -a.s., by definition of $\mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)$ and the fact that measures extend uniquely to the completed σ -algebras. We deduce that

$$Y_{t_1} \geq \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)}^{\mathbb{P}} \mathcal{Y}_{t_1}^{\mathbb{P}'}(t_2, Y_{t_2}), \quad \mathbb{P}\text{-a.s.},$$

by arbitrariness of \mathbb{P}' .

(ii) We now show that

$$C_{t_1}^{\mathbb{P}} := \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)}^{\mathbb{P}} \mathbb{E}^{\mathbb{P}'}[(K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'})^p | \mathcal{F}_{t_1}^+] < +\infty, \quad \mathbb{P}\text{-a.s.}$$

First of all, we have by definition

$$\begin{aligned} (K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'})^p &\leq C \left(\sup_{t_1 \leq t \leq t_2} |Y_t|^p + \left(\int_{t_1}^{t_2} |\widehat{f}^{\mathbb{P}',0}| ds \right)^p + \left(\int_{t_1}^{t_2} \|(\widehat{a}_s^{1/2})^\top Z_s\| ds \right)^p \right) \\ &\quad + C \left(\left| \int_{t_1}^{t_2} Z_s \cdot dX_s^{c,\mathbb{P}'} \right|^p + \left| \int_{t_1}^{t_2} dM_s^{\mathbb{P}'} \right|^p \right), \end{aligned}$$

for some constant $C > 0$, so that we obtain by BDG inequalities

$$(4.4) \quad \mathbb{E}^{\mathbb{P}'}[(K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'})^p] \leq C \left(\phi_f^{p,\kappa} + \|Y\|_{\mathbb{D}_0^p}^p + \|Z\|_{\mathbb{H}_0^p}^p + \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}}[[M^{\mathbb{P}}]_T^{\frac{p}{2}}] \right),$$

for some other constant $C > 0$ and hence $C_{t_1}^{\mathbb{P}} < +\infty$, \mathbb{P} -a.s. Next, we claim that the family

$$\{\mathbb{E}^{\mathbb{P}'}[(K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'})^p | \mathcal{F}_{t_1}^+], \mathbb{P}' \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)\},$$

is upward directed. Indeed, let us consider $(\mathbb{P}^1, \mathbb{P}^2) \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+) \times \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}^+)$, and let us define the following subsets of Ω :

$$A_1 := \{\omega \in \Omega : \mathbb{E}^{\mathbb{P}^1}[(K_{t_2}^{\mathbb{P}^1} - K_{t_1}^{\mathbb{P}^1})^p | \mathcal{F}_{t_1}^+](\omega) > \mathbb{E}^{\mathbb{P}^2}[(K_{t_2}^{\mathbb{P}^2} - K_{t_1}^{\mathbb{P}^2})^p | \mathcal{F}_{t_1}^+](\omega)\},$$

$$A_2 := \Omega \setminus A_1.$$

Then $A_1, A_2 \in \mathcal{F}_{t_1}^{\mathbb{P}^+}$, and we can define the following probability measure on (Ω, \mathcal{F}_T) :

$$\mathbb{P}^{1,2}(B) := \mathbb{P}^1(A_1 \cap B) + \mathbb{P}^2(A_2 \cap B), \quad \text{for any } B \in \mathcal{F}_T.$$

By Assumption 2.1(v), we know that $\mathbb{P}^{1,2} \in \mathcal{P}_0$, and by definition, we further have $\mathbb{P}^{1,2} \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)$ as well as, \mathbb{P} -a.s.,

$$\mathbb{E}^{\mathbb{P}^{1,2}}[(K_{t_2}^{\mathbb{P}^{1,2}} - K_{t_1}^{\mathbb{P}^{1,2}})^p | \mathcal{F}_{t_1}^+] = \mathbb{E}^{\mathbb{P}^1}[(K_{t_2}^{\mathbb{P}^1} - K_{t_1}^{\mathbb{P}^1})^p | \mathcal{F}_{t_1}^+] \vee \mathbb{E}^{\mathbb{P}^2}[(K_{t_2}^{\mathbb{P}^2} - K_{t_1}^{\mathbb{P}^2})^p | \mathcal{F}_{t_1}^+],$$

which proves the claim.

Therefore, by classical results for the essential supremum (see, e.g., Neveu [67]), there exists a sequence $(\mathbb{P}^n)_{n \geq 0} \subset \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)$ such that

$$\text{ess sup}_{\mathbb{P}' \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)} \mathbb{E}^{\mathbb{P}'}[(K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'})^p | \mathcal{F}_{t_1}^+] = \lim_{n \rightarrow \infty} \uparrow \mathbb{E}^{\mathbb{P}^n}[(K_{t_2}^{\mathbb{P}^n} - K_{t_1}^{\mathbb{P}^n})^p | \mathcal{F}_{t_1}^+].$$

Then using (4.4) and the monotone convergence theorem under \mathbb{P} , we deduce that

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}[C_{t_1}^{\mathbb{P}}] &\leq \lim_{n \rightarrow \infty} \uparrow \mathbb{E}^{\mathbb{P}}[\mathbb{E}^{\mathbb{P}^n}[(K_{t_2}^{\mathbb{P}^n} - K_{t_1}^{\mathbb{P}^n})^p | \mathcal{F}_{t_1}^+]] \\ &\leq C\left(\phi_f^{p, \kappa} + \|Y\|_{\mathbb{D}_0^p}^p + \|Z\|_{\mathbb{H}_0^p}^p + \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}}\left[[M^{\mathbb{P}}]_T^{\frac{p}{2}}\right]\right) < +\infty, \end{aligned}$$

which provides the desired result.

(iii) We now prove the reverse inequality. Since we will use a linearization argument, we work on the enlarged space, remembering that this is without loss of generality by Lemma 2.2. Fix $\mathbb{P} \in \mathcal{P}_0$. For every $\mathbb{P}' \in \mathcal{P}' \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)$, we extend the definition of $(Y, Z, (M^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}, (K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0})$ on $\overline{\Omega}$ as in (2.1), and denote

$$\delta Y := Y - \tilde{Y}^{\mathbb{P}' \otimes \mathbb{P}_0}, \quad \delta Z := Z - \tilde{Z}^{\mathbb{P}' \otimes \mathbb{P}_0} \quad \text{and} \quad \delta M^{\mathbb{P}'} := M^{\mathbb{P}'} - \tilde{M}^{\mathbb{P}' \otimes \mathbb{P}_0}.$$

By Assumption 2.1(i), there exist two bounded processes $\lambda^{\mathbb{P}'}$ and $\eta^{\mathbb{P}'}$ such that for all $t_1 \leq t \leq t_2$, $\mathbb{P}' \otimes \mathbb{P}_0$ -a.s.,

$$\begin{aligned} \delta Y_t &= \int_t^{t_2} (\lambda_s^{\mathbb{P}'} \delta Y_s + \eta_s^{\mathbb{P}'} \cdot (\hat{a}_s^{1/2})^\top \delta Z_s) ds - \int_t^{t_2} \delta Z_s \cdot \hat{a}_s^{1/2} dW_s^{\mathbb{P}'} \\ &\quad - \int_t^{t_2} d(\delta M_s^{\mathbb{P}'} - K_s^{\mathbb{P}'}). \end{aligned}$$

Define for $t_1 \leq t \leq t_2$ the following continuous process:

$$(4.5) \quad \Delta_t^{\mathbb{P}'} := \exp\left(\int_{t_1}^t \left(\lambda_s^{\mathbb{P}'} - \frac{1}{2} \|\eta_s^{\mathbb{P}'}\|^2\right) ds - \int_{t_1}^t \eta_s^{\mathbb{P}'} \cdot dW_s^{\mathbb{P}'}\right), \quad \mathbb{P}' \otimes \mathbb{P}_0\text{-a.s.}$$

Note that since $\lambda^{\mathbb{P}'}$ and $\eta^{\mathbb{P}'}$ are bounded, we have for all $p \geq 1$, for some constant $C_p > 0$, independent of \mathbb{P}' .

$$(4.6) \quad \mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} \left[\sup_{t_1 \leq t \leq t_2} (\Delta_t^{\mathbb{P}'})^p + \sup_{t_1 \leq t \leq t_2} (\Delta_t^{\mathbb{P}'})^{-p} \mid \overline{\mathcal{F}}_{t_1}^+ \right] \leq C_p, \quad \mathbb{P}' \otimes \mathbb{P}_0\text{-a.s.}$$

Then, by Itô’s formula, we obtain

$$(4.7) \quad \delta Y_{t_1} = \mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} \left[\int_{t_1}^{t_2} \Delta_t^{\mathbb{P}'} dK_t^{\mathbb{P}'} \mid \overline{\mathcal{F}}_{t_1}^+ \right],$$

because the martingale terms vanish when taking conditional expectation. We therefore deduce

$$\begin{aligned} \delta Y_{t_1} &\leq \left(\mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} \left[\sup_{t_1 \leq t \leq t_2} |\Delta_t^{\mathbb{P}'}|^{\frac{p+1}{p-1}} \mid \overline{\mathcal{F}}_{t_1}^+ \right] \right)^{\frac{p-1}{p+1}} \left(\mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} \left[(K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'})^{\frac{p+1}{2}} \mid \overline{\mathcal{F}}_{t_1}^+ \right] \right)^{\frac{2}{p+1}} \\ &\leq C(C_{t_1}^{\mathbb{P}'})^{\frac{1}{p+1}} \left(\mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} [K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'} \mid \overline{\mathcal{F}}_{t_1}^+] \right)^{\frac{1}{p+1}}. \end{aligned}$$

Remember Y and $K^{\mathbb{P}'}$ are extended on $\overline{\Omega}$ as in (2.1), then it only depends on X and not on B . Going back now to the canonical space Ω , it follows by Lemma 2.2 that

$$\delta Y'_{t_1} := Y_{t_1} - \tilde{\mathcal{Y}}_{t_1}^{\mathbb{P}'} \leq C(C_{t_1}^{\mathbb{P}'})^{\frac{1}{p+1}} \left(\mathbb{E}^{\mathbb{P}'} [K_{t_2}^{\mathbb{P}'} - K_{t_1}^{\mathbb{P}'} \mid \mathcal{F}_{t_1}^+] \right)^{\frac{1}{p+1}}.$$

By arbitrariness of \mathbb{P}' we deduce, thanks to (4.2), that

$$Y_{t_1} - \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(t_1, \mathbb{P}, \mathbb{F}_+)}^{\mathbb{P}} \mathcal{Y}_{t_1}^{\mathbb{P}'}(t_2, Y_{t_2}) \leq 0, \quad \mathbb{P}\text{-a.s.}$$

Finally, the uniqueness of Y is immediate by the representation (4.3). Then, since

$$\langle Y, X \rangle_t = \int_0^t \widehat{a}_s Z_s ds, \quad \mathbb{P}\text{-a.s.},$$

Z is also uniquely defined, $\widehat{a}_t dt \otimes \mathcal{P}_0$ -q.s. We therefore deduce that the processes $M^{\mathbb{P}} - K^{\mathbb{P}}$ are also uniquely defined for any $\mathbb{P} \in \mathcal{P}_0$. But, since they are $(\mathbb{F}_+^{\mathbb{P}}, \mathbb{P})$ -supermartingales, such that in addition $(K_t^{\mathbb{P}}, M_t^{\mathbb{P}}) \in \mathbb{L}_0^{\mathbb{P}}(\mathbb{F}_+^{\mathbb{P}}, \mathbb{P}) \times \mathbb{L}_0^{\mathbb{P}}(\mathbb{F}_+^{\mathbb{P}}, \mathbb{P})$ for any $t \in [0, T]$, and since $K^{\mathbb{P}}$ is $\mathbb{F}_+^{\mathbb{P}}$ -predictable, the uniqueness of $M^{\mathbb{P}}$ and $K^{\mathbb{P}}$ is a simple consequence of the uniqueness in the Doob–Meyer decomposition of these supermartingales. \square

With the previous theorem in hand, the following comparison result is an immediate consequence of the corresponding one for BSDEs (see, for instance, Lemma A.3 in the Appendix)

THEOREM 4.3. *For $i = 1, 2$, let f^i and ξ^i be respectively a generator map and a terminal condition satisfying the required properties in Assumptions 2.1 and 3.1. Let also Y^i be the first component of the solution to the 2BSDE with generator f^i and terminal condition ξ^i . Suppose in addition that for any $\mathbb{P} \in \mathcal{P}_0$ we have:*

- (i) $\xi^1 \leq \xi^2, \mathbb{P}$ -a.s.
- (ii) $\widehat{f}_s^{1,\mathbb{P}}(y_s^1, (\widehat{a}_s^{1/2})^\top z_s^1) \geq \widehat{f}_s^{2,\mathbb{P}}(y_s^2, (\widehat{a}_s^{1/2})^\top z_s^2), ds \times d\mathbb{P}$ -a.e., on $[0, T] \times \Omega$, where for $i = 1, 2, (y^i, z^i)$ are the first two components of the solution of the BSDE under \mathbb{P} with generator $\widehat{f}^{i,\mathbb{P}}$ and terminal condition ξ^i .

Then we have $Y_t^1 \leq Y_t^2, t \in [0, T], \mathcal{P}_0$ -q.s.

4.3. *A priori estimates and stability.* In this section, we give *a priori* estimates for 2BSDEs, which, as in the case of the classical BSDEs, play a very important role in the study of associated numerical schemes for instance. The proofs are actually based on the general results given very recently in [13].

THEOREM 4.4. *Let Assumptions 2.1 and 3.1 hold, and let $\xi \in \mathbb{L}_0^{p,\kappa}$. Consider a solution $(Y, Z, (M^\mathbb{P})_{\mathbb{P} \in \mathcal{P}_0}, (K^\mathbb{P})_{\mathbb{P} \in \mathcal{P}_0})$ to the 2BSDE (4.1). Then there exists a constant C_κ depending only on p, κ, T and the Lipschitz constant of f such that*

$$\|Y\|_{\mathbb{D}_0^p}^p + \|Z\|_{\mathbb{H}_0^p}^p + \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^\mathbb{P}[(K_T^\mathbb{P})^p] + \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^\mathbb{P}[[M^\mathbb{P}]_T^{\frac{p}{2}}] \leq C_\kappa (\|\xi\|_{\mathbb{L}_0^{p,\kappa}}^p + \phi_f^{p,\kappa}),$$

PROOF. First, by Lemma 3.5, we have for any $\mathbb{P} \in \mathcal{P}_0$:

$$Y_t = \operatorname{ess\,sup}_{\mathbb{P}' \in \mathcal{P}_0(t, \mathbb{P}, \mathbb{F}_+)}^\mathbb{P} \mathcal{Y}_t^{\mathbb{P}'}(T, \xi), \quad \mathbb{P}\text{-a.s.}$$

Furthermore, by Lemma A.1 (together with Lemma 2.2), we know that there exists a constant C (which may change from line to line) depending only on κ, T and the Lipschitz constant of \widehat{f} , such that for all \mathbb{P} ,

$$(4.8) \quad |\mathcal{Y}_t^\mathbb{P}(T, \xi)| \leq C \left(\mathbb{E}^\mathbb{P} \left[|\xi|^\kappa + \int_t^T |\widehat{f}_s^{\mathbb{P},0}|^\kappa ds \middle| \mathcal{F}_t^+ \right] \right)^{\frac{1}{\kappa}}, \quad \mathbb{P}\text{-a.s.}$$

Hence, we deduce immediately

$$\|Y\|_{\mathbb{D}_0^p}^p \leq C (\|\xi\|_{\mathbb{L}_0^{p,\kappa}}^p + \phi_f^{p,\kappa}).$$

Now, by extending the definition of $(Y, Z, (M^\mathbb{P})_{\mathbb{P} \in \mathcal{P}_0}, (K^\mathbb{P})_{\mathbb{P} \in \mathcal{P}_0})$ on the enlarged space $\overline{\Omega}$ (see (2.1)), one has, for every $\mathbb{P} \in \mathcal{P}_0$,

$$\begin{aligned} Y_t &= \xi - \int_t^T \widehat{f}_s^\mathbb{P}(Y_s, (\widehat{a}_s^{1/2})^\top Z_s) ds - \int_t^T Z_s \cdot \widehat{a}_s^{1/2} dW_s^\mathbb{P} \\ &\quad - \int_t^T dM_s^\mathbb{P} + \int_t^T dK_s^\mathbb{P}, \quad \mathbb{P} \otimes \mathbb{P}_0\text{-a.s.} \end{aligned}$$

Then for every $\mathbb{P} \in \mathcal{P}_0, (Y, Z, M^\mathbb{P}, K^\mathbb{P})$ can be interpreted as a super-solution of a BSDE in the enlarged space $\overline{\Omega}$. We can therefore use Theorem 2.1 of [13] (notice that the constants appearing there do not depend on the underlying probability

measure) to obtain the required estimates. Noticing once again that the norms of $Z, K^{\mathbb{P}}$ and $M^{\mathbb{P}}$ are the same on the enlarged space $\overline{\Omega}$ or on Ω , it follows then

$$\begin{aligned} & \|Z\|_{\mathbb{H}_0^p}^p + \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}}[(K_T^{\mathbb{P}})^p] + \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}}[[M^{\mathbb{P}}]_T^{\frac{p}{2}}] \\ & \leq C \left(\|\xi\|_{\mathbb{L}_0^{p,\kappa}}^p + \phi_f^{p,\kappa} + \|\xi\|_{\mathbb{L}_0^p}^p + \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}} \left[\int_0^T |\widehat{f}_s^{\mathbb{P},0}|^p ds \right] \right) \\ & \leq C (\|\xi\|_{\mathbb{L}_0^{p,\kappa}}^p + \phi_f^{p,\kappa}), \end{aligned}$$

for some constant $C > 0$, where we used the fact that by definition

$$\|\xi\|_{\mathbb{L}_0^p}^p \leq \|\xi\|_{\mathbb{L}_0^{p,\kappa}}^p \quad \text{and} \quad \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}} \left[\int_0^T |\widehat{f}_s^{\mathbb{P},0}|^p ds \right] \leq \phi_f^{p,\kappa}. \quad \square$$

Next, we also have the following estimates for the difference of two solutions of 2BSDEs, which plays a fundamental role for stability properties.

THEOREM 4.5. *Let Assumptions 2.1 and 3.1 hold, and let us be given two generators f^1 and f^2 such that Assumption 3.1 holds. Assume that $\xi^i \in \mathbb{L}_0^{p,\kappa}$ and $(Y^i, Z^i, (M^{i,\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}, (K^{i,\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0})$ is a solution to the 2BSDE with generator f^i and terminal condition ξ^i , for $i = 1, 2$. Define*

$$\begin{aligned} \phi_{f^1, f^2}^{p,\kappa} & := \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}} \left[\text{ess sup}_{0 \leq t \leq T}^{\mathbb{P}} \mathbb{E}^{\mathbb{P}} \left[\left(\int_0^T |\widehat{f}_s^{1,\mathbb{P}} - \widehat{f}_s^{2,\mathbb{P}}|^{\kappa} \right. \right. \right. \\ & \quad \left. \left. \left. \times (y_s^{1,\mathbb{P}}, (\widehat{a}_s^{1/2})^\top z_s^{1,\mathbb{P}}) ds \right)^{\frac{p}{\kappa}} \middle| \mathcal{F}_t^+ \right] \right] \\ \psi_{f^1, f^2}^p & := \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}} \left[\int_0^T |\widehat{f}_s^{1,\mathbb{P}} - \widehat{f}_s^{2,\mathbb{P}}|^p (Y_s^1, (\widehat{a}_s^{1/2})^\top Z_s^1) ds \right]. \end{aligned}$$

Then there exists a constant C_κ depending only on κ, T and the Lipschitz constant of f^1 and f^2 such that

$$\begin{aligned} & \|Y^1 - Y^2\|_{\mathbb{D}_0^p}^p \leq C_\kappa (\|\xi^1 - \xi^2\|_{\mathbb{L}_0^{p,\kappa}}^p + \phi_{f^1, f^2}^{p,\kappa}), \\ & \|Z^1 - Z^2\|_{\mathbb{H}_0^p}^p + \sup_{\mathbb{P} \in \mathcal{P}_0} \mathbb{E}^{\mathbb{P}} [[N^{1,\mathbb{P}} - N^{2,\mathbb{P}}]_T^{\frac{p}{2}}] \\ & \leq C_\kappa (\|\xi^1 - \xi^2\|_{\mathbb{L}_0^{p,\kappa}}^p + \phi_{f^1, f^2}^{p,\kappa} + \|\xi^1 - \xi^2\|_{\mathbb{L}_0^{p,\kappa}}^{\frac{p}{2} \wedge (p-1)} \\ & \quad + \psi_{f^1, f^2}^p + (\phi_{f^1, f^2}^{p,\kappa})^{\frac{p}{2} \wedge (p-1)}), \end{aligned}$$

where we have once more defined $N^{i,\mathbb{P}} := M^{i,\mathbb{P}} - K^{i,\mathbb{P}}$ for any $\mathbb{P} \in \mathcal{P}_0, i = 1, 2$.

PROOF. First of all, by Lemma A.1 (together with Lemma 2.2), we know that there exists a constant C depending only on κ, T and the Lipschitz constant of \widehat{f} , such that for all $\mathbb{P} \in \mathcal{P}_0$, we have \mathbb{P} -a.s.

$$(4.9) \quad |y_t^{1,\mathbb{P}} - y_t^{2,\mathbb{P}}| \leq C \left(\mathbb{E}^{\mathbb{P}} \left[|\xi^1 - \xi^2|^\kappa + \int_t^T |\widehat{f}_s^{1,\mathbb{P}} - \widehat{f}_s^{2,\mathbb{P}}|^\kappa (y_s^{1,\mathbb{P}}, (\widehat{a}_s^{1/2})^\top z_s^{1,\mathbb{P}}) ds \middle| \mathcal{F}_t^+ \right] \right)^{\frac{1}{\kappa}}.$$

This immediately provides the estimate for $Y^1 - Y^2$ by the representation formula (4.3) and the definition of the norms and of $\phi_{f^1, f^2}^{p, \kappa}$. Next, we argue exactly as in the proof of Theorem 4.4 by working on the enlarged space $\overline{\Omega}$ and using now Theorem 2.2 of [13] to obtain the required estimates. \square

4.4. *Existence through dynamic programming.* In this section, we will show that $\widehat{\mathcal{Y}}^+$ defined in Section 2 is indeed a solution to the 2BSDE (4.1), thus completing the proof of Theorem 4.1.

Recall that $\widehat{\mathcal{Y}}^+$ is defined by (3.5), and one has processes $(Z, (M^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}, (K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}) \in \mathbb{H}_0^p(\mathbb{F}^{\mathcal{P}_0}) \times \mathbb{M}_0^p((\mathbb{F}^{\mathcal{P}_0})^+) \times \mathbb{I}_0^p((\mathbb{F}^{\mathcal{P}_0})^+)$ given by Lemma 3.6, so that the only thing left for us is to show that the family $(K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}$ satisfies the minimality condition (4.2).

Next we again extend the definition of $(Y, Z, (M^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}, (K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0})$ and $\widehat{\mathcal{Y}}^+$, $\mathcal{Y}^{\mathbb{P}'}(T, \xi)$ on $\overline{\Omega}$ as in (2.1) (recall also Lemma 2.2). Then by (4.7), denoting $\delta \widehat{\mathcal{Y}}^+ := \widehat{\mathcal{Y}}^+ - \mathcal{Y}^{\mathbb{P}'}(T, \xi)$, we have for any $t \in [0, T]$, for any $\mathbb{P} \in \mathcal{P}_0$ and any $\mathbb{P}' \in \mathcal{P}_0(t, \mathbb{P}, \mathbb{F}^+)$

$$\begin{aligned} \delta \widehat{\mathcal{Y}}_t^+ &= \mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} \left[\int_t^T \Delta_s^{\mathbb{P}'} dK_s^{\mathbb{P}'} \middle| \overline{\mathcal{F}}_t^+ \right] \\ &\geq \mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} \left[\inf_{t \leq s \leq T} \Delta_s^{\mathbb{P}'} (K_T^{\mathbb{P}'} - K_t^{\mathbb{P}'}) \middle| \overline{\mathcal{F}}_t^+ \right], \quad \mathbb{P}\text{-a.s.}, \end{aligned}$$

where $\Delta^{\mathbb{P}'}$ is defined in (4.5). We therefore have

$$\begin{aligned} &\mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} [K_T^{\mathbb{P}'} - K_t^{\mathbb{P}'} \middle| \overline{\mathcal{F}}_t^+] \\ &\leq \left(\mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} \left[\inf_{t \leq s \leq T} \Delta_s^{\mathbb{P}'} (K_T^{\mathbb{P}'} - K_t^{\mathbb{P}'}) \middle| \overline{\mathcal{F}}_t^+ \right] \right)^{\frac{1}{2}} \\ &\quad \times \left(\mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} [(K_T^{\mathbb{P}'} - K_t^{\mathbb{P}'})^p \middle| \overline{\mathcal{F}}_t^+] \right)^{\frac{1}{2p}} \left(\mathbb{E}^{\mathbb{P}' \otimes \mathbb{P}_0} \left[\left(\inf_{t \leq s \leq T} \Delta_s^{\mathbb{P}'} \right)^{-q} \middle| \overline{\mathcal{F}}_t^+ \right] \right)^{\frac{1}{2q}} \\ &\leq C (C_t^{\mathbb{P}'})^{\frac{1}{2p}} (\delta \widehat{\mathcal{Y}}_t^+)^{\frac{1}{2}}, \end{aligned}$$

with $q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$.

Notice that by definition $K^{\mathbb{P}'}$ (defined on $\overline{\Omega}$) only depends on X and not on B , so that we can go back to Ω and obtain

$$\mathbb{E}^{\mathbb{P}'} [K_T^{\mathbb{P}'} - K_t^{\mathbb{P}'} | \mathcal{F}_t^+] \leq C(C_t^{\mathbb{P}'})^{\frac{1}{2p}} (\delta \widehat{\mathcal{Y}}_t^+)^{\frac{1}{2}}.$$

Then the result follows immediately thanks to Lemma 3.5.

REMARK 4.2. For other classes of 2BSDEs with possibly a non-Lipschitz generator, such as 2BSDEs under a monotonicity condition [77], quadratic 2BSDEs [79] or second-order reflected BSDEs [63], if a Doob–Meyer decomposition for the corresponding nonlinear supermartingales is available under any probability measure in the set \mathcal{P}_0 , then together with Proposition 2.1, we can generalize the well-posedness result in Theorem 4.1 to these classes of 2BSDEs when there is no regularity condition on the terminal condition and the generator. In particular, all probability measures in the nondominated set considered in the articles above do satisfy this property, which means that our result extends directly to their context.

5. Nonlinear optional decomposition and super-hedging duality. In this section, we show that under an additional assumption on the sets \mathcal{P}_0 , basically stating that it is rich enough, we can give a different definition of second-order BSDEs, which is akin to a nonlinear optional decomposition theorem, as initiated by [39, 43, 56] in a dominated model framework, and more recently by [69] for nondominated models.

5.1. *Saturated 2BSDEs.* We introduce the following definition.

DEFINITION 5.1. The set \mathcal{P}_0 is said to be saturated if, when $\mathbb{P} \in \mathcal{P}_0$, we have $\mathbb{Q} \in \mathcal{P}_0$ for every probability measure \mathbb{Q} on (Ω, \mathcal{F}) which is equivalent to \mathbb{P} and under which X is local martingale.

We give now an alternative definition for 2BSDEs of the form

$$(5.1) \quad \begin{aligned} Y_t = & \xi - \int_t^T \widehat{f}_s^{\mathbb{P}}(Y_s, (\widehat{a}_s^{1/2})^\top Z_s) ds - \left(\int_t^T Z_s \cdot dX_s^{c, \mathbb{P}} \right)^{\mathbb{P}} \\ & + K_T^{\mathbb{P}} - K_t^{\mathbb{P}}, \quad 0 \leq t \leq T. \end{aligned}$$

DEFINITION 5.2. $(Y, Z, (K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0}) \in \mathbb{D}^p(\mathbb{F}^{U+, \mathcal{P}_0}) \times \mathbb{H}^p(\mathbb{F}^{U, \mathcal{P}_0}) \times \mathbb{I}^{o, p}((\mathbb{F}_+^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0})$ is a saturated solution to 2BSDE (5.1) if (5.1) holds \mathcal{P}_0 -q.s. and if the family $\{K^{\mathbb{P}}, \mathbb{P} \in \mathcal{P}_0\}$ satisfies the minimality condition (4.2).

REMARK 5.1. In the above definition, two changes have occurred. First, the orthogonal martingales $M^{\mathbb{P}}$ have disappeared, and the nondecreasing processes $K^{\mathbb{P}}$ are assumed to be $\mathbb{F}_+^{\mathbb{P}}$ -optional instead of predictable.

We then have the following result.

THEOREM 5.1. *Let Assumptions 2.1, 3.1 hold and assume in addition that the set \mathcal{P}_0 is saturated. Then there is a unique saturated solution of the 2BSDE (5.1).*

PROOF. By Theorem 4.1, we know that the following 2BSDE is well-posed:

$$Y_t = \xi - \int_t^T \widehat{f}_s^{\mathbb{P}}(Y_s, (\widehat{a}_s^{1/2})^\top Z_s) ds - \left(\int_t^T Z_s \cdot dX_s^{c,\mathbb{P}} \right)^{\mathbb{P}} - \int_t^T dM_s^{\mathbb{P}} + K_T^{\mathbb{P}} - K_t^{\mathbb{P}}, \quad 0 \leq t \leq T, \mathcal{P}_0\text{-q.s.}$$

In particular, this means that the process

$$Y_t - \int_0^t \widehat{f}_s^{\mathbb{P}}(Y_s, (\widehat{a}_s^{1/2})^\top Z_s) ds,$$

is a $(\mathbb{F}_+^{\mathbb{P}}, \mathbb{P})$ -supermartingale in $\mathbb{D}_0^p(\mathbb{F}_+^{\mathbb{P}}, \mathbb{P})$ for every $\mathbb{P} \in \mathcal{P}_0$. Since \mathcal{P}_0 is saturated, it follows by Theorem 1 of [43] (see also Theorem 3.1 of [44]), that there exists a \mathbb{F} -predictable process $\widetilde{Z}^{\mathbb{P}}$ such that

$$Y_t - \int_0^t \widehat{f}_s^{\mathbb{P}}(Y_s, (\widehat{a}_s^{1/2})^\top Z_s) ds - \int_0^t \widetilde{Z}_s^{\mathbb{P}} \cdot dX_s^{c,\mathbb{P}} \text{ is nonincreasing,}$$

\mathbb{P} -a.s., for every $\mathbb{P} \in \mathcal{P}_0$.

Hence, we can write

$$Y_t = \xi - \int_t^T \widehat{f}_s^{\mathbb{P}}(Y_s, (\widehat{a}_s^{1/2})^\top Z_s) ds - \left(\int_t^T \widetilde{Z}_s^{\mathbb{P}} \cdot dX_s^{c,\mathbb{P}} \right)^{\mathbb{P}} + \widetilde{K}_T^{\mathbb{P}} - \widetilde{K}_t^{\mathbb{P}}, \quad 0 \leq t \leq T, \mathcal{P}_0\text{-q.s.},$$

where for any $\mathbb{P} \in \mathcal{P}_0$, $\widetilde{K}^{\mathbb{P}}$ is càdlàg, nondecreasing \mathbb{P} -a.s. and $\mathbb{F}_+^{\mathbb{P}}$ -optional. Moreover, by identification of the martingale parts, we deduce that we necessarily have $\widetilde{Z}^{\mathbb{P}} = Z, \widehat{a}_t dt \times \mathcal{P}_0\text{-q.s.}$ Finally, following the same arguments as in the proof of Theorem 4.4, we deduce that $(\widetilde{K}^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0} \in \mathbb{I}^{o,p}((\mathbb{F}_+^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0})$, which completes the proof. \square

5.2. *A super-hedging duality in uncertain, incomplete and nonlinear markets.*

The result of the previous section finds an immediate application to the so-called problem of robust super-hedging. Before discussing the related results in the literature, let us explain exactly what the problem is.

We consider a standard financial market (possibly incomplete) consisting of a nonrisky asset and n risky assets whose dynamics are uncertain. Concretely, let U be some (nonempty) Polish space, \mathcal{U} denote the collection of all U -valued \mathbb{F} -progressively measurable processes, $(\mu, \sigma) : [0, T] \times \Omega \times U \longrightarrow \mathbb{R}^d \times \mathbb{S}^d$ be

the drift and volatility coefficient function which are assumed to be bounded (for simplicity) and such that for some constant $L > 0$,

$$\|(\mu, \sigma)(t, \omega, u) - (\mu, \sigma)(t', \omega', u)\| \leq L(\sqrt{|t - t'|} + \|\omega_{t \wedge \cdot} - \omega'_{t \wedge \cdot}\|).$$

Recall that the canonical process X on the canonical space Ω is a standard Brownian motion under the Wiener measure \mathbb{P}_0 . Then the dynamics of risky assets are given by $S^{t, \omega, \nu}$, which is the unique solution to the SDE

$$S_s^{t, \omega, \nu} = \omega_t + \int_t^s \mu(r, S^{t, \omega, \nu}, \nu_r) dr + \int_t^s \sigma(r, S^{t, \omega, \nu}, \nu_r) dX_r, \quad s \in [t, T], \mathbb{P}_0\text{-a.s.},$$

with initial condition $S_s^{t, \omega, \nu} = \omega_s$ for $s \in [0, t]$ and $\nu \in \mathcal{U}$.

Then we define for every $(t, \omega) \in [0, T] \times \Omega$:

$$\mathcal{P}^{\mathcal{U}}(t, \omega) := \{\mathbb{P}_0 \circ (S^{t, \omega, \nu})^{-1}, \nu \in \mathcal{U}\}.$$

It is known (see Theorem 3.5 and Lemma 3.6 in [41] or Theorem 2.4 and Proposition 2.2 in [65] in a simpler context) that these sets do satisfy Assumption 2.1. We assume in addition that $\mathcal{P}_0^{\mathcal{U}}$ is saturated.

A portfolio strategy is then defined as a \mathbb{R}^n -valued and $\mathbb{F}^{\mathcal{P}_0^{\mathcal{U}}}$ -predictable process $(Z_t)_{t \in [0, T]}$, such that Z_t^i describes the number of units of asset i in the portfolio of the investor at time t . It is well known that under some constrained cases, the wealth $Y^{y_0, Z}$ associated to the strategy Z and initial capital $y_0 \in \mathbb{R}$ can be written, for every $\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}$, as

$$Y_t^{y_0, Z} := y_0 + \int_0^t \widehat{f}_s^{\mathbb{P}}(Y_s^{y_0, Z}, (\widehat{a}_s^{1/2})^\top Z_s) ds + \int_0^t Z_s \cdot dX_s^{c, \mathbb{P}}, \quad t \in [0, T], \mathbb{P}\text{-a.s.}$$

For instance, the classical case corresponds to

$$(5.2) \quad \widehat{f}_s^{\mathbb{P}}(y, z) = r_s y + z \cdot \theta_s^{\mathbb{P}},$$

where r_s is the risk-free rate of the market and $\theta^{\mathbb{P}}$ is the risk premium vector under \mathbb{P} , defined by $\theta_s^{\mathbb{P}} := (\widehat{a}_s^{1/2})^\oplus (b_s^{\mathbb{P}} - r_s \mathbf{1}_n)$, where $(\widehat{a}_s^{1/2})^\oplus$ denotes the Moore–Penrose pseudo-inverse of $\widehat{a}_s^{1/2}$.

The simplest example of a nonlinear $\widehat{f}^{\mathbb{P}}$ corresponds to the case where there are different lending and borrowing rates $\underline{r}_t \leq \bar{r}_t$, in which (see Example 1.1 in [37])

$$\widehat{f}_s^{\mathbb{P}}(y, z) = \underline{r}_s y + z \cdot \theta_s^{\mathbb{P}} - (\bar{r}_s - \underline{r}_s)(y - z \cdot \mathbf{1}_n)^-.$$

We will always assume that $\widehat{f}^{\mathbb{P}}$ satisfies our standing hypotheses in Assumptions 2.1 and 3.1.

Let us now be given some Borel random variable $\xi \in \mathbb{L}^p(\mathbb{F}^{U, \mathcal{P}_0^{\mathcal{U}}})$. The problem of super-hedging ξ corresponds to finding its super-replication price, defined as

$$P_{\text{sup}}(\xi) := \inf\{y_0 \in \mathbb{R} : \exists Z \in \mathcal{H}, Y_T^{y_0, Z} \geq \xi, \mathcal{P}_0^{\mathcal{U}}\text{-q.s.}\},$$

where the set of admissible trading strategies \mathcal{H} is defined as the set of $\mathbb{F}^{U, \mathcal{P}_0^{\mathcal{U}}}$ -predictable processes Z such that in addition, $(Y_t^{y_0, Z})_{t \in [0, T]}$ is an $\mathcal{E}^{\hat{f}^{\mathbb{P}}}$ -supermartingale under \mathbb{P} for any $\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}$, that is, for any $0 \leq s \leq t \leq T$

$$Y_s^{y_0, Z} \geq \mathcal{Y}_s^{\mathbb{P}}(t, Y_t^{y_0, Z}), \quad \mathbb{P}\text{-a.s.}$$

In the case where $\hat{f}^{\mathbb{P}}$ corresponds to our first example (5.2) with $r = 0$, and where the set of measures considered satisfy the predictable martingale representation property (i.e., the financial market is complete under any of the measures considered) this super-hedging price has been thoroughly studied in the recent literature; see among others [2, 7, 26, 27, 65, 70, 72, 76, 78, 83, 85, 86]. The extension to possibly incomplete markets has been carried out notably by [10] in discrete-time and more recently by [69] in continuous time for models possibly incorporating jumps. Our result below extends all the results for continuous processes to markets with nonlinear portfolio dynamics. Of course, the same proof would go through for the more general jump case, provided that a 2BSDE theory, extending that of [52, 53], is obtained in such a setting.

THEOREM 5.2. *Let (Y, Z) be the first two components of the saturated solution of the 2BSDE with generator $\hat{f}^{\mathbb{P}}$ and terminal condition ξ . Then*

$$P_{\text{sup}}(\xi) = \sup_{\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}} \mathbb{E}^{\mathbb{P}}[Y_0],$$

and $Z \in \mathcal{H}$ is a super-hedging strategy for ξ .

PROOF. First of all, assume that we have some $Z \in \mathcal{H}$ such that $Y_T^{y_0, Z} \geq \xi$, $\mathcal{P}_0^{\mathcal{U}}$ -q.s. Then, since $Y_T^{y_0, Z}$ is an $\mathcal{E}^{\hat{f}^{\mathbb{P}}}$ -supermartingale under \mathbb{P} for any $\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}$, we have

$$y_0 \geq \mathcal{Y}_0^{\mathbb{P}}(T, Y_T^{y_0, Z}), \quad \mathcal{P}_0^{\mathcal{U}}\text{-q.s.}$$

However, by the comparison result of Lemma A.3 (together with Lemma 2.2), we also have $\mathcal{Y}_0^{\mathbb{P}}(T, Y_T^{y_0, Z}) \geq \mathcal{Y}_0^{\mathbb{P}}(T, \xi)$, from which we deduce

$$y_0 \geq \mathcal{Y}_0^{\mathbb{P}}(T, \xi), \quad \mathbb{P}\text{-a.s.}$$

In particular, for any $\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}$, we deduce that

$$y_0 \geq \text{ess sup}_{\mathbb{P}' \in \mathcal{P}_0^{\mathcal{U}}(0, \mathbb{P}, \mathbb{F}^+)}^{\mathbb{P}} \mathcal{Y}_0^{\mathbb{P}'}(T, \xi) = Y_0, \quad \mathbb{P}\text{-a.s.},$$

where we have used Lemma 3.5. It therefore directly implies, since y_0 is deterministic, that

$$y_0 \geq \sup_{\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}} \mathbb{E}^{\mathbb{P}}[Y_0].$$

For the reverse inequality, let $(Y, Z, (K^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}}) \in \mathbb{D}^p(\mathbb{F}^{U+, \mathcal{P}_0^{\mathcal{U}}}) \times \mathbb{H}^p(\mathbb{F}^{U, \mathcal{P}_0^{\mathcal{U}}}) \times \mathbb{I}^{0,p}((\mathbb{F}_+^{\mathbb{P}})_{\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}})$ be the unique saturated solution to the 2BSDE with generator $\widehat{f}^{\mathbb{P}}$ and terminal condition ξ . Then we have for any $\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}$

$$Y_0 + \int_0^T \widehat{f}_s^{\mathbb{P}}(Y_s, (\widehat{a}_s^{1/2})^\top Z_s) ds + \int_0^T Z_s \cdot dX_s^{c, \mathbb{P}} = \xi + K_T^{\mathbb{P}} - K_t^{\mathbb{P}} \geq \xi, \quad \mathbb{P}\text{-a.s.}$$

However, since Y_0 is only $\mathcal{F}_0^{\mathcal{P}_0^{\mathcal{U}+}}$ -measurable, it is not, in general, deterministic, so that we cannot conclude directly. Let us nonetheless consider, for any $\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}$, $y_0^{\mathbb{P}}$ the smallest constant which dominates Y_0 , \mathbb{P} -a.s. We therefore want to show that for any $\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}$:

$$y_0^{\mathbb{P}} \leq \sup_{\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}} \mathbb{E}^{\mathbb{P}}[Y_0],$$

which can be done by following exactly the same arguments as in the proof of Theorem 3.2 in [69]. Finally, we do have $Z \in \mathcal{H}$, since by Lemma 3.4, Y is automatically an $\mathcal{E}^{\widehat{f}^{\mathbb{P}}}$ -supermartingale for every $\mathbb{P} \in \mathcal{P}_0^{\mathcal{U}}$. \square

6. Path-dependent PDEs. In the context of stochastic control theory, using the dynamic programming principle, we can characterize the value function as a viscosity solution of PPDE. Recall that μ, σ, U as well as \mathcal{U} are the same given in Section 5.2, we introduce a path-dependent PDE

$$(6.1) \quad \partial_t v(t, \omega) + G(t, \omega, v(t, \omega), \partial_\omega v, \partial_{\omega, \omega}^2 v) = 0,$$

where

$$G(t, \omega, y, z, \gamma) := \sup_{u \in U} \left\{ f(t, \omega, y, \sigma(\cdot)z, \mu, a)(t, \omega, u) + \mu(t, \omega, u) \cdot z + \frac{1}{2} a(t, \omega, u) : \gamma \right\}.$$

As in the survey of Ren, Touzi and Zhang [81] (see also [80]), one may define viscosity solutions of path dependent PDEs by using jets. For $\alpha \in \mathbb{R}, \beta \in \mathbb{R}^d, \gamma \in \mathbb{S}^d$, denote

$$\phi^{\alpha, \beta, \gamma}(t, x) := \alpha t + \beta \cdot x + \frac{1}{2} \gamma : (xx^T) \quad \text{for all } (t, x) \in \mathbb{R}^+ \times \mathbb{R}^d,$$

where $A_1 : A_2 := \text{Tr}[A_1 A_2]$. Let $\text{BUC}([0, T] \times \Omega)$ denote the set of all bounded functions in Ω which are in addition uniformly continuous w.r.t. the metric d defined by

$$d((t, \omega), (t', \omega')) := \sqrt{|t - t'|} + \|\omega_{t \wedge \cdot} - \omega'_{t' \wedge \cdot}\|_\infty.$$

Then define the semi-jets of a function $u \in \text{BUC}([0, T] \times \Omega)$ at the point $(t, \omega) \in [0, T] \times \Omega$ by

$$\underline{\mathcal{J}}u(t, \omega) := \left\{ (\alpha, \beta, \gamma) : u(t, \omega) = \max_{\tau \in \mathcal{T}_{H_\delta}} \bar{\mathcal{E}}_{t, \omega}[u_\tau - \phi_\tau^{\alpha, \beta, \gamma, t, \omega}], \text{ for some } \delta > 0 \right\},$$

$$\bar{\mathcal{J}}u(t, \omega) := \left\{ (\alpha, \beta, \gamma) : u(t, \omega) = \min_{\tau \in \mathcal{T}_{H_\delta}} \underline{\mathcal{E}}_{t, \omega}[u_\tau - \phi_\tau^{\alpha, \beta, \gamma, t, \omega}], \text{ for some } \delta > 0 \right\},$$

where

$$\bar{\mathcal{E}}_{t, \omega}[\xi] := \sup_{\mathbb{P} \in \mathcal{P}_{t, \omega}^{\mathcal{U}}} \mathbb{E}^{\mathbb{P}}[\xi], \quad \underline{\mathcal{E}}_{t, \omega}[\xi] := \inf_{\mathbb{P} \in \mathcal{P}_{t, \omega}^{\mathcal{U}}} \mathbb{E}^{\mathbb{P}}[\xi]$$

and $H_\delta(\omega') := \delta \wedge \inf\{s \geq 0 : |\omega'_s| \geq \delta\}$ and \mathcal{T}_{H_δ} denotes the collection of all \mathbb{F} -stopping times larger than H_δ , and $\phi_s^{\alpha, \beta, \gamma, t, \omega} := \phi^{\alpha, \beta, \gamma}(s - t, X_s - \omega_t)$.

DEFINITION 6.1. Let $u \in \text{BUC}([0, T] \times \Omega)$.

(i) u is a $\mathcal{P}^{\mathcal{U}}$ -viscosity subsolution (resp., super-solution) of the path dependent PDE (6.1), if at any point $(t, \omega) \in [0, T] \times \Omega$ it holds for all $(\alpha, \beta, \gamma) \in \underline{\mathcal{J}}u(t, \omega)$ (resp. $\bar{\mathcal{J}}u(t, \omega)$) that

$$-\alpha - G(t, \omega, u(t, \omega), \beta, \gamma) \leq (\text{resp. } \geq) 0.$$

(ii) u is a $\mathcal{P}^{\mathcal{U}}$ -viscosity solution of PPDE (6.1), if u is both a $\mathcal{P}^{\mathcal{U}}$ -viscosity subsolution and a $\mathcal{P}^{\mathcal{U}}$ -viscosity super-solution of (6.1).

Using the dynamic programming principle, and by exactly the same arguments as in [31], Section 4.3, we can characterize the value Y as viscosity solution of the above PPDE.

THEOREM 6.1. Let Assumptions 2.1 and 3.1 hold. Suppose in addition that $\omega \mapsto \xi(\omega)$ and $\omega \mapsto f(t, \omega, y, z, b, a)$ are uniformly continuous with respect to $\|\cdot\|_\infty$. Then the value function $v(t, \omega) := Y_t(\omega) \in \text{BUC}([0, T] \times \Omega)$ and v is a viscosity solution of PPDE (6.1).

Of course, in order to have a complete characterization of the solution to a 2BSDE as viscosity solution of the corresponding PPDE, the above result has to be complemented with a comparison theorem. In the case of a fully nonlinear PPDE, such a result has been recently achieved by Ren, Touzi and Zhang [82]. However, their main result, Theorem 4.2, needs to consider viscosity subsolutions and super-solutions in a smaller set than $\text{BUC}([0, T] \times \Omega)$. Namely, define for any $\ell > 0$ the set $\text{BUC}^\ell([0, T] \times \Omega)$ of all bounded functions in Ω which are in addition uniformly continuous w.r.t. the metric d^ℓ defined by

$$d^\ell((t, \omega), (t', \omega')) := \sqrt{|t - t'|} + \|\omega_{\cdot \wedge t} - \omega'_{\cdot \wedge t}\|_\ell,$$

where for any $\omega \in \Omega$

$$\|\omega\|_\ell := \left(\int_0^{T+1} \|\omega_s\|^\ell ds \right)^{\frac{1}{\ell}}.$$

We then have the following result.

THEOREM 6.2. *Let Assumptions 2.1, 3.1 hold and let G satisfy Assumption 4.1 of [82]. Suppose in addition that $\omega \mapsto \xi(\omega)$ and $\omega \mapsto f(t, \omega, y, z, b, a)$ are uniformly continuous with respect to $\|\cdot\|_\ell$ for some $\ell \leq p$. Then the value function $Y_t(\omega)$ is the unique viscosity solution of PPDE (6.1) in $BUC^\ell([0, T] \times \Omega)$.*

PROOF. The only thing to prove here is that Y does belong to $BUC^\ell([0, T] \times \Omega)$, since we can then apply immediately Theorem 4.2 of [82]. However, this regularity can be obtained from classical *a priori* estimates for BSDEs, and arguments similar to the ones used in Example 7.1 of [82]. \square

APPENDIX

In this appendix, we collect several results related to BSDE theory, which are used throughout the paper. We fix $r \in [0, T]$ and some $\mathbb{P} \in \mathcal{P}(r, \omega)$. A generator will here be a map $g : [r, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$, which is \mathbb{F}_+ -progressively measurable and uniformly Lipschitz in (y, z) , satisfying

$$\mathbb{E}^\mathbb{P} \left[\int_r^T |g_s(0, 0)|^p ds \right] < +\infty.$$

Similarly, a terminal condition will be a \mathcal{F}_T -measurable random variable in $\mathbb{L}_r^p(\mathbb{F}^+, \mathbb{P})$. To state our results, we will actually need to work on the enlarged canonical space $\bar{\Omega}$, but we remind the reader that by Lemma 2.2, it is purely a technical tool. Let $\bar{\mathbb{P}} := \mathbb{P} \otimes \mathbb{P}_0$. We will then say that $(y, z, m) \in \mathbb{D}_r^p(\bar{\mathbb{F}}^{\bar{\mathbb{P}}^+}, \bar{\mathbb{P}}) \times \mathbb{H}_r^p(\bar{\mathbb{F}}^{\bar{\mathbb{P}}}, \bar{\mathbb{P}}) \times \mathbb{M}_r^p(\bar{\mathbb{F}}^{\bar{\mathbb{P}}^+}, \bar{\mathbb{P}})$ is a solution to the BSDE with generator g and terminal condition ξ if

$$\begin{aligned} (A.1) \quad y_t &= \xi(X_\cdot) - \int_t^T g_s(y_s, (\hat{a}_s^{1/2})^\top z_s) ds \\ &\quad - \int_t^T z_s \cdot \hat{a}_s^{1/2} dW_s^\mathbb{P} - \int_t^T dm_s, \quad t \in [r, T], \bar{\mathbb{P}}\text{-a.s.} \end{aligned}$$

Similarly, if we are given a process $k \in \mathbb{I}_r^p(\bar{\mathbb{F}}^{\bar{\mathbb{P}}^+}, \bar{\mathbb{P}})$, we call (y, z, m, k) a super-solution of the BSDE with generator g and terminal condition ξ if

$$\begin{aligned} (A.2) \quad y_t &= \xi - \int_t^T g_s(y_s, (\hat{a}_s^{1/2})^\top z_s) ds - \int_t^T z_s \cdot \hat{a}_s^{1/2} dW_s^\mathbb{P} \\ &\quad - \int_t^T dm_s + \int_t^T dk_s, \quad t \in [r, T], \bar{\mathbb{P}}\text{-a.s.} \end{aligned}$$

A.1. Technical results for BSDEs.

LEMMA A.1 (Estimates and stability). *Let Assumption 2.1 hold. Then, for $i = 1, 2$, let us denote by (y^i, z^i, m^i) the solution of the BSDE (A.1) with generator g^i and terminal condition ξ^i . Then, for any $\kappa \in (1, p]$, there exists some constant $C > 0$ such that*

$$|y_t^i| \leq C \left(\mathbb{E}^{\mathbb{P}} \left[|\xi^i|^\kappa + \int_r^T |g_s^i(0, 0)|^\kappa ds \middle| \mathcal{F}_t^+ \right] \right)^{\frac{1}{\kappa}}, \quad t \in [r, T], \bar{\mathbb{P}}\text{-a.s.},$$

and

$$\|z^i\|_{\mathbb{H}^p(\bar{\mathbb{P}})}^p + \|m^i\|_{\mathbb{M}^p(\bar{\mathbb{P}})}^p \leq C \left(\|\xi^i\|_{\mathbb{L}^p(\bar{\mathbb{P}})}^p + \mathbb{E}^{\mathbb{P}} \left[\int_r^T |g_s^i(0, 0)|^p ds \right] \right).$$

Denoting $\delta\xi := \xi^1 - \xi^2$, $\delta y := y^1 - y^2$, $\delta z := z^1 - z^2$, $\delta m := m^1 - m^2$, $\delta g := (g^1 - g^2)(\cdot, y^1, z^1)$, we also have

$$|\delta y_t| \leq C \left(\mathbb{E}^{\mathbb{P}} \left[|\delta\xi|^\kappa + \int_r^T |\delta g_s|^\kappa ds \middle| \mathcal{F}_t^+ \right] \right)^{\frac{1}{\kappa}}, \quad t \in [r, T], \bar{\mathbb{P}}\text{-a.s.},$$

and

$$\|\delta z\|_{\mathbb{H}^p(\bar{\mathbb{P}})}^p + \|\delta m\|_{\mathbb{M}^p(\bar{\mathbb{P}})}^p \leq C \left(\|\delta\xi\|_{\mathbb{L}^p(\bar{\mathbb{P}})}^p + \mathbb{E}^{\mathbb{P}} \left[\int_r^T |\delta g_s|^p ds \right] \right).$$

PROOF. See Section 4 of [13]. \square

LEMMA A.2. *For any \mathbb{F} -stopping times $0 \leq r \leq \rho \leq \tau \leq T$, any decreasing sequence of \mathbb{F} -stopping times $(\tau_n)_{n \geq 1}$ converging \mathbb{P} -a.s. to τ , and any \mathbb{F}^+ -progressively measurable and right-continuous process $V \in \mathbb{D}_r^p(\mathbb{F}_+^p, \mathbb{P})$, if $y(\cdot, V)$ denotes the first component of the solution to the BSDE (A.1) on $[r, \cdot]$ with terminal condition V and some generator g , we have*

$$\mathbb{E}^{\mathbb{P}} [|y_\rho(\tau, V_\tau) - y_\rho(\tau_n, V_{\tau_n})|] \xrightarrow{n \rightarrow +\infty} 0.$$

PROOF. First of all, by Lemma (2.7), we have

$$y_\rho(\tau, V_\tau) - y_\rho(\tau_n, V_{\tau_n}) = y_\rho(\tau, V_\tau) - y_\rho(\tau, y_\tau(\tau_n, V_{\tau_n})).$$

By Lemma A.1, we therefore have for any $\kappa \in (1, p]$:

$$\mathbb{E}^{\mathbb{P}} [|y_\rho(\tau, V_\tau) - y_\rho(\tau, y_\tau(\tau_n, V_{\tau_n}))|] \leq C \mathbb{E}^{\mathbb{P}} [|V_\tau - y_\tau(\tau_n, V_{\tau_n})|^\kappa].$$

Next, again by a linearization argument, we can find bounded processes λ and η , which are $\bar{\mathbb{F}}$ -progressively measurable such that

$$y_\tau(\tau_n, V_{\tau_n}) = \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} \left[\mathcal{E} \left(\int_\tau^{\tau_n} \eta_s \cdot dW_s^{\mathbb{P}} \right) \times \left(e^{\int_\tau^{\tau_n} \lambda_s ds} V_{\tau_n} - \int_\tau^{\tau_n} e^{\int_\tau^s \lambda_u du} g_s(0, 0) ds \right) \middle| \mathcal{F}_\tau^+ \right].$$

Hence, choosing $\kappa < \tilde{p} < p$,

$$\begin{aligned} & \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} [|y_\rho(\tau, V_\tau) - y_\rho(\tau, y_\tau(\tau_n, V_{\tau_n}))|] \\ & \leq C \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} \left[\mathcal{E} \left(\int_\tau^{\tau_n} \eta_s \cdot dW_s^\mathbb{P} \right)^\kappa e^{\kappa \int_\tau^{\tau_n} \lambda_s ds} |V_{\tau_n} - V_\tau|^\kappa \right] \\ & \quad + C \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} \left[\left| 1 - \mathcal{E} \left(\int_\tau^{\tau_n} \eta_s \cdot dW_s^\mathbb{P} \right) e^{\int_\tau^{\tau_n} \lambda_s ds} \right|^\kappa |V_\tau|^\kappa \right] \\ & \quad + C \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} \left[\mathcal{E} \left(\int_\tau^{\tau_n} \eta_s \cdot dW_s^\mathbb{P} \right)^\kappa \int_\tau^{\tau_n} e^{\kappa \int_\tau^{\tau_n} \lambda_s ds} |g_s(0, 0)|^\kappa ds \right] \\ & \leq C (\mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} [|V_{\tau_n} - V_\tau|^{\tilde{p}}])^{\frac{\kappa}{\tilde{p}}} \\ & \quad + C \left(\mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} \left[\left| 1 - \mathcal{E} \left(\int_\tau^{\tau_n} \eta_s \cdot dW_s^\mathbb{P} \right) e^{\int_\tau^{\tau_n} \lambda_s ds} \right|^{\frac{p}{p-\kappa}} \right] \right)^{\frac{p-\kappa}{p}} \\ & \quad + C \mathbb{E}^{\mathbb{P} \otimes \mathbb{P}_0} \left[\int_\tau^{\tau_n} e^{\tilde{p} \int_\tau^{\tau_n} \lambda_s ds} |g_s(0, 0)|^{\tilde{p}} ds \right], \end{aligned}$$

where we have used Hölder inequality, that λ is bounded and the fact that since η is also bounded, the Doléans–Dade exponential appearing above has finite moments of any order. Now the terms inside the expectations on the right-hand side all converge in probability to 0 and are clearly uniformly integrable by de la Vallée–Poussin criterion since $V \in \mathbb{D}_r^p(\mathbb{F}_+^\mathbb{P}, \mathbb{P})$ and $\tilde{p} < p$. We can therefore conclude by dominated convergence. \square

LEMMA A.3 (Comparison). *Let Assumption 2.1 hold. Then, for $i = 1, 2$, let us denote by (y^i, z^i, m^i, k^i) the super-solution of the BSDE (A.2) with generator g^i and terminal condition ξ^i . If it holds \mathbb{P} -a.s. that*

$$\xi_1 \geq \xi_2, \quad k^1 - k^2 \text{ is nondecreasing} \quad \text{and} \quad g^1(s, y_s^1, z_s^1) \geq g^2(s, y_s^1, z_s^1),$$

then we have for all $t \in [0, T]$

$$y_t^1 \geq y_t^2, \quad \mathbb{P}\text{-a.s.}$$

PROOF. We remind the reader that since $W^\mathbb{P}$ and $m^i, i = 1, 2$ are orthogonal and since $W^\mathbb{P}$ is actually continuous, we not only have $[W^\mathbb{P}, m^i] = 0, \bar{\mathbb{P}}\text{-a.s.}$, but also

$$\langle W^\mathbb{P}, m^i \rangle = \langle W^\mathbb{P}, m^{i,c,\bar{\mathbb{P}}} \rangle = \langle W^\mathbb{P}, m^{i,d,\bar{\mathbb{P}}} \rangle = 0, \quad \bar{\mathbb{P}}\text{-a.s.},$$

where $m^{i,c,\bar{\mathbb{P}}}$ (resp., $m^{i,d,\bar{\mathbb{P}}}$) is the continuous (resp., purely discontinuous) martingale part of m^i , under the measure $\bar{\mathbb{P}}$.

Then, since the g^i are uniformly Lipschitz, there exist two processes λ and η which are bounded, $\bar{\mathbb{P}}\text{-a.s.}$, and which are respectively $\bar{\mathbb{F}}_+^{\bar{\mathbb{P}}}$ -progressively measurable and $\bar{\mathbb{F}}^{\bar{\mathbb{P}}}$ -predictable, such that

$$g^2(s, y_s^1, z_s^1) - g^2(s, y_s^2, z_s^2) = \lambda_t (y_s^1 - y_s^2) + \eta_s (z_s^1 - z_s^2), \quad ds \times d\bar{\mathbb{P}}\text{-a.e.}$$

For any $0 \leq t \leq s \leq T$, let us define the following continuous, positive and $\overline{\mathbb{F}}_+$ -progressively measurable process

$$A_{t,s} := \exp\left(\int_t^s \lambda_u du - \int_t^s \eta_u \cdot dW_u^{\mathbb{P}} - \frac{1}{2} \int_t^s \|\eta_u\|^2 du\right).$$

By Itô's formula, we deduce classically that

$$y_t^1 - y_t^2 = \mathbb{E}^{\overline{\mathbb{P}}}\left[A_{t,T}(\xi^1 - \xi^2) + \int_t^T A_{t,s}[(g^1 - g^2)(s, y_s^1, z_s^1) ds + d(k_s^1 - k_s^2)] \middle| \overline{\mathcal{F}}_{t+}\right],$$

from which we deduce immediately that $y_t^1 \geq y_t^2$, $\overline{\mathbb{P}}$ -a.s. \square

REFERENCES

- [1] AİD, R., POSSAMAİ, D. and TOUZI, N. (2016). A principal–agent model for pricing electricity volatility demand. Preprint.
- [2] AVELLANEDA, M., LEVY, A. and PARÁS, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Appl. Math. Finance* **2** 73–88.
- [3] BARRIEU, P. and EL KAROUI, N. (2013). Monotone stability of quadratic semimartingales with applications to unbounded general quadratic BSDEs. *Ann. Probab.* **41** 1831–1863.
- [4] BAYRAKTAR, E. and SİRBU, M. (2012). Stochastic Perron's method and verification without smoothness using viscosity comparison: The linear case. *Proc. Amer. Math. Soc.* **140** 3645–3654.
- [5] BAYRAKTAR, E. and SİRBU, M. (2013). Stochastic Perron's method for Hamilton–Jacobi–Bellman equations. *SIAM J. Control Optim.* **51** 4274–4294.
- [6] BERTSEKAS, D. P. and SHREVE, S. E. (1978). *Stochastic Optimal Control: The Discrete Time Case, Vol. 23*. Academic Press, New York.
- [7] BIAGINI, S., BOUCHARD, B., KARDARAS, C. and NUTZ, M. (2017). Robust fundamental theorem for continuous processes. *Math. Finance*. **27** 963–987. [MR3705159](#)
- [8] BOUCHARD, B., MOREAU, L. and NUTZ, M. (2014). Stochastic target games with controlled loss. *Ann. Appl. Probab.* **24** 899–934. [MR3199977](#)
- [9] BOUCHARD, B. and NUTZ, M. (2012). Weak dynamic programming for generalized state constraints. *SIAM J. Control Optim.* **50** 3344–3373.
- [10] BOUCHARD, B. and NUTZ, M. (2015). Arbitrage and duality in nondominated discrete-time models. *Ann. Appl. Probab.* **25** 823–859. [MR3313756](#)
- [11] BOUCHARD, B. and NUTZ, M. (2016). Stochastic target games and dynamic programming via regularized viscosity solutions. *Math. Oper. Res.* **41** 109–124.
- [12] BOUCHARD, B., POSSAMAİ, D. and TAN, X. (2016). A general Doob–Meyer–Mertens decomposition for g -supermartingale systems. *Electron. J. Probab.* **21** 1–21.
- [13] BOUCHARD, B., POSSAMAİ, D., TAN, X. and ZHOU, C. (2015). A unified approach to a priori estimates for supersolutions of BSDEs in general filtrations. *Ann. Inst. Henri Poincaré B, Probab. Stat.* To appear.
- [14] BOUCHARD, B. and TOUZI, N. (2011). Weak dynamic programming principle for viscosity solutions. *SIAM J. Control Optim.* **49** 948–962.
- [15] CHERIDITO, P., SONER, H. M., TOUZI, N. and VICTOIR, N. (2007). Second–order backward stochastic differential equations and fully nonlinear parabolic pdes. *Comm. Pure Appl. Math.* **60** 1081–1110.
- [16] CLAISSE, J., TALAY, D. and TAN, X. (2016). A pseudo–Markov property for controlled diffusion processes. *SIAM J. Control Optim.* **54** 1017–1029.

- [17] COHEN, S. N. (2012). Representing filtration consistent nonlinear expectations as g -expectations in general probability spaces. *Stochastic Process. Appl.* **122** 1601–1626.
- [18] COQUET, F., HU, Y., MÉMIN, J. and PENG, S. (2002). Filtration–consistent nonlinear expectations and related g -expectations. *Probab. Theory Related Fields* **123** 1–27.
- [19] CVITANIĆ, J., KARATZAS, I. and SONER, H. M. (1998). Backward stochastic differential equations with constraints on the gains-process. *Ann. Probab.* **24** 1522–1551. [MR1675035](#)
- [20] CVITANIĆ, J., POSSAMAÏ, D. and TOUZI, N. (2014). Moral hazard in dynamic risk management. *Manage. Sci.* To appear.
- [21] CVITANIĆ, J., POSSAMAÏ, D. and TOUZI, N. (2015). Dynamic programming approach to principal–agent problems. Preprint. Available at [arXiv:1510.07111](#).
- [22] CVITANIĆ, J. and ZHANG, J. (2012). *Contract Theory in Continuous–Time Models*. Springer, Berlin.
- [23] DARLING, R. W. R. and PARDOUX, É. (1997). Backwards SDE with random terminal time and applications to semilinear elliptic PDE. *Ann. Probab.* **25** 1135–1159. [MR1457614](#)
- [24] DELLACHERIE, C. (1985). Quelques résultats sur les maisons de jeu analytiques. *Séminaire de Probabilités de Strasbourg XIX* 222–229.
- [25] DELLACHERIE, C. and MEYER, P.-A. (1979). *Probabilities and Potential. Mathematics Studies 29*. North-Holland, Amsterdam.
- [26] DENIS, L. and MARTINI, C. (2006). A theoretical framework for the pricing of contingent claims in the presence of model uncertainty. *Ann. Appl. Probab.* **16** 827–852. [MR2244434](#)
- [27] DOLINSKY, Y. and SONER, H. M. (2014). Martingale optimal transport and robust hedging in continuous time. *Probab. Theory Related Fields* **160** 391–427.
- [28] DOOB, J. L. (1984). *Classical Potential Theory and Its Probabilistic Counterpart. Grundlehren der Mathematischen Wissenschaften 262*. Springer, Berlin.
- [29] DUMITRESCU, R., QUENEZ, M.-C. and SULEM, A. (2016). A weak dynamic programming principle for combined optimal stopping/stochastic control with \mathcal{E}^f -expectations. *SIAM J. Control Optim.* **54** 2090–2115.
- [30] EKREN, I., KELLER, C., TOUZI, N. and ZHANG, J. (2014). On viscosity solutions of path dependent PDEs. *Ann. Probab.* **42** 204–236. [MR3161485](#)
- [31] EKREN, I., TOUZI, N. and ZHANG, J. (2016). Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part I. *Ann. Probab.* **44** 1212–1253. [MR3474470](#)
- [32] EKREN, I., TOUZI, N. and ZHANG, J. (2016). Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II. *Ann. Probab.* **44** 2507–2553. [MR3531674](#)
- [33] EL KAROUI, N. (1981). Les aspects probabilistes du contrôle stochastique. In *École D’été de Probabilités de Saint–Flour IX—1979. Lecture Notes in Math.* **876** 73–238. Springer, Berlin.
- [34] EL KAROUI, N. and HUANG, S.-J. (1997). A general result of existence and uniqueness of backward stochastic differential equations. In *Backward Stochastic Differential Equations* (N. El Karoui and L. Mazliak, eds.). *Chapman & Hall/CRC Research Notes in Mathematics Series 364* 27–36. Longman, Harlow.
- [35] EL KAROUI, N., HUU NGUYEN, D. and JEANBLANC-PICQUÉ, M. (1987). Compactification methods in the control of degenerate diffusions: Existence of an optimal control. *Stochastics* **20** 169–219.
- [36] EL KAROUI, N., KAPOUDJIAN, C., PARDOUX, É., PENG, S. and QUENEZ, M.-C. (1997). Reflected solutions of backward SDE’s, and related obstacle problems for PDE’s. *Ann. Probab.* **25** 702–737. [MR1434123](#)
- [37] EL KAROUI, N., PENG, S. and QUENEZ, M.-C. (1997). Backward stochastic differential equations in finance. *Math. Finance* **7** 1–71.
- [38] EL KAROUI, N. and QUENEZ, M.-C. (1991). Programmation dynamique et évaluation des actifs contingents en marché incomplet. *C. R. Acad. Sci., Sér. 1 Math.* **313** 851–854.

- [39] EL KAROUI, N. and QUENEZ, M.-C. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM J. Control Optim.* **33** 29–66.
- [40] EL KAROUI, N. and TAN, X. (2013). Capacities, measurable selection and dynamic programming part I: Abstract framework. Preprint. Available at [arXiv:1310.3363](https://arxiv.org/abs/1310.3363).
- [41] EL KAROUI, N. and TAN, X. (2013). Capacities, measurable selection and dynamic programming part II: Application in stochastic control problems. Preprint. Available at [arXiv:1310.3364](https://arxiv.org/abs/1310.3364).
- [42] FLEMING, W. H. and SONER, H. M. (2006). *Controlled Markov Processes and Viscosity Solutions*, 2nd ed. *Stochastic Modelling and Applied Probability* **25**. Springer, New York.
- [43] FÖLLMER, H. and KABANOV, YU. M. (1997). Optional decomposition and Lagrange multipliers. *Finance Stoch.* **2** 69–81.
- [44] FÖLLMER, H. and KRAMKOV, D. O. (1997). Optional decompositions under constraints. *Probab. Theory Related Fields* **109** 1–25.
- [45] FREMLIN, D. H. (2008). *Measure Theory, Vol. 5: Set–Theoretic Measure Theory*. Torres Fremlin.
- [46] HAMADÈNE, S. and LEPETIER, J.-P. (1995). Backward equations, stochastic control and zero–sum stochastic differential games. *Stochastics* **54** 221–231.
- [47] HU, M., JI, S., PENG, S. and SONG, Y. (2014). Backward stochastic differential equations driven by G -Brownian motion. *Stochastic Process. Appl.* **124** 759–784.
- [48] HU, M., JI, S., PENG, S. and SONG, Y. (2014). Comparison theorem, Feynman–Kac formula and Girsanov transformation for BSDEs driven by G -Brownian motion. *Stochastic Process. Appl.* **124** 1170–1195.
- [49] HU, Y., MA, J., PENG, S. and YAO, S. (2008). Representation theorems for quadratic f -consistent nonlinear expectations. *Stochastic Process. Appl.* **118** 1518–1551.
- [50] JACOD, J. and SHIRYAEV, A. (2013). *Limit Theorems for Stochastic Processes*, Vol. 288. Springer, Berlin.
- [51] KARANDIKAR, R. L. (1995). On pathwise stochastic integration. *Stochastic Process. Appl.* **57** 11–18.
- [52] KAZI-TANI, N., POSSAMAÏ, D. and ZHOU, C. (2015). Second–order BSDEs with jumps: Formulation and uniqueness. *Ann. Appl. Probab.* **25** 2867–2908.
- [53] KAZI-TANI, N., POSSAMAÏ, D. and ZHOU, C. (2015). Second order BSDEs with jumps: Existence and probabilistic representation for fully–nonlinear PIDEs. *Electron. J. Probab.* **20** 1–31.
- [54] KOBYLANSKI, M. (1997). Résultats d’existence et d’unicité pour des équations différentielles stochastiques rétrogrades avec des générateurs à croissance quadratique. *C. R. Acad. Sci., Ser. I Math.* **324** 81–86.
- [55] KOBYLANSKI, M. (2000). Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.* **28** 558–602. [MR1782267](https://doi.org/10.1214/1782267)
- [56] KRAMKOV, D. O. (1996). Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. *Probab. Theory Related Fields* **105** 459–479.
- [57] LARSON, P. B. (2009). The filter dichotomy and medial limits. *J. Math. Log.* **9** 159–165.
- [58] LENGART, É. (1980). Tribus de Meyer et théorie des processus. *Séminaire de Probabilités de Strasbourg* **XIV** 500–546.
- [59] LEPETIER, J.-P. and SAN MARTÍN, J. (1997). Backward stochastic differential equations with continuous coefficient. *Statist. Probab. Lett.* **32** 425–430.
- [60] LYONS, T. J. (1995). Uncertain volatility and the risk–free synthesis of derivatives. *Appl. Math. Finance* **2** 117–133.
- [61] MASTROLIA, T. and POSSAMAÏ, D. (2015). Moral hazard under ambiguity. Preprint. Available at [arXiv:1511.03616](https://arxiv.org/abs/1511.03616).
- [62] MATOUSSI, A., PIOZIN, L. and POSSAMAÏ, D. (2014). Second–order BSDEs with general reflection and game options under uncertainty. *Stochastic Process. Appl.* **124** 2281–2321.

- [63] MATOUSSI, A., POSSAMAÏ, D. and ZHOU, C. (2013). Second order reflected backward stochastic differential equations. *Ann. Appl. Probab.* **23** 2420–2457. [MR3127940](#)
- [64] MATOUSSI, A., POSSAMAÏ, D. and ZHOU, C. (2015). Robust utility maximization in nondominated models with 2BSDE: The uncertain volatility model. *Math. Finance* **25** 258–287.
- [65] NEUFELD, A. and NUTZ, M. (2013). Superreplication under volatility uncertainty for measurable claims. *Electron. J. Probab.* **18** 1–14.
- [66] NEUFELD, A. and NUTZ, M. (2014). Measurability of semimartingale characteristics with respect to the probability law. *Stochastic Process. Appl.* **124** 3819–3845.
- [67] NEVEU, J. (1975). *Discrete-Parameter Martingales*, Vol. 10. North-Holland, Amsterdam.
- [68] NUTZ, M. (2012). Pathwise construction of stochastic integrals. *Electron. Commun. Probab.* **17** 1–7.
- [69] NUTZ, M. (2015). Robust superhedging with jumps and diffusion. *Stochastic Process. Appl.* **125** 4543–4555. [MR3406595](#)
- [70] NUTZ, M. and SONER, H. M. (2012). Superhedging and dynamic risk measures under volatility uncertainty. *SIAM J. Control Optim.* **50** 2065–2089.
- [71] NUTZ, M. and VAN HANDEL, R. (2013). Constructing sublinear expectations on path space. *Stochastic Process. Appl.* **123** 3100–3121.
- [72] NUTZ, M. and ZHANG, J. (2015). Optimal stopping under adverse nonlinear expectation and related games. *Ann. Appl. Probab.* **25** 2503–2534. [MR3375882](#)
- [73] PARDOUX, É. (1999). BSDEs, weak convergence and homogenization of semilinear PDEs. In *Nonlinear Analysis, Differential Equations and Control* (F. H. Clarke, R. J. Stern and G. Sabidussi, eds.). *NATO Science Series* **528** 503–549. Springer, Berlin.
- [74] PENG, S. (1997). Backward SDE and related g -expectation. In *Backward Stochastic Differential Equations* (N. El Karoui and L. Mazliak, eds.). *Pitman Research Notes in Mathematics* **364** 141–160. Longman, Harlow.
- [75] PENG, S. (1999). Monotonic limit theorem of BSDE and nonlinear decomposition theorem of Doob–Meyer’s type. *Probab. Theory Related Fields* **113** 473–499.
- [76] PENG, S. (2013). Nonlinear expectation theory and stochastic calculus under Knightian uncertainty. In *Real Options, Ambiguity, Risk and Insurance* (A. Bensoussan, S. Peng and J. Sung, eds.). *Studies in Probability, Optimization and Statistics* **5**. IOS Press, Amsterdam.
- [77] POSSAMAÏ, D. (2013). Second order backward stochastic differential equations under a monotonicity condition. *Stochastic Process. Appl.* **123** 1521–1545.
- [78] POSSAMAÏ, D., ROYER, G. and TOUZI, N. (2013). On the robust superhedging of measurable claims. *Electron. Commun. Probab.* **18** 1–13.
- [79] POSSAMAÏ, D. and ZHOU, C. (2013). Second order backward stochastic differential equations with quadratic growth. *Stochastic Process. Appl.* **123** 3770–3799.
- [80] REN, Z., TOUZI, N. and ZHANG, J. (2014). Comparison of viscosity solutions of semi-linear path-dependent PDEs. Preprint. Available at [arXiv:1410.7281](#).
- [81] REN, Z., TOUZI, N. and ZHANG, J. (2014). An overview of viscosity solutions of path-dependent PDEs. In *Stochastic Analysis and Applications 2014: In Honour of Terry Lyons* (D. Crisan, B. Hambly and T. Zariphopoulou, eds.). *Springer Proceedings in Mathematics and Statistics* **100** 397–453. Springer, Berlin.
- [82] REN, Z., TOUZI, N. and ZHANG, J. (2015). Comparison of viscosity solutions of fully nonlinear degenerate parabolic path-dependent PDEs. Preprint. Available at [arXiv:1511.05910](#).
- [83] SONER, H. M., TOUZI, N. and ZHANG, J. (2011). Martingale representation theorem for the G -expectation. *Stochastic Process. Appl.* **121** 265–287.
- [84] SONER, H. M., TOUZI, N. and ZHANG, J. (2012). Wellposedness of second order backward SDEs. *Probab. Theory Related Fields* **153** 149–190.
- [85] SONER, H. M., TOUZI, N. and ZHANG, J. (2013). Dual formulation of second order target problems. *Ann. Appl. Probab.* **23** 308–347. [MR3059237](#)

- [86] SONG, Y. (2011). Some properties on G -evaluation and its applications to G -martingale decomposition. *Sci. China Math.* **54** 287–300.
- [87] STROOCK, D. W. and VARADHAN, S. R. S. (1997). *Multidimensional Diffusion Processes*. *Grundlehren der Mathematischen Wissenschaften* **233**. Springer, Berlin.

D. POSSAMAÏ
X. TAN
CEREMADE
UNIVERSITÉ PARIS DAUPHINE
PSL RESEARCH UNIVERSITY
PLACE DU MARÉCHAL DE LATTRE DE
TASSIGNY
75775 PARIS CEDEX 16
FRANCE
E-MAIL: possamai@ceremade.dauphine.fr
tan@ceremade.dauphine.fr

C. ZHOU
DEPARTMENT OF MATHEMATICS
NATIONAL UNIVERSITY OF SINGAPORE
10 LOWER KENT RIDGE ROAD
119076, SINGAPORE
SINGAPORE
E-MAIL: matzc@nus.edu.sg