BSDEs with weak terminal condition

Bruno Bouchard, Romuald Elie, Anthony Réveillac

BSDEs with weak terminal condition

Bruno Bouchard* Romuald Elie† Anthony Réveillac‡

Université Paris-Dauphine
CEREMADE UMR CNRS 7534
Place du Maréchal De Lattre De Tassigny
75775 Paris cedex 16 France

October 18, 2012

Abstract

We introduce a new class of Backward Stochastic Differential Equations in which the T-terminal value Y_T of the solution (Y, Z) is not fixed as a random variable, but only satisfies a weak constraint of the form $E[\Psi(Y_T)] \geq m$, for some (possibly random) non-decreasing map Ψ and some threshold m. We name them *BSDEs with weak terminal condition* and obtain a representation of the minimal time t-values Y_t such that (Y, Z) is a supersolution of the BSDE with weak terminal condition. It provides a non-Markovian BSDE formulation of the PDE characterization obtained for Markovian stochastic target problems under controlled loss in Bouchard, Elie and Touzi [3]. We then study the main properties of this minimal value. In particular, we analyze its continuity and convexity with respect to the m-parameter appearing in the weak terminal condition, and show how it can be related to a dual optimal control problem in Meyer form. These last properties generalize to a non Markovian framework previous results on quantile hedging and hedging under loss constraints obtained in Föllmer and Leukert [10, 11], and in Bouchard, Elie and Touzi [3]. Finally, we observe a surprisingly strong connection between BSDEs with weak terminal condition and 2nd order BSDEs in the quasi linear case.

Key words: Backward stochastic differential equations, optimal control, stochastic target.

MSC Classification (2000): Primary: 60H10; 93E20; Secondary: 49L20; 91G80

*and CREST, bouchard@ceremade.dauphine.fr
†elie@ceremade.dauphine.fr
‡anthony.reveillac@ceremade.dauphine.fr

1 Introduction

Solving a backward stochastic differential equation (hereafter BSDE), with terminal data $\xi \in \mathbf{L}_2(\mathcal{F}_T)$ and driver g, consists in finding a pair of predictable processes (Y, Z), with certain integrability properties, such that the dynamics of Y satisfies $dY_t = -g(t, Y_t, Z_t)dt + Z_t dW_t$ and $Y_T = \xi$ (where W denotes a standard Brownian motion). It can be rephrased in: find an initial data Y_0 and a control process Z such that the solution Y^Z of the controlled stochastic differential equation

$$Y_t^Z = Y_0 - \int_0^t g(s, Y_s^Z, Z_s)ds + \int_0^t Z_s dW_s \ , \qquad 0 \leq t \leq T \ , \tag{1.1}$$

satisfies $Y_T^Z = \xi$. In cases where the previous problem does not admit a solution, a weaker formulation is to find an initial data Y_0 and a control Z such that

$$Y_T^Z \geq \xi \ \ \mathbb{P} - \text{a.s.} \tag{1.2}$$

In most applications, one is interested in the minimal initial condition Y_0 and in the associated control Z. This is for instance the case in the financial literature in which Y_0 represents the cost of the cheapest super-replication strategy for the contingent claim ξ, and Z provides the associated hedging strategy, see e.g. [9].

Motivated by situations where this minimal value Y_0 is too large for practical applications, it was suggested to relax the strong constraint (1.2) into a weaker one of the form

$$E\left[\ell(Y_T^Z - \xi)\right] \geq m \ , \tag{1.3}$$

where m is a given threshold and ℓ is a non-decreasing map. For $\ell(x) = \mathbf{1}_{\{x \geq 0\}}$, this corresponds to matching the criteria $Y_T^Z \geq \xi$ at least with probability m. In financial terms, this is the so-called quantile hedging problem, see Föllmer and Leukert [10]. More generally, ℓ is viewed as a loss function, one typical example being $\ell(x) := -(x^-)^q$ with $q \geq 1$, see Föllmer and Leukert [11] for general non-Markovian but linear dynamics. Such problems were coined "stochastic target problems with controlled loss" by Bouchard, Elie and Touzi [3] who consider a non-linear Markovian formulation in a Brownian diffusion setting, see also Moreau [12] for the jump diffusions setting.

The aim of this paper is to study the non-linear non-Markovian setting in which the terminal constraint is of the form

$$E\left[\Psi(Y_T^Z)\right] \geq m. \tag{1.4}$$

In the above, $m \in \mathbb{R}$ and Ψ is a (possibly random) non-decreasing real-valued map. Our problem can then be written as

$$\text{Find the minimal } Y_0 \text{ such that (1.1) and (1.4) hold for some } Z. \tag{1.5}$$

This leads to the introduction of a new class of BSDEs which we call BSDEs with weak terminal condition. More precisely, we refer to this problem by saying that we want to solve the BSDE with driver g and weak terminal condition (Ψ, m) to insist on the fact that the

terminal condition Y_T^Z is not fixed as a random variable, but only has to satisfy the weak constraint (1.4).

The first step in our analysis lies in a reformulation based on the martingale representation theorem, as suggested in [3]. More precisely, if Y_0 and Z are such that (1.4) holds, then the martingale representation Theorem implies that we can find an element α in the set $\mathbf{A}_0$, of predictable square integrable processes, such that

$$\Psi(Y_T^Z) \geq M_T^\alpha := m + \int_0^T \alpha_s dW_s.$$

On the other hand, since Ψ is non-decreasing, one can introduce its right inverse Φ and note that the solution (Y^α, Z^α) of the BSDE

$$Y_t^\alpha = \Phi(M_T^\alpha) + \int_t^T g(s, Y_s^\alpha, Z_s^\alpha)ds - \int_t^T Z_s^\alpha dW_s,\ 0 \leq t \leq T\ , \tag{1.6}$$

actually solves (1.1) and (1.4). We indeed show that the solution of (1.5) is given by

$$\inf\{Y_0^\alpha,\ \alpha \in \mathbf{A}_0\}. \tag{1.7}$$

This leads to study its dynamical counterpart

$$\mathcal{Y}_\tau^\alpha := \operatorname{essinf}\{Y_\tau^{\alpha'},\ \alpha' \in \mathbf{A}_0 \text{ s.t. } \alpha' = \alpha \text{ on } [\![0,\tau]\!]\}\ ,\quad 0 \leq \tau \leq T\ . \tag{1.8}$$

We verify that the family $\{\mathcal{Y}^\alpha,\ \alpha \in \mathbf{A}_0\}$ satisfies a dynamic programming principle which can be seen as a counterpart of the geometric dynamic programming principle of Soner and Touzi [20] used in [3]. In particular, this implies that $\{\mathcal{Y}^\alpha,\ \alpha \in \mathbf{A}_0\}$ is a g-submartingale family to which we can apply the non-linear Doob-Meyer decomposition of [16]. This provides a representation of the family $\{\mathcal{Y}^\alpha,\ \alpha \in \mathbf{A}_0\}$ in terms of minimal supersolutions to a family of BSDEs with driver g and (strong) terminal conditions $\{\Phi(M_T^\alpha),\ \alpha \in \mathbf{A}_0\}$. This representation allows in particular to characterize the family $\{\mathcal{Y}^\alpha,\ \alpha \in \mathbf{A}_0\}$ uniquely. Under additional convexity assumptions on the coefficients g and Φ, we observe that the essential infimum in (1.8) is attained. Hence, there exists an optimal $\hat{\alpha} \in \mathbf{A}_0$ such that solving the BSDE with weak terminal condition (Ψ, m) boils down to solving the BSDE with dynamics (1.6) and strong terminal condition $\Phi(M_T^{\hat{\alpha}})$. In a Markovian framework, our approach provides in particular a BSDE formulation for the PDEs derived in [3].

We then study in details important properties of this family and focus in particular on the regularity of $\mathcal{Y}^\alpha$ with respect to the threshold parameter m. We exhibit, under weak conditions, a stability property of the solution with respect to the variations of the parameter m. We also observe that $\mathcal{Y}^\alpha$ is convex with respect to the threshold parameter. This observation allows us in particular to conclude that Φ (whenever it is deterministic) can be replaced by its more regular convex envelope in order to compute $\mathcal{Y}^\alpha$ on $[0,T)$. This was already observed in the restrictive Markovian setting of [3], in which it is proved by using PDE technics. We provide here a pure probabilistic argument. Similarly, it was also observed in [10], [11] and [3] that (1.5) admits a dual linear problem when g is linear. We extend this result via probabilistic arguments to the semi-linear setting, for which the dual

formulation takes the form of a stochastic control problem in Meyer form.

The rest of the paper is organized as follows. In Section 2, we provide a precise formulation for (1.5) and relate this problem to a g-submartingale family satisfying a dynamic programming principle. Attainability of the optimal control $\hat{\alpha} \in \mathbf{A}_0$ is also discussed. Section 3 collects the continuity and convexity properties as well as the dual formulation of the problem. The connection with PDEs in a Markovian framework is given in Section 4. Finally, Section 5 contains the proof of the BSDE representation for $\{\mathcal{Y}^\alpha,\ \alpha \in \mathbf{A}_0\}$.

We close this introduction with a series of notations that will be used all over this paper. Let $d \geq 1$ and $T > 0$ be fixed. We denote by $W := (W_t)_{t\in[0,T]}$ a d-dimensional Brownian motion defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{P}$-augmented natural filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$. The components of W will be denoted by $W = (W^1, \cdots, W^d)$ and E will stand for the expectation with respect to $\mathbb{P}$. For simplicity, we assume that $\mathcal{F} = \mathcal{F}_T$. Throughout the paper we will make use of the following spaces.

- $\mathbf{L}_p(U, \mathcal{G})$ denotes the set of p-integrable $\mathcal{G}$-measurable random variables with values in U, $p \geq 0$, U a Borel set of $\mathbb{R}^n$ for some $n \geq 1$ and $\mathcal{G} \subset \mathcal{F}$. When U and $\mathcal{G}$ can be clearly identified by the context, we omit them. This will be in particular the case when $\mathcal{G} = \mathcal{F}$.

- $\mathcal{T}$ denotes the set of $\mathbb{F}$-stopping times in $[0, T]$. For $\tau_1 \in \mathcal{T}$, $\mathcal{T}_{\tau_1}$ is the set of stopping times τ_2 in $\mathcal{T}$ such that $\tau_2 \geq \tau_1$ $\mathbb{P}-$a.s. The notation $E_\tau[\cdot]$ stands for the conditional expectation given $\mathcal{F}_\tau$, $\tau \in \mathcal{T}$.

- $\mathbf{S}_2$ denotes the set of $\mathbb{R}$-valued, *càdlàg*[1] and $\mathbb{F}$-adapted stochastic processes $X = (X_t)_{t\in[0,T]}$ such that $\|X\|_{\mathbf{S}_2} := E[\sup_{t\in[0,T]} |X_t|^2]^{1/2} < \infty$.

- $\mathbf{H}_2$ denotes the set of $\mathbb{R}^n$-valued, $\mathbb{F}$-predictable stochastic processes $X = (X_t)_{t\in[0,T]}$ such that $\|X\|_{\mathbf{H}_2} := E\left[\int_0^T |X_t|^2 dt\right]^{1/2} < \infty$. In the following, the dimension n will be given by the context.

- $\mathbf{K}_2$ denotes the set of non-decreasing $\mathbb{R}$-valued and $\mathbb{F}$-adapted stochastic processes $X = (X_t)_{t\in[0,T]}$ such that $\|X\|_{\mathbf{S}_2} < \infty$.

Inequalities between random variables are understood in the $\mathbb{P}-$a.s.-sense.

2 BSDE with weak terminal condition

2.1 Definitions and problem reformulation

We first define the main object of this paper.

Definition 2.1 (Solution to a BSDE with weak terminal condition)**.** *Given a measurable map $\Psi : \mathbb{R} \times \Omega \mapsto U$, with $U \subset \mathbb{R} \cup \{-\infty\}$, $\tau \in \mathcal{T}$ and $\mu \in \mathbf{L}_0(U, \mathcal{F}_\tau)$, we say that*

[1] right-continuous with left limits

$(Y, Z) \in \mathbf{S}_2 \times \mathbf{H}_2$ *is a supersolution of the BSDE with generator* $g : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ *and weak terminal condition* (Ψ, μ, τ)*, in short* $\text{BSDE}(g, \Psi, \mu, \tau)$*, if*

$$Y \geq Y_T + \int_t^T g(s, Y_s, Z_s) - \int_t^T Z_s dW_s, \quad t \in [0, T], \qquad \textit{and} \tag{2.1}$$

$$E_\tau \left[\Psi(Y_T)\right] \geq \mu. \tag{2.2}$$

Before discussing the well-posedness of Equation (2.1)-(2.2), let us emphasize that the difference with classical BSDEs lies in the fact that we do not prescribe a terminal condition to Y in the classical $\mathbb{P}-$a.s.-sense but only impose a weak condition in expectation form (which justifies the terminology of *BSDE with weak terminal condition*). Even if we were asking for equalities in (2.1)-(2.2), this would obviously be too weak to expect uniqueness, as any random variable ξ satisfying $E_\tau \left[\Psi(\xi)\right] = \mu$ could serve as a terminal condition.

However, when Ψ is non-decreasing, the set

$$\Gamma(\tau, \mu) := \left\{Y_\tau : (Y, Z) \in \mathbf{S}_2 \times \mathbf{H}_2 \text{ is a supersolution of } \text{BSDE}(g, \Psi, \mu, \tau)\right\}, \tag{2.3}$$

defined for any $\tau \in \mathcal{T}$ and $\mu \in \mathbf{L}_0(U, \mathcal{F}_\tau)$, can be characterized by its lower-bound, whenever it is achieved.

Throughout the paper, we shall restrict to the case where g is Lipschitz continuous with linear growth, Ψ^+ is bounded, and the domain of Ψ is bounded from below, in order to avoid un-necessary technicalities.

Standing Assumption ($\mathbf{H}_\Psi$): For $\mathbb{P}-$a.e. $\omega \in \Omega$, the map $y \in \mathbb{R} \mapsto \Psi(\omega, y)$ is non-decreasing and valued in $[0, 1] \cup \{-\infty\}$, its right-inverse $\Phi(\omega, \cdot)$ is such that $\Phi : \Omega \times [0, 1] \mapsto [0, 1]$ is measurable.

The above assumption means that $\Psi(\omega, \cdot) \in [0, 1]$ on $[0, \infty)$ and $\Psi(\omega, \cdot) = -\infty$ on $(-\infty, 0)$. In particular, the constraint in expectation (2.2) implies $Y_T \geq 0$ $\mathbb{P}-$a.s. Obviously the set $[0, 1]$ is chosen for ease of notations and can be replaced by any closed interval. By right-inverse we mean $\Phi(x) := \inf\{y \in \mathbb{R},\ \Psi(y) \geq x\}$. Immediate computations show that

$$\Phi \circ \Psi \leq \text{Id}. \tag{2.4}$$

Standing Assumption ($\mathbf{H}_g$) g is a measurable map from $\Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d$ to $\mathbb{R}$ and $g(\cdot, y, z)$ is $\mathbb{F}$-predictable, for each $(y, z) \in \mathbb{R} \times \mathbb{R}^d$. There exists a constant $K_g > 0$ and a random variable $\chi_g \in \mathbf{L}_2(\mathbb{R}_+)$, such that

$$\begin{gathered} |g(t, 0, 0)| \leq \chi_g \ \ \mathbb{P}-\text{a.s.} \\ |g(t, y_1, z_1) - g(t, y_2, z_2)| \leq K_g(|y_1 - y_2| + |z_1 - z_2|) \ \ \mathbb{P}-\text{a.s.} \\ \forall (t, y_i, z_i) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d,\ i = 1, 2. \end{gathered}$$

Let $\mathbf{A}_{\tau,\mu}$ denote the set elements $\alpha \in \mathbf{H}_2$ such that

$$M^{(\tau,\mu),\alpha} := \mu + \int_\tau^{\tau \vee \cdot} \alpha_s dW_s \ \text{ takes values in } [0, 1]. \tag{2.5}$$

Then, (2.2) is equivalent to $\Psi(Y_T) \geq M_T^{(\tau,\mu),\alpha}$ for some $\alpha \in \mathbf{A}_{\tau,\mu}$. In view of (2.4), this is equivalent to $Y_T \geq \Phi(M_T^{(\tau,\mu),\alpha})$ for some $\alpha \in \mathbf{A}_{\tau,\mu}$. This implies that supersolutions of BSDE(g, Ψ, μ, τ) can be characterized in terms of g-expectations whose definition is recalled below.

Definition 2.2 (g-expectation)**.** *Given $\tau_2 \in \mathcal{T}$ and $\xi \in \mathbf{L}_2(\mathbb{R}, \mathcal{F}_{\tau_2})$, let $(Y, Z) \in \mathbf{S}_2 \times \mathbf{H}_2$ denote the solution of*

$$Y = \xi + \int_{\cdot\wedge\tau_2}^{\tau_2} g(s, Y_s, Z_s) - \int_{\cdot\wedge\tau_2}^{\tau_2} Z_s dW_s.$$

Then, we define the (conditional) g-expectation of ξ at the stopping time $\tau_1 \leq \tau_2$ as $\mathcal{E}^g_{\tau_1,\tau_2}[\xi] := Y_{\tau_1}$. When $\tau_2 \equiv T$, we only write $\mathcal{E}^g_{\tau_1}[\xi]$, and say that (Y, Z) solves BSDE(g, ξ).

Note that existence and uniqueness hold under Assumption **(H$_g$)**. In the following, we shall adopt the terminology of Peng [17] and call g-martingale (resp. g-submartingale) a process Y such that $\mathcal{E}^g_{t,s}[Y_s] = Y_t$ (resp. $\mathcal{E}^g_{t,s}[Y_s] \geq Y_t$), for all $t \leq s \leq T$.

Proposition 2.1. *Fix $\tau \in \mathcal{T}$, $\mu \in \mathbf{L}_0([0,1], \mathcal{F}_\tau)$. Then, $(Y, Z) \in \mathbf{S}_2 \times \mathbf{H}_2$ is a supersolution of* BSDE(g, Ψ, μ, τ) *if and only if (Y, Z) satisfies* (2.1) *and there exists $\alpha \in \mathbf{A}_{\tau,\mu}$ such that $Y_t \geq \mathcal{E}^g_t[\Phi(M_T^{(\tau,\mu),\alpha})]$ for $t \in [0, T]$ $\mathbb{P}-$ a.s.*

Proof. Let (Y, Z) be a super solution of BSDE(g, Ψ, μ, τ). Then there exists some element ρ in $\mathbf{L}_0([0,1], \mathcal{F}_\tau)$ with $\rho \geq \mu$, $\mathbb{P}-$a.s. and $\tilde{\alpha}$ in $\mathbf{A}_{\tau,\rho}$ such that $\Psi(Y_T) = M_T^{(\tau,\rho),\tilde{\alpha}}$. Set $\theta^{\tilde{\alpha}} := \inf\{s \geq \tau,\ M_s^{(\tau,\mu),\tilde{\alpha}} = 0\}$. It is clear that $\theta^{\tilde{\alpha}}$ belongs to $\mathcal{T}$ and that $\alpha := \tilde{\alpha}\mathbf{1}_{[0,\theta^{\tilde{\alpha}})}$ belongs to $\mathbf{A}_{\tau,\mu}$ and satisfies $M_T^{(\tau,\rho),\tilde{\alpha}} \geq M_T^{(\tau,\mu),\alpha}$, $\mathbb{P}-$a.s., since $M_T^{(\tau,\rho),\tilde{\alpha}} \geq 0$ by definition of $\mathbf{A}_{\tau,\rho}$. The monotonicity of Φ and Relation (2.4) imply that

$$Y_T \geq (\Phi \circ \Psi)(Y_T) \geq \Phi(M_T^{(\tau,\mu),\alpha}).$$

By comparison for Lipschitz BSDEs, we obtain $Y_t \geq \mathcal{E}^g_t[\Phi(M_T^{(\tau,\mu),\alpha})]$ for $t \in [0, T]$. Conversely, let $\alpha \in \mathbf{A}_{\tau,\mu}$ be such that $Y_t \geq \mathcal{E}^g_t[\Phi(M_T^{(\tau,\mu),\alpha})]$ for $t \in [0, T]$ and assume that (Y, Z) satisfies (2.1). Then, by definition of Φ, it holds that

$$\Psi(Y_T) \geq (\Psi \circ \Phi)(M_T^{(\tau,\mu),\alpha}) \geq M_T^{(\tau,\mu),\alpha}.$$

Taking the conditional expectation on both sides leads to (2.2). □

In view of Proposition 2.1, the lower bound of $\Gamma(\tau, \mu)$ (which we recall, has been defined in (2.3)) can be expressed in terms of

$$\mathcal{Y}_\tau(\mu) := \operatorname{ess\,inf}_{\alpha \in \mathbf{A}_{\tau,\mu}} \mathcal{E}^g_\tau\left[\Phi(M_T^{(\tau,\mu),\alpha})\right],\ \tau \in \mathcal{T},\ \mu \in \mathbf{L}_0([0,1], \mathcal{F}_\tau). \tag{2.6}$$

This is the statement of the next proposition.

Proposition 2.2. $\operatorname{essinf} \Gamma(\tau, \mu) = \mathcal{Y}_\tau(\mu)$, *for all $\tau \in \mathcal{T}$ and $\mu \in \mathbf{L}_0([0,1], \mathcal{F}_\tau)$.*

Proof. The fact that $Y_\tau \in \Gamma(\tau, \mu)$ implies $Y_\tau \geq \mathcal{Y}_\tau(\mu)$ follows from Proposition 2.1. On the other hand, the same proposition implies that each $\mathcal{E}^g_\tau[\Phi(M_T^{(\tau,\mu),\alpha})]$ with $\alpha \in \mathbf{A}_{\tau,\mu}$ belongs to $\Gamma(\tau, \mu)$. □

Remark 2.1. For later use, note that the assumptions **($\mathbf{H}_g$)** and **($\mathbf{H}_\Psi$)** ensure that we can find $\eta \in \mathbf{S}_2$ such that $|\mathcal{E}^g_t(\Phi(M))| \vee |\mathcal{Y}_t(\mu)| \leq \eta_t$, for all $t \leq T$ and $\mu \in \mathbf{L}_0([0,1], \mathcal{F}_t)$, $M \in \mathbf{L}_0([0,1])$. See (i) of Proposition 6.2 in the Appendix.

Remark 2.2. Note that $\mathcal{Y}_\tau(\mu) = \mathcal{Y}_\tau(\mu_1)\mathbf{1}_A + \mathcal{Y}_\tau(\mu_2)\mathbf{1}_{A^c}$ whenever $\mu := \mu_1\mathbf{1}_A + \mu_2\mathbf{1}_{A^c}$ for $A \in \mathcal{F}_\tau$, $\mu_1, \mu_2 \in \mathbf{L}_0([0,1], \mathcal{F}_\tau)$, and $\tau \in \mathcal{T}$. Indeed, $\alpha := \mathbf{1}_{[\tau,T]}(\alpha_1\mathbf{1}_A + \alpha_2\mathbf{1}_{A^c}) \in \mathbf{A}_{\tau,\mu}$ for all $\alpha_i \in \mathbf{A}_{\tau,\mu_i}$ with $i = 1, 2$. Since $\mathcal{E}^g_\tau\left[\Phi(M_T^{(\tau,\mu),\alpha})\right] = \mathcal{E}^g_\tau\left[\Phi(M_T^{(\tau,\mu_1),\alpha_1})\right]\mathbf{1}_A + \mathcal{E}^g_\tau\left[\Phi(M_T^{(\tau,\mu_2),\alpha_2})\right]\mathbf{1}_{A^c}$, this implies $\mathcal{Y}_\tau(\mu) \leq \mathcal{Y}_\tau(\mu_1)\mathbf{1}_A + \mathcal{Y}_\tau(\mu_2)\mathbf{1}_{A^c}$. The converse inequality follows from the previous identity applied with $\alpha_1 := \alpha\mathbf{1}_A$ and $\alpha_2 := \alpha\mathbf{1}_{A^c}$ for any $\alpha \in \mathbf{A}_{\tau,\mu}$ so that $\alpha_i \in \mathbf{A}_{\tau,\mu_i}$ for $i = 1, 2$.

Remark 2.3. Before going on with the study of the set Γ, let us notice that a similar analysis can be carried out for weak constraints of the form $\mathcal{E}^h_\tau[\Psi(Y_T)] \geq \mu$ in place of $E_\tau[\Psi(Y_T)] \geq \mu$ in (2.2), with $\mathcal{E}^h$ defined as the h-expectation associated to some random map h satisfying similar conditions as g. In finance, the latter condition interprets as a risk-measure constraints, see e.g. [17], while our condition is more related to expected loss constraints, see [11]. Again, we try to avoid un-necessary additional technicalities and stick to the case $h \equiv 0$.

2.2 BSDE characterization of the minimal initial condition

The main result of this section is a BSDE characterization for the lower bound of the set $\Gamma(\tau, \mu)$ of time-τ initial conditions of supersolutions of BSDE(g, Ψ, μ, τ). In particular, this extends to a non Markovian framework the PDE characterization of [3].

For ease of notations, we now fix $m_o \in [0,1]$ and set

$$\begin{cases} M^\alpha := M^{(0,m_o),\alpha}\ ,\ \mathbf{A}^\alpha_\tau := \{\alpha' \in \mathbf{A}_{\tau,M^\alpha_\tau} :\ \alpha' = \alpha\ dt \times d\mathbb{P} \text{ on } [\![0,\tau]\!]\}, \\ \mathbf{A}_0 := \mathbf{A}_{0,m_o} \text{ and } \mathcal{Y}^\alpha_\cdot := \mathcal{Y}_\cdot(M^\alpha_\cdot) \text{ for } \alpha \in \mathbf{A}_0, \end{cases}$$

where we recall that $M^{(0,m_o),\alpha}$ and $\mathbf{A}_{0,m_o}$ are given in (2.5).

Theorem 2.1. *For any $\alpha \in \mathbf{A}_0$, $\mathcal{Y}^\alpha$ is indistinguishable from a làdlàg g-submartingale, and the following dynamic programming principle holds:*

(i) $\mathcal{Y}^\alpha_{\tau_1} = \operatorname{ess\,inf}_{\bar\alpha \in \mathbf{A}^\alpha_{\tau_1}} \mathcal{E}^g_{\tau_1,\tau_2}[\mathcal{Y}^{\bar\alpha}_{\tau_2}]$, *for each $\tau_1 \in \mathcal{T}$, $\tau_2 \in \mathcal{T}_{\tau_1}$.*

Under the additional assumption that

$$m \in [0,1] \mapsto \Phi(\omega, m) \text{ is continuous for } \mathbb{P}\text{-a.e. } \omega \in \Omega, \tag{2.7}$$

the following holds:

(ii) $\mathcal{Y}^\alpha$ *is indistinguishable from a càdlàg g-submartingale, for each $\alpha \in \mathbf{A}_0$.*

(iii) There exists a family $(\mathcal{Z}^\alpha, \mathcal{K}^\alpha)_{\alpha \in \mathbf{A}_0} \subset \mathbf{H}_2 \times \mathbf{K}_2$ satisfying

$$\sup_{\alpha \in \mathbf{A}_0} \|(\mathcal{Y}^\alpha, \mathcal{Z}^\alpha, \mathcal{K}^\alpha)\|_{\mathbf{S}_2 \times \mathbf{H}_2 \times \mathbf{K}_2} \ <\ \infty\,, \tag{2.8}$$

and such that, for all $\alpha \in \mathbf{A}_0$*, we have*

$$\mathcal{Y}^\alpha = \Phi(M_T^\alpha) + \int_\cdot^T g(s, \mathcal{Y}_s^\alpha, \mathcal{Z}_s^\alpha) ds - \int_\cdot^T \mathcal{Z}_s^\alpha dW_s + \mathcal{K}^\alpha - \mathcal{K}_T^\alpha \ \text{ on } \ [0,T], \quad (2.9)$$

$$\mathcal{K}_{\tau_1}^\alpha = \operatorname*{ess\,inf}_{\bar\alpha \in \mathbf{A}_{\tau_1}^\alpha} E\left[\mathcal{K}_{\tau_2}^{\bar\alpha} | \mathcal{F}_{\tau_1}\right] \ , \ \forall\, \tau_1 \in \mathcal{T},\ \tau_2 \in \mathcal{T}_{\tau_1}, \quad (2.10)$$

and

$$(\mathcal{Y}^\alpha, \mathcal{Z}^\alpha, \mathcal{K}^\alpha)\mathbf{1}_{[0,\tau]} = (\mathcal{Y}^{\bar\alpha}, \mathcal{Z}^{\bar\alpha}, \mathcal{K}^{\bar\alpha})\mathbf{1}_{[0,\tau]}, \ \ \forall\, \tau \in \mathcal{T},\ \bar\alpha \in \mathbf{A}_\tau^\alpha. \quad (2.11)$$

(iv) $(\mathcal{Y}^\alpha, \mathcal{Z}^\alpha, \mathcal{K}^\alpha)_{\alpha \in \mathbf{A}_0}$ *is the unique family of* $\mathbf{S}_2 \times \mathbf{H}_2 \times \mathbf{K}_2$ *satisfying* (2.8)-(2.9)-(2.10)-(2.11) *for all* $\alpha \in \mathbf{A}_0$.

The proof of this theorem is reported in Section 5.

Remark 2.3. (i) The precise continuity assumption needed in the proof is : $\Phi(M_T^{\alpha_n})$ converges in $\mathbf{L}_2$ to $\Phi(M_T^\alpha)$ whenever $\|M_T^{\alpha_n} - M_T^\alpha\|_{\mathbf{L}_2}$ tends to 0, for any sequence $(\alpha_n)_n \subset \mathbf{A}_0$. However, this condition implies that Φ is continuous, as soon as random variables with non-absolutely continuous law with respect to the Lebesgue measure might be considered (which is the case here).

(ii) We shall see in Proposition 3.3 below that Φ can be replaced by its m-convex envelope, under mild assumptions. In this case, the continuity assumption of the second part of Theorem 2.1 can be replaced by a similar continuity assumption on the convex envelope of Φ, which is clearly a much weaker condition as it only concerns the right boundary point 1 (since Φ is non-decreasing). A typical example is given in Remark 3.4 below.

2.3 Representation as a BSDE with strong terminal condition

The previous section raises in particular one natural question: Does there exist an admissible control $\hat\alpha$ on the whole time interval $[0,T]$ allowing to match all time t-values of the minimal solution of a BSDE with weak terminal condition? Rephrasing, we wonder about the existence of a control $\hat\alpha$ in $\mathbf{A}_0$ such that

$$\mathcal{Y}_t^{\hat\alpha} = \mathcal{E}_t^g\left[\Phi(M_T^{\hat\alpha})\right] \ , \qquad 0 \le t \le T \ .$$

Hereby, solving the BSDE with weak terminal condition $(\Psi, m_o, 0)$ boils down to solving the classical BSDE with the optimal strong terminal one $\Phi(M_T^{\hat\alpha})$: along the optimal path $\hat\alpha$, the compensator $\mathcal{K}^{\hat\alpha}$ of the BSDE (2.9) must degenerate to 0.

Not surprisingly, the existence of an optimal control requires the addition of convexity assumptions on the coefficients of the BSDE. We shall therefore assume that:

($\mathbf{H}_{conv}$) For all $(\lambda, m_1, m_2, t, y_1, y_2, z_1, z_2) \in [0,1] \times [0,1]^2 \times [0,T] \times \mathbb{R}^2 \times [\mathbb{R}^d]^2$:

$$\begin{aligned} \Phi(\lambda m_1 + (1-\lambda)m_2) &\le \lambda\Phi(m_1) + (1-\lambda)\Phi(m_2) \ \mathbb{P}-\text{a.s.} \\ g(t, \lambda y_1 + (1-\lambda)y_2, \lambda z_1 + (1-\lambda)z_2) &\le \lambda g(t, y_1, z_1) + (1-\lambda) g(t, y_2, z_2) \ \mathbb{P}-\text{a.s.} \end{aligned}$$

Remark 2.4. We recall the following result which is based on standard comparison arguments, see e.g. [19, Proposition 7]: For any $\tau \in \mathcal{T}$, the map $\mathcal{E}^g_\tau[\Phi(\cdot)] : \mathbf{L}_0([0,1]) \to \mathbf{L}_0$ is convex under Assumption $(\mathbf{H}_{conv})$.

Proposition 2.3. *Assume that Assumptions* $(\mathbf{H}_{conv})$ *and* (2.7) *hold. Then, for any* $(\tau, \alpha) \in \mathcal{T} \times \mathbf{H}_2$*, there exists* $\hat{\alpha}^{\tau,\alpha} \in \mathbf{A}^\alpha_\tau$ *such that*

$$\mathcal{Y}^\alpha_\tau = \mathcal{E}^g_\tau \left[\Phi(M_T^{\hat{\alpha}^{\tau,\alpha}})\right] = \mathcal{E}^g_{\tau,\tau'} \left[\mathcal{Y}^{\hat{\alpha}^{\tau,\alpha}}_{\tau'}\right], \ \forall\, \tau' \in \mathcal{T}_\tau.$$

Remark 2.5. As detailed in Remark 3.3 below, the convexity assumption on the terminal map Φ can be avoided in some cases. In particular, if Φ is deterministic then it can be replaced by its convex envelope. Then, only the convexity assumption on g has to hold.

Proof. Lemma 5.1 below provides a sequence $(\alpha^n)_n$ valued in $\mathbf{A}^\alpha_\tau$ such that

$$\mathcal{Y}^\alpha_\tau = \lim_{n\to\infty} \downarrow \mathcal{E}^g_\tau \left[\Phi(M_T^{\alpha^n})\right], \ \mathbb{P} - \text{a.s.} \tag{2.12}$$

Since the sequence $(M_T^{\alpha^n})_n$ is bounded in $[0,1]$, we can find sequences of non-negative real numbers $(\lambda^n_i)_{i\geq n}$ with $\sum_{i\geq n} \lambda^n_i = 1$, such that only a finite number of λ^n_i do not vanish, for each n, and such that the sequence of convex combinations $(\tilde{M}^n_T)_n$ given by

$$\tilde{M}^n_T \ := \ \sum_{i\geq n} \lambda^n_i M_T^{\alpha^i} \tag{2.13}$$

converges $\mathbb{P} - $ a.s. to some $\hat{M}_T \in \mathbf{L}_0([0,1])$. By dominated convergence, the convergence holds in $\mathbf{L}_2$, in particular $E_\tau[\hat{M}_T] = M^\alpha_\tau$, and the martingale representation Theorem implies that we can find $\hat{\alpha} \in \mathbf{A}^\alpha_\tau$ such that $\hat{M}_T = M^{\hat{\alpha}}_T$. Using the convexity of Φ and g, see Remark 2.4, we deduce that

$$\tilde{Y}^n_\tau := \sum_{i\geq n} \lambda^n_i \mathcal{E}^g_\tau \left[\Phi(M_T^{\alpha^i})\right] \geq \mathcal{E}^g_\tau \left[\Phi(\tilde{M}^n_T)\right].$$

By (2.12), $\tilde{Y}^n_\tau \to \mathcal{Y}^\alpha_\tau$ $\mathbb{P}-$a.s. On the other hand, the convergence $\tilde{M}^n_T \to M^{\hat{\alpha}}_T$ in $\mathbf{L}_2$ combined with the boundedness and a.s. continuity of Φ implies that $\Phi(\tilde{M}^n_T) \to \Phi(M^{\hat{\alpha}}_T)$ in $\mathbf{L}_2$, after possibly passing to a subsequence. Therefore the convergence $\mathcal{E}^g_\tau \left[\Phi(\tilde{M}^n_T)\right] \to \mathcal{E}^g_\tau \left[\Phi(M^{\hat{\alpha}}_T)\right]$ $\mathbb{P} - $ a.s. follows by Proposition 6.1 below. This gives $\mathcal{Y}^\alpha_\tau \geq \mathcal{E}^g_\tau \left[\Phi(M^{\hat{\alpha}}_T)\right]$, while the converse holds by definition of $\mathcal{Y}^\alpha_\tau$.

It remains to show that $\mathcal{Y}^\alpha_\tau = \mathcal{E}^g_{\tau,\tau'} \left[\mathcal{Y}^{\hat{\alpha}}_{\tau'}\right]$, for $\tau' \in \mathcal{T}_\tau$. To see this, first note that the above implies that $\mathcal{Y}^\alpha_\tau = \mathcal{E}^g_{\tau,\tau'} \left[\mathcal{E}^g_{\tau'}[\Phi(M^{\hat{\alpha}}_T)]\right] \geq \mathcal{E}^g_{\tau,\tau'} \left[\mathcal{Y}^{\hat{\alpha}}_{\tau'}\right]$ by standard comparison arguments and the fact that $\mathcal{E}^g_{\tau'}[\Phi(M^{\hat{\alpha}}_T)] \geq \mathcal{Y}^{\hat{\alpha}}_{\tau'}$ by definition. As above, we can find a sequence $(\hat{\alpha}^n) \in \mathbf{A}^{\hat{\alpha}}_{\tau'}$ such that $\mathcal{E}^g_{\tau'} \left[\Phi(M_T^{\hat{\alpha}^n})\right] \to \mathcal{Y}^{\hat{\alpha}}_{\tau'}$ $\mathbb{P} - $ a.s. In view of Remark 2.1, the convergence holds in $\mathbf{L}_2$ and Proposition 6.1 below implies

$$\mathcal{Y}^\alpha_\tau \leq \mathcal{E}^g_{\tau,\tau'} \left[\mathcal{E}^g_{\tau'} \left[\Phi(M_T^{\hat{\alpha}^n})\right]\right] \to \mathcal{E}^g_{\tau,\tau'} \left[\mathcal{Y}^{\hat{\alpha}}_{\tau'}\right],$$

where we used the fact that $\hat{\alpha}^n \in \mathbf{A}^{\hat{\alpha}}_{\tau'} \subset \mathbf{A}^\alpha_\tau$ to obtain the left hand-side. □

2.4 Link with second order BSDEs

Note that the formulation of Theorem 2.1(iii), is very close to the one obtained in [21] for 2BSDEs. The connection with 2BSDEs can actually be established more precisely, at least at an informal level, in the case where the driver g and Φ are deterministic maps, and g does not depend on its Z-component.

Given $\alpha \in \mathbf{A}_0$, let (Y^α, Z^α) denote the solution of the classical BSDE$(g, \Phi(M_T^\alpha))$. If α belongs to the subset $\mathbf{A}_{0,>0}$ of elements of $\mathbf{A}_0$ that have $dt \times d\mathbb{P}$-a.e. non-zero components, we can define $\tilde{Y}^\alpha = -Y^\alpha$ and $\tilde{Z}^{\alpha,i} := -Z^{\alpha,i}/\alpha^i$, $i = 1, \dots, d$. Let us also set $\Xi(b) := -\Phi(m_o + \sum_{i=1}^d b^i)$, $f(\cdot,\cdot) := -g(\cdot,-\cdot)$, and $B^{\alpha,i} := \int_0^\cdot \alpha_s^i dW_s^i$. Then,

$$\tilde{Y}^\alpha \;=\; \Xi(B_T^\alpha) + \int_\cdot^T f(s, \tilde{Y}_s^\alpha)ds - \int_\cdot^T \tilde{Z}_s^\alpha dB_s^\alpha \;\text{ on } [0,T]\,,\ \mathbb{P}-\text{a.s.}$$

Let $\mathbb{P}_o$ denote the Wiener measure and $\mathbb{P}^\alpha = \mathbb{P}_o \circ (B^\alpha)^{-1}$ denote the pull-back measure associated to B^α on the canonical space. Then, the canonical process B has the same law under $\mathbb{P}^\alpha$ than B^α under $\mathbb{P}_o$. This implies, at an informal level, that $\tilde{Y}^\alpha$ has the same law under $\mathbb{P}_o$ than $\bar{Y}^{\mathbb{P}^\alpha}$ under $\mathbb{P}^\alpha$, where $(\bar{Y}^{\mathbb{P}^\alpha}, \bar{Z}^{\mathbb{P}^\alpha})$ denotes the solution of the BSDE

$$\bar{Y}^{\mathbb{P}^\alpha} \;=\; \Xi(B_T) + \int_\cdot^T f(s, \bar{Y}_s^{\mathbb{P}^\alpha})ds - \int_\cdot^T \bar{Z}_s^{\mathbb{P}^\alpha} dB_s \;\text{ on } [0,T]\,,\ \mathbb{P}^\alpha-\text{a.s.}$$

An informal density argument then leads to

$$-\mathcal{Y}_0 = \operatorname{ess\,sup}_{\alpha\in\mathbf{A}_0} \tilde{Y}_0^\alpha = \operatorname{ess\,sup}_{\alpha\in\mathbf{A}_{0,>0}} \tilde{Y}_0^\alpha = \operatorname{ess\,sup}_{\alpha\in\mathbf{A}_{0,>0}} \bar{Y}_0^{\mathbb{P}^\alpha}.$$

In view of [21, Theorem 4.3], this corresponds to the time 0 value of the Y-component of the 2BSDE with driver f and terminal condition $\Xi(B_T)$, for the family of probability measures $\{\mathbb{P}^\alpha, \alpha \in \mathbf{A}_{0,>0}\}$. The corresponding time t values could be similarly related. Note that in our setting $(\mathcal{Y}_s^\alpha, \mathcal{Z}_s^\alpha)_{s\le t}$ only depends on the path of the control α up to time t, see (2.11). In 2BSDEs, this dependency is incorporated in the dependency of the solution on the path of the canonical process: the solution is progressively measurable with respect to the right limit of the raw filtration. The dependencies are therefore similar. The difference lies in the fact that the solution of a 2BSDE, when it exists, is defined at the same time for all measures $\mathbb{P}^\alpha$. This requires a non-trivial aggregation procedure which does not appear in our setting.

Since this connection is, at least for the moment, more of rhetorical nature, we will not elaborate further on it in this paper.

3 Main properties of the minimal initial condition process

In this section, we emphasize remarkable properties of the map $\mathcal{Y}_t : \mu \in \mathbf{L}_0([0,1], \mathcal{F}_t) \mapsto \mathcal{Y}_t(\mu)$, for $t \in [0,T)$. We first derive the continuity of this map under a weak continuity assumption on $\mathcal{E}^g[\Phi(\cdot)]$. Then, we verify that this map (or more precisely its l.s.c. envelope) is convex, and discuss the propagation of the convexity property to the time boundary $T-$. Finally, we retrieve, in this non-Markovian setting, a dual representation of the map $\mathcal{Y}_0$, using solely probabilistic arguments.

3.1 Continuity

Our continuity result is stated in terms of the quantities

$$Err_t(\eta) := \operatorname{esssup}\left\{\mathcal{R}_t(M,M') \ : \ M,M' \in \mathbf{L}_0([0,1]) \ , \ \ E_t[|M-M'|^2] \le \eta\right\},$$

defined for $\eta \in \mathbf{L}_0([0,1])$, in which

$$\mathcal{R}_t(M,M') := |\mathcal{E}_t^g[\Phi(M)] - \mathcal{E}_t^g[\Phi(M')]|.$$

Observe that classical a priori estimates on BSDEs ensure that $Err_t(\eta_n) \to 0$ as $\eta_n \to 0$ $\mathbb{P}-$a.s. with $(\eta_n)_n \subset \mathbf{L}_0([0,1])$, whenever Φ is a deterministic Lipschitz map, see e.g. Proposition 6.1 below. This observation remains valid when Φ is simply continuous, via a classical convolution density argument for Lipschitz maps on bounded domains. The next result indicates that this property ensures regularity on the map: $\mu \mapsto \mathcal{Y}_t(\mu)$.

Proposition 3.1. *Let $t<T$, $\mu_1,\mu_2 \in \mathbf{L}_0([0,1],\mathcal{F}_t)$. Then,*

$$|\mathcal{Y}_t(\mu_1) - \mathcal{Y}_t(\mu_2)| \le Err_t(\Delta(\mu_1,\mu_2)) + Err_t(\Delta(\mu_2,\mu_1)),$$

where

$$\Delta(\mu_i,\mu_j) := (1-\frac{\mu_i}{\mu_j})\mathbf{1}_{\{\mu_i<\mu_j\}} + \frac{\mu_i-\mu_j}{1-\mu_j}\mathbf{1}_{\{\mu_i>\mu_j\}}, \ i,j=1,2.$$

Moreover,

$$|\mathcal{Y}_t(\mu_1) - \mathcal{Y}_t(\mu_2)|\mathbf{1}_{\{\mu_1=0\}} \le \mathcal{R}_t(\mu_2,0)$$

and

$$|\mathcal{Y}_t(\mu_1) - \mathcal{Y}_t(\mu_2)|\mathbf{1}_{\{\mu_1=1\}} \le \operatorname{esssup}\left\{\mathcal{R}_t(1,M) \ : \ M \in \mathbf{L}_0([0,1]) \ , \ \ E_t[|1-M|^2] \le 1-\mu_2\right\},$$

In particular, if $Err_t(\eta_n) \to 0$ $\mathbb{P}-a.s.$ as $\eta_n \to 0$ $\mathbb{P}-a.s.$, for all $(\eta_n)_n \subset \mathbf{L}_0([0,1])$, then $\mu \in \mathbf{L}_0((0,1),\mathcal{F}_t) \mapsto \mathcal{Y}_t(\mu)$ is continuous for the sequential $\mathbb{P}-a.s.$ convergence and the strong $\mathbf{L}_2$ convergence.

Proof. Step 1. Fix $\mu_1,\mu_2 \in \mathbf{L}_0([0,1],\mathcal{F}_t)$. Given $\alpha_2 \in \mathbf{A}_{t,\mu_2}$, we define

$$\lambda \ := \ \frac{1-\mu_1}{1-\mu_2}\mathbf{1}_{\{\mu_2<\mu_1\}} + \frac{\mu_1}{\mu_2}\mathbf{1}_{\{\mu_1<\mu_2\}} + \mathbf{1}_{\{\mu_1=\mu_2\}} \ ,$$

which is by construction valued in $[0,1]$. Since $M^{(t,\mu_2),\alpha_2}$ takes values in $[0,1]$,

$$M^{(t,\mu_1),\lambda\alpha_2} = \mu_1 - \lambda\mu_2 + \lambda M^{(t,\mu_2),\alpha_2} \ \in \ [\mu_1-\lambda\mu_2, \mu_1+\lambda(1-\mu_2)] \ \subset \ [0,1] \ .$$

In particular, $\lambda\alpha_2 \in \mathbf{A}_{t,\mu_1}$. Thus, (2.6) leads to

$$\mathcal{Y}_t(\mu_1) \ \le \ \mathcal{E}_t^g[\Phi(M_T^{(t,\mu_2),\alpha_2})] + (\mathcal{E}_t^g[\Phi(M_T^{(t,\mu_1),\lambda\alpha_2})] - \mathcal{E}_t^g[\Phi(M_T^{(t,\mu_2),\alpha_2})]) \ . \tag{3.1}$$

Besides,

$$M_T^{(t,\mu_1),\lambda\alpha_2} - M_T^{(t,\mu_2),\alpha_2} = \mu_1 - \lambda\mu_2 + (\lambda-1)M_T^{(t,\mu_2),\alpha_2}$$

so that, since $M_T^{(t,\mu_2),\alpha_2}$ belongs to $[0,1]$, we have

$$\mu_1 - 1 + \lambda(1-\mu_2) \le M_T^{(t,\mu_1),\lambda\alpha_2} - M_T^{(t,\mu_2),\alpha_2} \le \mu_1 - \lambda\mu_2.$$

In addition,

$$\begin{aligned} \mu_1 - \lambda\mu_2 &= 0\,, \qquad \text{if } \mu_1 < \mu_2\,, \text{ and} \\ \mu_1 - 1 + \lambda(1-\mu_2) &= 0\,, \qquad \text{if } \mu_1 \ge \mu_2\,. \end{aligned}$$

This directly leads to

$$E_t[|M_T^{(t,\mu_1),\lambda\alpha_2} - M_T^{(t,\mu_2),\alpha_2}|] \;\le\; \Delta(\mu_1,\mu_2)\,.$$

Since these two processes belong to $[0,1]$, we get

$$E_t[|M_T^{(t,\mu_1),\lambda\alpha_2} - M_T^{(t,\mu_2),\alpha_2}|^2] \le \Delta(\mu_1,\mu_2).$$

Hence, the arbitrariness of $\alpha_2 \in \mathbf{A}_{t,\mu_2}$ together with (2.6) and (3.1) provides

$$\mathcal{Y}_t(\mu_1) \;\le\; \mathcal{Y}_t(\mu_2) + Err_t(\Delta(\mu_1,\mu_2))\,.$$

Interchanging the roles of μ_1 and μ_2 leads to

$$\mathcal{Y}_t(\mu_2) \;\le\; \mathcal{Y}_t(\mu_1) + Err_t(\Delta(\mu_2,\mu_1))\,.$$

Step 2. We next consider the case where $\mathbb{P}[\mu_1 = 0] > 0$. Without loss of generality, we can assume that $\mu_1 \equiv 0$. Fix $\alpha \in \mathbf{A}_{t,\mu_2}$. Since $\mathbf{A}_{t,\mu_1} = \{0\}$, $M_T^{(t,\mu_2),\alpha} \ge 0$ and Φ is non-decreasing, comparison implies that

$$\mathcal{Y}_t(0) = \mathcal{E}_t^g[\Phi(0)] \le \mathcal{E}_t^g[\Phi(M_T^{(t,\mu_2),\alpha})].$$

In particular, $\mathcal{Y}_t(0) = \mathcal{E}_t^g[\Phi(0)] \le \mathcal{Y}_t(\mu_2) \le \mathcal{E}_t^g[\Phi(M_T^{(t,\mu_2),0})] = \mathcal{E}_t^g(\Phi(\mu_2))$.

Step 3. We now consider the case where $\mathbb{P}[\mu_1 = 1] > 0$. Again, we can assume that $\mu_1 \equiv 1$ so that $\mathbf{A}_{t,\mu_1} = \{0\}$. By comparison as above, one has

$$\mathcal{Y}_t(1) = \mathcal{E}_t^g[\Phi(1)] \ge \mathcal{Y}_t(\mu_2).$$

On the other hand, since $M^{(t,\mu_2),\alpha}$ is a martingale taking values in $[0,1]$, we have

$$E_t[|1 - M_T^{(t,\mu_2),\alpha}|^2] \le E_t[1 - M_T^{(t,\mu_2),\alpha}] = 1 - \mu_2, \quad \alpha \in \mathbf{A}_{t,\mu_2},$$

from which the result follows. □

3.2 Convexity

In [3] and [12], it is shown that the map $m \in [0,1] \mapsto \mathcal{Y}_0(m)$ is convex. This is done in a Markovian framework using PDE arguments. In this section, we provide a probabilistic proof of this result which hereby extends to our setting. The result is stated for the lower-semicontinuous envelope $\mathcal{Y}_{t*}$ of $\mathcal{Y}_t$ defined as

$$\mathcal{Y}_{t*}(\mu) := \lim_{\varepsilon \to 0} \operatorname{essinf}\{\mathcal{Y}_t(\mu') : \ |\mu' - \mu| \leq \varepsilon, \ \mu' \in \mathbf{L}_0([0,1], \mathcal{F}_t)\}, \tag{3.2}$$

for any $t \in [0,T]$. We refer to Proposition 3.1, the discussion before it and to (ii) of Remark 2.3 for conditions ensuring that $\mathcal{Y}_* = \mathcal{Y}$.

We first make precise the notion of convexity adapted to our non-Markovian setting. Fix a time $t \in [0,T]$.

Definition 3.1 ($\mathcal{F}_t$-convexity)**.**

(i) In the following, we say that a subset $D \subset \mathbf{L}_\infty(\mathbb{R}, \mathcal{F}_t)$ is $\mathcal{F}_t$-convex if $\lambda\mu_1 + (1-\lambda)\mu_2 \in D$, for all $\mu_1, \mu_2 \in D$ and $\lambda \in \mathbf{L}_0([0,1], \mathcal{F}_t)$.

(ii) Let D be an $\mathcal{F}_t$-convex subset of $\mathbf{L}_\infty(\mathbb{R}, \mathcal{F}_t)$. A map $\mathcal{J} : D \mapsto \mathbf{L}_2(\mathbb{R}, \mathcal{F}_t)$ is said to be $\mathcal{F}_t$-convex if

$$\mathrm{Epi}(\mathcal{J}) := \{(\mu, Y) \in D \times \mathbf{L}_2(\mathbb{R}, \mathcal{F}_t) : \ Y \geq \mathcal{J}(\mu)\}$$

is $\mathcal{F}_t$-convex.

(iii) Let $\mathrm{Epi}^c(\mathcal{Y}_t)$ be defined as the set of elements of the form $\sum_{n \leq N} \lambda_n(\mu_n, Y_n)$ with $(\mu_n, Y_n, \lambda_n)_{n \leq N} \subset \mathrm{Epi}(\mathcal{Y}_t) \times \mathbf{L}_0([0,1], \mathcal{F}_t)$ such that $\sum_{n \leq N} \lambda_n = 1$, for some $N \geq 1$. We then denote by $\overline{\mathrm{Epi}}^c(\mathcal{Y}_t)$ its closure in $\mathbf{L}_2$. Finally, the $\mathcal{F}_t$-convex envelope of $\mathcal{Y}_t$ is defined as

$$\mathcal{Y}_t^c(\mu) := \operatorname{essinf}\{Y \in \mathbf{L}_2(\mathbb{R}, \mathcal{F}_t) : (\mu, Y) \in \overline{\mathrm{Epi}}^c(\mathcal{Y}_t)\}. \tag{3.3}$$

We can now state the convexity property. It requires a right continuity property in time, which holds under the conditions of Theorem 2.1(ii), also recall (ii) of Remark 2.3.

Proposition 3.2. *Assume that $\mathcal{Y}_t(\mu) = \mathcal{Y}_{t+}(\mu)$ for any $\mu \in \mathbf{L}_0([0,1], \mathcal{F}_t)$ and $t < T$. Then, the map $\mu \in \mathbf{L}_0([0,1], \mathcal{F}_t) \mapsto \mathcal{Y}_{t*}(\mu)$ is $\mathcal{F}_t$-convex, for all $t < T$.*

Proof. Fix $t \in [0,T)$ and set $D := \mathbf{L}_0([0,1], \mathcal{F}_t)$ for ease of notations. The proof is divided in several steps.

Step 1. $(\mu, \mathcal{Y}_t^c(\mu)) \in \overline{\mathrm{Epi}}^c(\mathcal{Y}_t)$, *for all $\mu \in D$.*

Indeed, the family $F := \{Y \in \mathbf{L}_2(\mathbb{R}, \mathcal{F}_t) : (\mu, Y) \in \overline{\mathrm{Epi}}^c(\mathcal{Y}_t)\}$ is directed downward (for every fixed element μ in D) since $Y^1 \mathbf{1}_{\{Y^1 \leq Y^2\}} + Y^2 \mathbf{1}_{\{Y^1 > Y^2\}} \in F$, by $\mathcal{F}_t$-convexity of $\overline{\mathrm{Epi}}^c(\mathcal{Y}_t)$, for all $Y^1, Y^2 \in F$. It then follows from [13, Proposition VI.1.1] that there exists a sequence $(Y^n)_{n \geq 1} \subset F$ such that $Y^n \downarrow \mathcal{Y}_t^c(\mu)$ $\mathbb{P}-$a.s. Since Y^1 and $\mathcal{Y}_t^c(\mu) \in \mathbf{L}_2$, the monotone convergence Theorem implies that $Y^n \to \mathcal{Y}_t^c(\mu)$ in $\mathbf{L}_2$, as n goes to infinity. The set $\overline{\mathrm{Epi}}^c(\mathcal{Y}_t)$ being closed in $\mathbf{L}_2$, this proves our claim.

Step 2. *Let $\eta \in \mathbf{S}_2$ be as in Remark 2.1. Then, $|\mathcal{Y}^c_t(\mu)| \leq \eta_t$, for all $t \leq T$ and $\mu \in D$.*

We first observe that $\mathcal{Y} \geq \mathcal{Y}^c$ by construction. Remark 2.1 thus implies that $\mathcal{Y}^c_t(\mu) \leq \eta_t$. On the other hand, let $(Y^n)_{n\geq 1}$ be as in the step above. We claim that it satisfies $Y^n \geq -\eta_t$, for each $n \geq 1$. Then, the lower bound $\mathcal{Y}^c_t(\mu) \geq -\eta_t$ is obtained by passing to the limit. To see this, it suffices to prove this property for any $Y \in \mathbf{L}_2(\mathbb{R}, \mathcal{F}_t)$ such that $(\mu, Y) \in \overline{\text{Epi}}^c(\mathcal{Y}_t)$. But, such an element (μ, Y) is obtained by taking the $\mathbf{L}_2$ limit of elements of the form $\sum_{n\leq N} \lambda_n(\mu_n, Y_n)$ with $(\mu_n, Y_n, \lambda_n)_{n\leq N} \subset \text{Epi}(\mathcal{Y}_t) \times \mathbf{L}_0([0,1], \mathcal{F}_t)$, such that $\sum_{n\leq N} \lambda_n = 1$. Each Y_n of the latter family is bounded from below by $-\eta_t$ by Remark 2.1, and hence so is Y.

Step 3. *The map $\mu \in D \mapsto \mathcal{Y}^c_t(\mu)$ is $\mathcal{F}_t$-convex.*

Fix $\mu^1, \mu^2 \in D$ and $\lambda \in \mathbf{L}_0([0,1], \mathcal{F}_t)$. Step 1 implies that $(\mu^i, \mathcal{Y}^c_t(\mu^i)) \in \overline{\text{Epi}}^c(\mathcal{Y}_t)$ for $i = 1, 2$. Clearly, $\overline{\text{Epi}}^c(\mathcal{Y}_t)$ is $\mathcal{F}_t$-convex. It follows that $(\lambda\mu^1 + (1-\lambda)\mu^2, \lambda\mathcal{Y}^c_t(\mu^1) + (1-\lambda)\mathcal{Y}^c_t(\mu^2)) \in \overline{\text{Epi}}^c(\mathcal{Y}_t)$, so that $\lambda\mathcal{Y}^c_t(\mu^1) + (1-\lambda)\mathcal{Y}^c_t(\mu^2) \geq \mathcal{Y}^c_t(\lambda\mu^1 + (1-\lambda)\mu^2)$. Now, for any Y^i such that $(\mu^i, Y^i) \in \text{Epi}(\mathcal{Y}^c_t)$, one has $Y^i \geq \mathcal{Y}^c_t(\mu^i)$, $i = 1, 2$. This fact combined with the previous inequality thus implies $\lambda Y^1 + (1-\lambda)Y^2 \geq \mathcal{Y}^c_t(\lambda\mu^1 + (1-\lambda)\mu^2)$. This means that $\text{Epi}(\mathcal{Y}^c_t)$ is $\mathcal{F}_t$-convex.

Step 4. *$\mathcal{Y}_{t*}(\mu) \geq \mathcal{Y}^c_t(\mu)$, for all $\mu \in D$.*

Fix $\varepsilon > 0$ and set $D^\varepsilon_\mu := \{|\mu' - \mu| \leq \varepsilon,\ \mu' \in \mathbf{L}_0([0,1], \mathcal{F}_t)\}$. It follows from Remark 2.2 that the family $\{\mathcal{Y}_t(\mu') :\ \mu' \in D^\varepsilon_\mu\}$ is directed downward. Then, we can find a sequence $(\mu^\varepsilon_n)_{n\geq 1} \subset D^\varepsilon_\mu$ such that

$$\mathcal{Y}_t(\mu^\varepsilon_n) \to Z_\varepsilon(\mu) := \text{essinf}\{\mathcal{Y}_t(\mu') :\ \mu' \in D^\varepsilon_\mu\} \quad \mathbb{P} - \text{a.s.}$$

Since $(Z_\varepsilon(\mu))_{\varepsilon>0}$ is non-decreasing, $\lim_{N\to\infty} Z_{1/N}(\mu) = \mathcal{Y}_{t*}(\mu)$, recall (3.2). Note that Remark 2.1 implies that $(\mathcal{Y}_t(\mu^{1/N}_n))_{n\geq 1} \to_n Z_{1/N}(\mu)$ in $\mathbf{L}^2$ and define

$$k_N := \min\{n \geq 1 : \|\mathcal{Y}_t(\mu^{1/N}_n) - Z_{1/N}(\mu)\|_{\mathbf{L}^2} \leq 1/N\}.$$

Then, $(\mu^{1/N}_{k_N}, \mathcal{Y}_t(\mu^{1/N}_{k_N})) \to (\mu, \mathcal{Y}_{t*}(\mu))$ in $\mathbf{L}^2$ as $N \to \infty$. Since $\text{Epi}(\mathcal{Y}_t) \subset \overline{\text{Epi}}^c(\mathcal{Y}_t)$ and the latter is closed under $\mathbf{L}^2$-convergence, this implies that $(\mu, \mathcal{Y}_{t*}(\mu)) \in \overline{\text{Epi}}^c(\mathcal{Y}_t)$. We conclude by appealing to the definition of $\mathcal{Y}^c_t$ in (3.3).

Step 5. *$\mathcal{Y}^c_t(\mu) \geq \mathcal{Y}_{t*}(\mu)$, for all $\mu \in D$.*

In view of Steps 3 and 4, the result of Step 5 actually proves that $\mathcal{Y}_{t*} = \mathcal{Y}^c_t$ is $\mathcal{F}_t$-convex. We now proceed to the proof of Step 5 which is itself divided in two parts.

Step 5.a It follows from Step 1, that there exists a sequence

$$(\mu_n, Y_n, \lambda^N_n)_{n\geq 1, N\geq 1} \subset \text{Epi}(\mathcal{Y}_t) \times \mathbf{L}_0([0,1], \mathcal{F}_t) \tag{3.4}$$

such that $\sum_{n\leq N} \lambda^N_n = 1$, for all N, and

$$(\hat{\mu}_N, \hat{Y}_N) := \sum_{n\leq N} \lambda^N_n(\mu_n, Y_n) \to (\mu, \mathcal{Y}^c_t(\mu)) \ \text{ in } \mathbf{L}_2. \tag{3.5}$$

Fix $N \geq 1$ and $\varepsilon > 0$. Let $\hat{\alpha}^N \in \mathbf{H}_2$ be such that $\hat{\mu}_N = m_o + \int_0^t \hat{\alpha}_s^N dW_s$. Since the family $(\lambda_n^N)_{n \leq N}$ is composed of $\mathcal{F}_t$-measurable random variables summing to 1, one can find $\alpha^N \in \mathbf{H}_2$ and a random variable $\xi_N^\varepsilon \in \mathbf{L}_2(\mathcal{F}_{t+\varepsilon})$ such that

$$\hat{\mu}_N + \int_t^{t+\varepsilon} \alpha_s^N dW_s = \xi_N^\varepsilon \text{ and } \mathbb{P}\left[\xi_N^\varepsilon = \mu_n | \mathcal{F}_t\right] = \lambda_n^N, \quad \text{for } n \leq N. \tag{3.6}$$

Without loss of generality, we can assume that $\alpha^N = \hat{\alpha}^N \; dt \times d\mathbb{P}$ on $[0,t]$. Then, (i) of Theorem 2.1 and Remark 2.2 yield

$$\mathcal{Y}_t(\hat{\mu}_N) = \mathcal{Y}_t^{\hat{\alpha}^N} \leq \mathcal{E}^g_{t,t+\varepsilon}(\mathcal{Y}_{t+\varepsilon}^{\alpha^N}) = \mathcal{E}^g_{t,t+\varepsilon}(\mathcal{Y}_{t+\varepsilon}(\xi_N^\varepsilon)) = \mathcal{E}^g_{t,t+\varepsilon}\left(\sum_{n \leq N} 1_{\xi_N^\varepsilon = \mu_n} \mathcal{Y}_{t+\varepsilon}(\mu_n)\right). \tag{3.7}$$

We claim that

$$\liminf_{\varepsilon \to 0} \mathcal{E}^g_{t,t+\varepsilon}\left(\sum_{n \leq N} 1_{\xi_N^\varepsilon = \mu_n} \mathcal{Y}_{t+\varepsilon}(\mu_n)\right) \leq \sum_{n \leq N} \lambda_n^N \mathcal{Y}_t(\mu_n). \tag{3.8}$$

Then, (3.7), (3.8), (3.4) and (3.5) lead to

$$\mathcal{Y}_t(\hat{\mu}_N) \leq \sum_{n \leq N} \lambda_n^N \mathcal{Y}_t(\mu_n) \leq \sum_{n \leq N} \lambda_n^N Y_n = \hat{Y}_N.$$

Appealing to (3.5), we deduce that

$$\liminf_{N \to \infty} \mathcal{Y}_t(\hat{\mu}_N) \leq \mathcal{Y}_t^c(\mu).$$

Since $\hat{\mu}_N \to \mu \; \mathbb{P}-$ a.s., this together with Remark 2.2 implies that

$$Z_\varepsilon(\mu) \leq \liminf_{N \to \infty} \mathcal{Y}_t(\bar{\mu}_N) = \liminf_{N \to \infty} \left(\mathcal{Y}_t(\hat{\mu}_N) 1_{\{|\hat{\mu}_N - \mu| \leq \varepsilon\}} + \mathcal{Y}_t(\mu) 1_{\{|\hat{\mu}_N - \mu| > \varepsilon\}}\right) \leq \mathcal{Y}_t^c(\mu),$$

for all $\varepsilon > 0$, where

$$\bar{\mu}_N := \hat{\mu}_N 1_{\{|\hat{\mu}_N - \mu| \leq \varepsilon\}} + \mu 1_{\{|\hat{\mu}_N - \mu| > \varepsilon\}} \in D_\mu^\varepsilon,$$

see Step 4 for the definitions of $Z_\varepsilon(\mu)$ and D_μ^ε. Since $Z_\varepsilon(\mu) \uparrow \mathcal{Y}_{t*}(\mu)$ as $\varepsilon \downarrow 0$ by (3.2), this shows the required result.

Step 5.b It finally remains to prove the claim (3.8).
Remark 2.1 and (ii) of Proposition 6.2 in the Appendix imply that

$$\begin{aligned}
\mathcal{E}^g_{t,t+\varepsilon}\left(\sum_{n \leq N} 1_{\xi_N^\varepsilon = \mu_n} \mathcal{Y}_{t+\varepsilon}(\mu_n)\right) &\leq E_t\left[\sum_{n \leq N} 1_{\xi_N^\varepsilon = \mu_n} \mathcal{Y}_{t+\varepsilon}(\mu_n)\right] + \eta_\varepsilon \\
&\leq E_t\left[\sum_{n \leq N} 1_{\xi_N^\varepsilon = \mu_n} \mathcal{Y}_t(\mu_n)\right] + \eta_\varepsilon \\
&+ \sum_{n \leq N} E_t\left[|\mathcal{Y}_{t+\varepsilon}(\mu_n) - \mathcal{Y}_t(\mu_n)|\right],
\end{aligned}$$

where $\eta_\varepsilon \to 0$ $\mathbb{P}-$a.s. as $\varepsilon \to 0$. The right-hand side of (3.6) then leads to

$$\begin{aligned}\mathcal{E}^g_{t,t+\varepsilon}\left(\sum_{n\leq N} 1_{\xi^\varepsilon_N=\mu_n}\mathcal{Y}_{t+\varepsilon}(\mu_n)\right) &\leq \sum_{n\leq N}\lambda^N_n\mathcal{Y}_t(\mu_n)+\eta_\varepsilon \\ &+ \sum_{n\leq N} E_t\left[|\mathcal{Y}_{t+\varepsilon}(\mu_n)-\mathcal{Y}_t(\mu_n)|\right].\end{aligned}$$

Recall that $\mathcal{Y}_{t+}(\mu_n) = \mathcal{Y}_t(\mu_n)$ by assumption, and that $(\mathcal{Y}(\mu_n))_n$ is bounded by some $\eta \in \mathbf{S}_2$, see Remark 2.1. Sending $\varepsilon \to 0$ in the above inequality and appealing to the Lebesgue dominated convergence Theorem proves (3.8). □

In the context of PDEs, convexity in the domain propagates up to the boundary, which leads to a boundary layer phenomenon. In [3] and [12] this translates in the fact that the natural T-time boundary condition should be stated in terms of the m-convex envelope of Φ. We observe hereafter that this property extends to our non-Markovian setting, whenever Φ is deterministic.

We recall from Theorem 2.1 (i) that $\mathcal{Y}$ can be associated to a làdlàg process, up to indistinguishability. As opposed to Proposition 3.2, we shall not need to impose any right-continuity for the following.

Proposition 3.3. *Assume that Φ is deterministic and that its convex envelope $\hat{\Phi}$ is continuous on $[0,1]$. Then,*

$$\mathcal{Y}^\alpha_{T-} = \hat{\Phi}(M^\alpha_T) \;\; \textit{and} \;\; \mathcal{Y}^\alpha_\tau = \operatorname{ess\,inf}_{\alpha'\in\mathbf{A}^\alpha_\tau} \mathcal{E}^g_\tau\left[\hat{\Phi}(M^{\alpha'}_T)\right],$$

for all $\alpha \in \mathbf{A}_0$ and $\tau \in \mathcal{T}$ such that $\tau < T$.

Before proving this result, let us make some observations.

Remark 3.2. Since Φ is non-decreasing, its convex envelope is continuous on $[0,1)$ and coincides with Φ at points $1-$ and 1. Hence $\hat{\Phi}$ is continuous on $[0,1]$ if and only if Φ is left-continuous at 1.

Remark 3.3. In Section 2.3, we observed that the essential infimum in the dynamic programming principle is attained whenever Φ and g are convex. Hence, the previous proposition allows straightforwardly to avoid the convexity requirement on Φ, whenever it is deterministic.

Remark 3.4. The proof below can easily be adapted to the case where $\Phi(\omega,m) = \phi(m)\xi(\omega)$ for some non-negative random variable ξ and a deterministic map ϕ. This is due to the fact that the m-convex envelope of Φ is fully characterized by the convex envelope $\hat{\phi}$ of ϕ: $\hat{\Phi}(\omega,m) = \hat{\phi}(m)\xi(\omega)$. This allows one to follow the construction used in our proof. In particular, in the quantile hedging problem of Fölmer and Leukert [10], one has $\Phi(\omega,m) = \mathbf{1}_{\{m>0\}}\xi(\omega)$ $(m \in [0,1])$, with ξ taking non-negative values, so that $\hat{\Phi}(\omega,m) = m\xi(\omega)$, see also [3].

Remark 3.5. Our proof could also be adapted to Markovian settings in which the randomness in Φ is driven by the terminal value of an SDE. This will be done in the context of Section 4 below, see Proposition 4.3. Again, it is clear that a more general probabilistic argument could be used as well under suitable regularity conditions.

Proof of Proposition 3.3. We prove each assertion separately.
Step 1. By definition of the convex envelope, we can find a measurable map $m \in [0,1] \mapsto (\underline{\wp}(m), \overline{\wp}(m), \varepsilon(m)) \in [0,1]^3$ such that $\underline{\wp}(m) \leq m \leq \overline{\wp}(m)$, $\varepsilon(m)\underline{\wp}(m)+(1-\varepsilon(m))\overline{\wp}(m) = m$ and

$$\hat{\Phi}(m) = \varepsilon(m)\Phi(\underline{\wp}(m)) + (1-\varepsilon(m))\Phi(\overline{\wp}(m)) \ ,$$

for any $m \in [0,1]$. Let $t_n \uparrow T$. Then, one can find $\alpha^n \in \mathbf{A}^\alpha_{t_n}$ and $\xi^n \in \mathbf{L}_0([0,1])$ such that $M^{\alpha^n}_T = M^\alpha_{t_n} + \int_{t_n}^T \alpha^n_s dW_s = \xi^n$, where $\mathbb{P}\left[\xi^n = \underline{\wp}(M^\alpha_{t_n})|\mathcal{F}_{t_n}\right] = \varepsilon(M^\alpha_{t_n})$ and $\mathbb{P}\left[\xi^n = \overline{\wp}(M^\alpha_{t_n})|\mathcal{F}_{t_n}\right] = 1-\varepsilon(M^\alpha_{t_n})$. It follows from the above and (iii) of Proposition 6.2 in the Appendix that

$$\mathcal{Y}^\alpha_{t_n} \quad \leq \quad E_{t_n}\left[\Phi(\xi^n)\right] + \eta_n = \hat{\Phi}(M^\alpha_{t_n}) + \eta_n,$$

where $\eta_n \to 0$ as $n \to \infty$. Since $\mathcal{Y}$ can be considered làdlàg, up to indistinguishability (by Proposition 5.2), passing to the limit implies that

$$\mathcal{Y}^\alpha_{T-} \quad \leq \quad \hat{\Phi}(M^\alpha_T). \tag{3.9}$$

We now prove the converse inequality. We use (iii) in Proposition 6.2 in the Appendix and Jensen's inequality to deduce that

$$Y^{\alpha'}_{t_n} \quad := \quad \mathcal{E}^g_{t_n,T}[\Phi(M^{\alpha'}_T)] \geq E_{t_n}\left[\hat{\Phi}(M^{\alpha'}_T)\right] - \bar{\eta}_n \geq \hat{\Phi}(M^\alpha_{t_n}) - \bar{\eta}_n \ , \quad \alpha' \in \mathbf{A}^\alpha_{t_n},$$

where $\bar{\eta}_n \to 0$ as $n \to \infty$. Combining the arbitrariness of $\alpha' \in \mathbf{A}^\alpha_{t_n}$ with the làdlàg property of $\mathcal{Y}$, we get that

$$\mathcal{Y}^\alpha_{T-} \geq \liminf_{n\to\infty} \operatorname{ess\,inf}_{\alpha' \in \mathbf{A}^\alpha_{t_n}} Y^{\alpha'}_{t_n} \geq \hat{\Phi}(M^\alpha_T) \ .$$

Step 2. It follows from Theorem 2.1 (i) that

$$\mathcal{Y}^\alpha_\tau = \operatorname{ess\,inf}_{\alpha' \in \mathbf{A}^\alpha_\tau} \mathcal{E}^g_{\tau,t_n\vee\tau}[\mathcal{Y}^{\alpha'}_{t_n\vee\tau}] \ , \quad n \in \mathbb{N} \ .$$

The process $\mathcal{Y}^{\alpha'}_{.\vee\tau}$ being làdlàg, $\lim_{n\to\infty} \mathcal{Y}^{\alpha'}_{t_n\vee\tau} = \mathcal{Y}^{\alpha'}_{T-}$ is well-defined. Moreover, it follows from the bound in Remark 2.1 that the convergence holds in $\mathbf{L}_2$. In view of the stability result of Proposition 6.1 and Step 1. above, passing to the limit as $n \to \infty$ leads to

$$\mathcal{Y}^\alpha_\tau \leq \operatorname{ess\,inf}_{\alpha' \in \mathbf{A}^\alpha_\tau} \mathcal{E}^g_\tau[\mathcal{Y}^{\alpha'}_{T-}] = \operatorname{ess\,inf}_{\alpha' \in \mathbf{A}^\alpha_\tau} \mathcal{E}^g_\tau[\hat{\Phi}(M^{\alpha'}_T)].$$

Since $\Phi \geq \hat{\Phi}$, the reverse inequality holds by definition of $\mathcal{Y}^\alpha_\tau$ in (2.6). □

3.3 Dual representation

In this section, we provide a dual formulation for the minimal initial condition at time 0, $m \mapsto \mathcal{Y}_0(m)$. It requires the introduction of the Fenchel transforms of g and Φ.

We therefore define

$$\tilde{\Phi} \ : \ (\omega, l) \in \Omega \times \mathbb{R} \mapsto \sup_{m \in [0,1]} (ml - \Phi(\omega, m))$$

and

$$\tilde{g} \ : \ (\omega, t, u, v) \in \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \mapsto \sup_{(y,z) \in \mathbb{R} \times \mathbb{R}^d} \left(yu + z^\top v - g(\omega, t, y, z) \right).$$

Remark 3.1. It follows from the assumption $(\mathbf{H}_g)$ that the domain of $\tilde{g}(\omega, t, \cdot)$, $\mathrm{dom}(\tilde{g}(\omega, t, \cdot))$, is contained in $[-K_g, K_g]^{d+1}$ for $\mathbb{P}-$a.e. $\omega \in \Omega$ and all $t \leq T$. The assumption $(\mathbf{H}_\Psi)$ ensures that the domain of $\tilde{\Phi}(\omega, \cdot)$ is the all real line, $\mathbb{P}-$a.s..

In the following, we denote by $\mathbf{\Lambda}$ the set of predictable processes λ with values in $\mathbb{R} \times \mathbb{R}^d$ such that $\lambda_t(\omega) \in \mathrm{dom}(\tilde{g}(\omega, t, \cdot))$ for Leb $\times \mathbb{P}$-a.e. $(\omega, t) \in \Omega \times [0,T]$.

To $\lambda = (\nu, \vartheta) \in \mathbf{\Lambda}$, we associate the process L^λ defined by

$$L_t^\lambda = 1 + \int_0^t L_s^\lambda \nu_s ds + \int_0^t L_s^\lambda \vartheta_s dW_s \ , \ t \in [0,T].$$

Our dual formulation for $\mathcal{Y}_0$ is stated in terms of

$$\mathcal{X}_0(l) := \inf_{\lambda \in \mathbf{\Lambda}} \mathrm{X}_0^{l,\lambda}, \ l > 0,$$

where

$$\mathrm{X}_0^{l,\lambda} := E\left[\int_0^T L_s^\lambda \tilde{g}(s, \lambda_s) ds + L_T^\lambda \tilde{\Phi}(l / L_T^\lambda) \right] , \quad \lambda \in \mathbf{\Lambda} \ , \quad l > 0.$$

The fact that the Fenchel transform of $\mathcal{X}_0$ provides a lower bound for $\mathcal{Y}_0$ is straightforward, and detailed in Proposition 3.4 below for the convenience of the reader. For ease of notations, we now write $\mathbf{A}_m$ for $\mathbf{A}_{0,m}$, $M^{m,\alpha}$ for $M^{(0,m),\alpha}$, and denote by $(Y^{m,\alpha}, Z^{m,\alpha})$ the solution of the BSDE$(g, \Phi(M_T^{m,\alpha}))$, $\alpha \in \mathbf{A}_m$.

Proposition 3.4. *$\mathcal{Y}_0(m) \geq \sup_{l>0} (lm - \mathcal{X}_0(l))$, for all $m \in [0,1]$.*

Proof. Fix $\alpha \in \mathbf{A}_m$ and $\lambda = (\nu, \vartheta) \in \mathbf{\Lambda}$. Then, it follows from the definition of $\tilde{\Phi}$ and $\tilde{g}$ that

$$\begin{aligned} E\left[Y_T^{m,\alpha} L_T^\lambda \right] &= Y_0^{m,\alpha} + E\left[\int_0^T L_s^\lambda \left(\nu_s Y_s^{m,\alpha} + \vartheta_s^\top Z_s^{m,\alpha} - g(s, Y_s^{m,\alpha}, Z_s^{m,\alpha}) \right) ds \right] \\ &\leq Y_0^{m,\alpha} + E\left[\int_0^T L_s^\lambda \tilde{g}(s, \lambda_s) ds \right], \end{aligned}$$

and

$$Y_T^{m,\alpha} L_T^\lambda = \Phi(M_T^{m,\alpha}) L_T^\lambda \geq l M_T^{m,\alpha} - L_T^\lambda \tilde{\Phi}(l / L_T^\lambda),$$

for $l > 0$. Note that, in the above, we have cancelled the expectation of the local martingale part $\int_0^T (L_s^\lambda Z_s^{m,\alpha} + Y_s^{m,\alpha} L_s^\lambda \vartheta_s) dW_s$ although $L^\lambda Z^{m,\alpha}$ might not belong to $\mathbf{H}_2$. If not, one may use a localization argument since all other terms belongs to L^1 uniformly in time. Combining the above and using the martingale property of $M^{m,\alpha}$ yields

$$Y_0^{m,\alpha} \geq lm - E\left[\int_0^T L_s^\lambda \tilde{g}(s, \lambda_s) ds + L_T^\lambda \tilde{\Phi}(l/L_T^\lambda)\right] = lm - X_0^{l,\lambda} .$$

The result follows from the arbitrariness of $l > 0$, $\lambda \in \mathbf{\Lambda}$, and $\alpha \in \mathbf{A}_m$. □

We now show that equality is satisfied in Proposition 3.4 whenever existence holds in the dual problem. This is proved under the following assumptions. Let C_b^1 be the set of continuously differentiable maps with bounded first derivatives.

Assumption $(\mathbf{H_d^1})$ The following holds for Leb $\times \mathbb{P}$-a.e. $(t, \omega) \in [0, T] \times \Omega$:

(a) the maps $\tilde{\Phi}(\omega, \cdot)$ and $\tilde{g}(\omega, \cdot)$ are C_b^1 on their domain, and $\mathrm{dom}(\tilde{g}(\omega, t, \cdot))$ is closed;

(b) $|\nabla \tilde{\Phi}(\omega, \cdot)| + |\nabla \tilde{g}(\omega, t, \cdot)| \leq \chi_{\tilde{\Phi}, \tilde{g}}(\omega)$, for some $\chi_{\tilde{\Phi}, \tilde{g}} \in \mathbf{L}_2(\mathbb{R})$;

(c) $\Phi(\omega, m) = \sup_{l>0} \Big(lm - \tilde{\Phi}(\omega, l) \Big)$, for all $m \in [0, 1]$;

(d) $g(\omega, t, y, z) = \max_{(u,v) \in \mathrm{dom}(\tilde{g}(\omega,t,\cdot))} \big(yu + z^\top v - \tilde{g}(\omega, t, u, v) \big)$, for all $(y, z) \in \mathbb{R} \times \mathbb{R}^d$.

In the above, $\nabla \tilde{\Phi}$ and $\nabla \tilde{g}$ stands for the gradient with respect to l and (u, v) respectively.

Note that (a) and (b) are of technical nature, while (c) and (d) mean that Φ and g are convex, i.e. coincide with their bi-dual. The latter is a minimal requirement if one wants the duality to hold.

Proposition 3.5. *Let Assumption* $(\mathbf{H_d^1})$ *hold. Assume further that there exists* $\hat{l} > 0$ *and* $\hat{\lambda} \in \mathbf{\Lambda}$ *such that*

$$\sup_{l>0} (lm - \mathcal{X}_0(l)) = \hat{l}m - \mathcal{X}_0(\hat{l}) = \hat{l}m - \mathrm{X}_0^{\hat{l},\hat{\lambda}}. \tag{3.10}$$

Then, there exists $\hat{\alpha} \in \mathbf{A}_m$ *such that*

$$\mathcal{Y}_0(m) = Y_0^{m,\hat{\alpha}} = \hat{l}m - \mathcal{X}_0(\hat{l}).$$

It satisfies

$$g(\cdot, Y^{m,\hat{\alpha}}, Z^{m,\hat{\alpha}}) = \hat{\lambda}^\top (Y^{m,\hat{\alpha}}, Z^{m,\hat{\alpha}}) - \tilde{g}(\cdot, \hat{\lambda}) \;\; \textit{and} \;\; \Phi(M_T^{m,\hat{\alpha}}) = M_T^{m,\hat{\alpha}} \hat{l}/L_T^{\hat{\lambda}} - \tilde{\Phi}(\hat{l}/L_T^{\hat{\lambda}}). \tag{3.11}$$

Before to provide the proof, let us make the following observation which pertains for the case of a linear driver g.

Remark 3.2. Assume that g is linear, i.e. there exist bounded predictable processes A^Y and A^Z such that $g : (\omega, t, y, z) \mapsto g(\omega, t, 0, 0) + A^Y_t(\omega)y + A^Z_t(\omega)z$. In this case, $\mathbf{\Lambda} = \{(A^Y, A^Z)\}$ and therefore

$$\mathcal{X}_0(l) = E\left[\int_0^T L_s \tilde{g}(s, A^Y_s, A^Z_s)ds + L_T\tilde{\Phi}(l/L_T)\right] ,$$

with L given by

$$L_t = 1 + \int_0^t L_s A^Y_s\, ds + \int_0^t L_s A^Z_s\, dW_s\ ,\ t \in [0, T].$$

Then, the dual formulation of Proposition 3.5 above drops down to finding $\hat{l}$ which maximizes $lm - \mathcal{X}_0(l)$. This generalizes the result of [10] and [3] obtained for quantile hedging problems in linear models of financial markets.

Proof of Proposition 3.5. We split the proof in two steps.

Step 1. For ease of notations, we set $\hat{L} := L^{\hat{\lambda}}$. By optimality of $\hat{l}$, one has

$$\hat{l}m - E\left[\hat{L}_T\tilde{\Phi}(\hat{l}/\hat{L}_T)\right] \geq m(\hat{l} + \iota) - E\left[\hat{L}_T\tilde{\Phi}((\hat{l} + \iota)/\hat{L}_T)\right] ,$$

for all $\iota > -\hat{l}$. Since $\tilde{\Phi}$ is by construction $\mathbb{P}-$a.s. convex, this implies that $\zeta_\iota := \nabla\tilde{\Phi}((\hat{l} + \iota)/\hat{L}_T)$ satisfies $m\iota \leq E[\zeta_\iota]\iota$, for all $\iota > -\hat{l}$, recall $(\mathbf{H^1_d})$ (a) and (b). Taking ι of the form $-1/n$ and then $1/n$, for $n \to \infty$, and using $(\mathbf{H^1_d})$ (a) and (b) then leads to

$$m = E[\zeta] \ \text{ where } \ \zeta := \nabla\tilde{\Phi}(\hat{l}/\hat{L}_T). \tag{3.12}$$

We now appeal to $(\mathbf{H^1_d})$ (c) to deduce that

$$\Phi(\zeta) = \zeta(\hat{l}/\hat{L}_T) - \tilde{\Phi}(\hat{l}/\hat{L}_T). \tag{3.13}$$

By construction, $\tilde{\Phi}$ is $\mathbb{P}-$a.s. 1-Lipschitz and non-decreasing, i.e. $\zeta \in \mathbf{L}_0([0, 1])$. In view of (3.12), the martingale representation Theorem then implies that we can find $\hat{\alpha} \in \mathbf{A}_m$ such that $\hat{M}_T := M^{m,\hat{\alpha}}_T = \zeta$.

Step 2. We now write $(\hat{\nu}, \hat{\vartheta}) := \hat{\lambda}$ and fix $\lambda = (\nu, \theta) \in \mathbf{\Lambda}$ to be chosen later on. Clearly, $\mathbf{\Lambda}$ is convex. Hence, $\lambda^\varepsilon := (1-\varepsilon)(\hat{\nu}, \hat{\vartheta}) + \varepsilon(\nu, \vartheta) \in \mathbf{\Lambda}$, $\varepsilon \in [0, 1]$. Moreover, direct computations show that

$$\frac{\partial}{\partial\varepsilon}L^{\lambda^\varepsilon}|_{\varepsilon=0} = \hat{L}\hat{R} \ \text{ where } \ \hat{R} := \int_0^\cdot (\delta\nu_s - \delta\vartheta_s\hat{\vartheta}_s)ds + \int_0^\cdot \delta\vartheta_s dW_s,$$

in which we use the notations $\delta\lambda := (\delta\nu, \delta\vartheta) := (\nu - \hat{\nu}, \vartheta - \hat{\vartheta})$.

Recalling that elements of $\mathbf{\Lambda}$ take bounded values, see Remark 3.1, and arguing as in Step 1, one easily checks that the optimality condition $\mathrm{X}_0^{\hat{l},\lambda^\varepsilon} \geq \mathrm{X}_0^{\hat{l},\hat{\lambda}}$, for all $\varepsilon \in [0, 1]$, implies that $\hat{\eta} := \nabla\tilde{g}(\cdot, \hat{\lambda})$ satisfies

$$\begin{aligned} 0 \ &\leq \ E\left[\int_0^T \hat{L}_s\left(\hat{R}_s\tilde{g}(s, \hat{\lambda}_s) + \hat{\eta}_s^\top\delta\lambda_s\right)ds + \hat{R}_T\hat{L}_T(\tilde{\Phi}(\hat{l}/\hat{L}_T) - (\hat{l}/\hat{L}_T)\nabla\tilde{\Phi}(\hat{l}/\hat{L}_T))\right] \\ &= \ E\left[\int_0^T \hat{L}_s\left(\hat{R}_s\tilde{g}(s, \hat{\lambda}_s) + \hat{\eta}_s^\top\delta\lambda_s\right)ds - \hat{R}_T\hat{L}_T\Phi(\hat{M}_T)\right] , \end{aligned} \tag{3.14}$$

in which we used (3.12), (3.13) and the relation $\zeta = \hat{M}_T$ to deduce the second equality. Let $(\hat{Y}, \hat{Z})$ be defined by

$$\hat{Y} := \hat{L}^{-1}_{\cdot} E_{\cdot}\left[\hat{L}_T \Phi(\hat{M}_T) - \int_{\cdot}^{T} \hat{L}_s \tilde{g}(s, \hat{\lambda}_s) ds\right] \text{ and } \hat{Z} := \bar{Z} - \hat{Y}\hat{\vartheta}, \tag{3.15}$$

where $\bar{Z} \in \mathbf{H}_2$ is implicitly given by

$$\hat{L}_t \hat{Y}_t = \hat{L}_T \Phi(\hat{M}_T) - \int_t^T \hat{L}_s \tilde{g}(s, \hat{\lambda}_s) ds - \int_t^T \hat{L}_s \bar{Z}_s dW_s, \quad 0 \le t \le T\,. \tag{3.16}$$

The above combined with (3.14) implies

$$0 \;\le\; E\left[\int_0^T \hat{L}_s \left(\hat{R}_s \tilde{g}(s, \hat{\lambda}_s) + \hat{\eta}_s^\top \delta\lambda_s\right) ds - \hat{R}_T \hat{L}_T \hat{Y}_T\right].$$

Recalling the definition of $\hat{R}$ and $\hat{\eta}$ and applying Itô's Lemma, this leads to

$$0 \le E\left[\int_0^T \hat{L}_s \left(\hat{\eta}_s - (\hat{Y}_s, \hat{Z}_s)\right)^\top \delta\lambda_s ds\right] = E\left[\int_0^T \hat{L}_s \left(\nabla \tilde{g}(s, \hat{\lambda}_s) - (\hat{Y}_s, \hat{Z}_s)\right)^\top \delta\lambda_s ds\right]. \tag{3.17}$$

By Assumption $(\mathbf{H^1_d})$ (a), Remark 3.1 and [1, Theorem 18.19, p. 605], one can choose $\bar{\lambda} \in \mathbf{\Lambda}$ such that

$$\bar{\lambda} = \operatorname{argmin}\{f(\cdot, u, v),\ (u, v) \in \operatorname{dom}(\tilde{g}(\cdot))\} \text{ Leb} \times \mathbb{P}\text{−a.e.}$$

where

$$f : (\omega, s, u, v) \mapsto \left(\nabla \tilde{g}(\omega, s, \hat{\lambda}_s(\omega)) - (\hat{Y}_s(\omega), \hat{Z}_s(\omega))\right)^\top (u - \hat{\nu}_s(\omega), v - \hat{\vartheta}_s(\omega)).$$

Considering now Relation (3.17) with λ chosen to be equal to $\bar{\lambda}\mathbf{1}_{\{f(\cdot,\bar{\lambda})<0\}}$, we see that, for Leb $\times \mathbb{P}$-a.e. $(\omega, t) \in \Omega \times [0, T]$, the gradient $\Delta_t(\omega)$ at $\hat{\lambda}_t(\omega)$ of the convex map

$$(u, v) \in \operatorname{dom}(\tilde{g}(\omega, t, \cdot)) \mapsto F(\omega, t, u, v) := \tilde{g}(\omega, t, u, v) - u\hat{Y}_t(\omega) - v^\top \hat{Z}_t(\omega)$$

satisfies

$$\Delta_t(\omega)^\top (b - \hat{\lambda}_t(\omega)) \ge 0\,, \qquad \text{for all } b \in \operatorname{dom}(\tilde{g}(\omega, t, \cdot)).$$

This implies that $\hat{\lambda}_t(\omega)$ minimizes $F(\omega, t, \cdot)$ for Leb $\times \mathbb{P}$-a.e. $(\omega, t) \in \Omega \times [0, T]$ and therefore we compute

$$\tilde{g}(\cdot, \hat{\lambda}) = \hat{\lambda}^\top(\hat{Y}, \hat{Z}) - g(\cdot, \hat{Y}, \hat{Z}) \ \text{ Leb} \times \mathbb{P} - a.e.$$

by $(\mathbf{H^1_d})$ (d). Combining the above identity with (3.16) leads to $(\hat{Y}, \hat{Z}) = (Y^{m,\hat{\alpha}}, Z^{m,\hat{\alpha}})$. Then, by using (3.12), (3.13) and (3.15), in which $\hat{L}_0 = 1$, we obtain

$$\begin{aligned} Y_0^{m,\hat{\alpha}} &= E\left[\hat{L}_T \Phi(\hat{M}_T) - \int_0^T \hat{L}_s \tilde{g}(s, \hat{\lambda}_s) ds\right] \\ &= E\left[\hat{L}_T \left(\zeta \hat{l}/\hat{L}_T - \tilde{\Phi}(\hat{l}/\hat{L}_T)\right) - \int_0^T \hat{L}_s \tilde{g}(s, \hat{\lambda}_s) ds\right] \\ &= \hat{l}m - E\left[\hat{L}_T \tilde{\Phi}(\hat{l}/\hat{L}_T) + \int_0^T \hat{L}_s \tilde{g}(s, \hat{\lambda}_s) ds\right]. \end{aligned}$$

In view of Proposition 3.4, this concludes the proof. □

We now state the reciprocal statement: existence in the primal problem provides existence in the dual one. Here again, we need to impose some additional technical conditions.

Assumption $(\mathbf{H^2_d})$ The following holds for Leb $\times \mathbb{P}$-a.e. $(t,\omega) \in [0,T] \times \Omega$:

(a) the maps $\Phi(\omega,\cdot)$ and $g(\omega,t,\cdot)$ are C^1_b on $[0,1]$ and $\mathbb{R} \times \mathbb{R}^d$ respectively;

(b) $|\nabla\Phi(\omega,\cdot)| \leq \chi_\Phi(\omega)$, for some $\chi_\Phi \in \mathbf{L}_2(\mathbb{R})$.

Proposition 3.6. *Let Assumption* $(\mathbf{H^2_d})$ *hold. Let* $l > 0$ *be fixed and assume that there exists* $\hat{m} \in [0,1]$ *and* $\hat{\alpha} \in \mathbf{A}_{\hat{m}}$ *such that*

$$\sup_{m\in[0,1]} \sup_{\alpha\in\mathbf{A}_m} (ml - \mathcal{Y}_0(m)) = \hat{m}l - Y_0^{\hat{m},\hat{\alpha}}. \tag{3.18}$$

Then, there exists $\hat{\lambda} \in \mathbf{\Lambda}$ *such that*

$$\mathcal{Y}_0(\hat{m}) = \hat{m}l - \mathcal{X}_0(l) = \hat{m}l - \mathrm{X}_0^{l,\hat{\lambda}} ,$$

and $\hat{\lambda}$ *satisfies* (3.11) *with* $m = \hat{m}$ *and* $\hat{l} = l$.

Proof. Given $\varepsilon \in [0,1]$, a martingale M with values in $[0,1]$, $m := M_0$, we set $m_\varepsilon := \hat{m} + \varepsilon(m - \hat{m})$, $M^\varepsilon := \hat{M} + \varepsilon(M - \hat{M})$, where $\hat{M} := M^{\hat{m},\hat{\alpha}}$. For ease of notation, we set $(\hat{Y},\hat{Z}) := (Y^{\hat{m},\hat{\alpha}}, Z^{\hat{m},\hat{\alpha}})$ and denote by $(Y^\varepsilon, Z^\varepsilon)$ the solution of BSDE$(g, \Phi(M^\varepsilon_T))$, $\delta m := m - \hat{m}$, $(\delta M, \delta Y^\varepsilon, \delta Z^\varepsilon) := (M - \hat{M}, Y^\varepsilon - \hat{Y}, Z^\varepsilon - \hat{Z})$.

Step 1. *We first show that* $\varepsilon^{-1}(\delta Y^\varepsilon_s, \delta Z^\varepsilon_s)$ *converges in* $\mathbf{S}_2 \times \mathbf{H}_2$ *as* $\varepsilon \to 0$ *to the solution* $(\nabla Y, \nabla Z)$ *of*

$$\nabla Y_t = \nabla\Phi(\hat{M}_T)\delta M_T + \int_t^T \nabla g(s,\hat{Y}_s,\hat{Z}_s)^\top (\nabla Y_s, \nabla Z_s)ds - \int_t^T \nabla Z_s dW_s \ , \ t \leq T. \tag{3.19}$$

First note that existence and uniqueness of the solution to the above BSDE in guaranteed by Assumption $(\mathbf{H^2_d})$.

Letting $\xi^\varepsilon := \varepsilon^{-1}(\Phi(M^\varepsilon_T) - \Phi(\hat{M}_T))$, one easily checks that $\varepsilon^{-1}(\delta Y^\varepsilon_s, \delta Z^\varepsilon_s)$ solves

$$\frac{\delta Y^\varepsilon_s}{\varepsilon} = \xi^\varepsilon - \int_s^T \frac{\delta Z^\varepsilon_r}{\varepsilon} dW_r + \int_s^T \left(A^{Y,\varepsilon}_r \frac{\delta Y^\varepsilon_r}{\varepsilon} + A^{Z,\varepsilon}_r \frac{\delta Z^\varepsilon_r}{\varepsilon} \right) dr,$$

where

$$A^{Y,\varepsilon}_r := \int_0^1 \partial_y g(r, \hat{Y}_r + \theta\delta Y^\varepsilon_r, \hat{Z}_r)d\theta \ \text{ and } \ A^{Z,\varepsilon}_r := \int_0^1 \partial_z g(r, Y^\varepsilon_r, \hat{Z}_r + \theta\delta Z^\varepsilon_r)d\theta.$$

In the above, $\partial_y g$ and $\partial_z g$ denotes respectively the partial gradients of g with respect to y and z, recall $(\mathbf{H^2_d})$. The Assumption $\mathbf{(H_g)}$ implies $|A^{Y,\varepsilon}| + |A^{Z,\varepsilon}| \leq K_g$.

We now set $U^\varepsilon := \varepsilon^{-1}\delta Y^\varepsilon_s - \nabla Y$, $V^\varepsilon := \varepsilon^{-1}\delta Z^\varepsilon_s - \nabla Z$ and $\zeta^\varepsilon := \xi^\varepsilon - \nabla\Phi(\hat{M}_T)\delta M$. The pair $(U^\varepsilon, V^\varepsilon)$ is an element of $\mathbf{S}_2 \times \mathbf{H}_2$ and solves

$$U^\varepsilon_s = \zeta^\varepsilon - \int_s^T V^\varepsilon_r dW_r + \int_s^T \left(A^{Y,\varepsilon}_r U^\varepsilon_r + A^{Z,\varepsilon}_r V^\varepsilon_r + R^\varepsilon_r \right) dr \ , \quad 0 \leq s \leq T \ ,$$

with

$$R_r^\varepsilon := \nabla Z_r(A_r^{Z,\varepsilon} - \partial_z g(r,\hat{Y}_r,\hat{Z}_r)) + \nabla Y_r(A_r^{Y,\varepsilon} - \partial_y g(r,\hat{Y}_r,\hat{Z}_r)) \ , \quad 0 \le r \le T \ .$$

Hence, by stability for Lipschitz BSDEs (see Proposition 6.1 in the Appendix) there exists a constant $C > 0$ (which does not depend on ε) such that

$$\|U^\varepsilon\|^2_{\mathbf{S}_2} + \|V^\varepsilon\|^2_{\mathbf{H}_2} \le C\left(\|\zeta^\varepsilon\|^2_{\mathbf{L}_2} + \|R_r^\varepsilon\|^2_{\mathbf{H}_2}\right). \tag{3.20}$$

The result of Step 1. will follow if we prove that the right-hand side of the inequality (3.20) vanishes as ε tends to zero. The convergence of $\|R_r^\varepsilon\|^2_{\mathbf{H}_2}$ to 0 follows from Assumption $(\mathbf{H^2_d})$ and the convergence of M_T^ε to M_T. As for the second term, it suffices to prove that $(Y^\varepsilon, Z^\varepsilon)_\varepsilon$ converges in $\mathbf{S}_2 \times \mathbf{H}_2$ to $(\hat{Y}, \hat{Z})$, and to appeal to $\mathbf{(H}_g\mathbf{)}$ and $(\mathbf{H^2_d})$. The latter is obtained by standard stability results, see Proposition 6.1 below, which imply the existence of a constant $C > 0$ (which does not depend on ε) such that

$$\|Y^\varepsilon - \hat{Y}\|^2_{\mathbf{S}_2} + \|Z^\varepsilon - \hat{Z}\|^2_{\mathbf{H}_2} \le C\|\Phi(M_T^\varepsilon) - \Phi(\hat{M}_T)\|^2_{\mathbf{L}_2} \ \longrightarrow_{\varepsilon\to 0} 0.$$

In the latter, the convergence follows from Lebesgue's dominated convergence Theorem and assumption $(\mathbf{H^2_d})$.

Step 2. By optimality of $(\hat{m}, \hat{\alpha})$, $Y_0^\varepsilon - m_\varepsilon l - \hat{Y}_0 + \hat{m} l \ge 0$, for any $\varepsilon > 0$. In view of Step 1, dividing by $\varepsilon > 0$ and sending $\varepsilon \to 0$ leads to

$$0 \le \nabla\Phi(\hat{M}_T)\delta M_T - l\delta m + \int_0^T \nabla g(s,\hat{Y}_s,\hat{Z}_s)^\top(\nabla Y_s, \nabla Z_s)ds - \int_0^T \nabla Z_s dW_s = \nabla Y_0 - l\delta m,$$

after possibly passing to a subsequence.

Set $\hat{L} := L^{\hat{\lambda}}$ where $\hat{\lambda} := \nabla g(\cdot, \hat{Y}, \hat{Z})$. Observe that the latter belongs to $\mathbf{\Lambda}$. For later use, also notice that

$$g(\cdot, \hat{Y}, \hat{Z}) = (\hat{\nu}, \hat{\vartheta})^\top(\hat{Y}, \hat{Z}) - \tilde{g}(\cdot, \hat{\nu}, \hat{\vartheta}), \tag{3.21}$$

see e.g. [18]. Then, it follows from (3.19) that $\hat{L}\nabla Y$ is a martingale. The previous inequality thus implies that

$$0 \le \hat{L}_0\nabla Y_0 - l\delta m = E\left[\hat{L}_T\nabla Y_T\right] - l\delta m = E\left[\hat{L}_T\delta M_T\left(\nabla\Phi(\hat{M}_T) - l/\hat{L}_T\right)\right],$$

in which we used the fact that $\hat{L}_0 = 1$ and $E[\delta M_T] = \delta m$. Since M_T can be any arbitrary random variable with values in $[0,1]$, this shows that, $\mathbb{P}-$a.s., $\hat{M}_T(\omega)$ minimizes $m \in [0,1] \mapsto \Phi(\omega, m) - ml/\hat{L}_T(\omega)$. Hence,

$$\hat{M}_T l - \hat{L}_T\Phi(\hat{M}_T) = \hat{L}_T\tilde{\Phi}(l/\hat{L}_T),$$

see e.g. [18]. Combining the above identity together with (3.21) and using Itô's Lemma leads to $l\hat{m} - \hat{Y}_0 = \mathrm{X}_0^{\hat{l},\hat{\lambda}}$. One concludes by appealing to Proposition 3.4. □

4 PDE characterization in the Markovian case

4.1 The Markovian framework

In this section, we specialize to a Markovian framework. Given $(t,x,m) \in [0,T] \times \mathbb{R}^d \times [0,1]$ and $\alpha \in \mathbf{H}_2$, we let $(X^{t,x}, M^{t,m,\alpha})$ denote the unique strong solution of

$$\begin{aligned} X_s^{t,x} &= x + \int_{t \vee s}^s b(X_r^{t,x}) dr + \int_{t \vee s}^s \sigma(X_r^{t,x}) dW_r, \ s \in [0,T] \\ M_s^{t,m,\alpha} &= m + \int_{t \vee s}^s \alpha_r dW_r. \end{aligned}$$

In the above,

$$(b, \sigma) : \mathbb{R}^d \mapsto \mathbb{R}^d \times \mathbb{R}^{d \times d} \text{ is Lipschitz continuous.} \tag{4.1}$$

Given two deterministic maps $\ell : \mathbb{R}^d \times \mathbb{R} \mapsto [0,1] \cup \{-\infty\}$ and $g : \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$, we set

$$\Psi_{t,x}(\cdot) := \ell(X_T^{t,x}, \cdot) \text{ and } g_{t,x} := g(X^{t,x}, \cdot) \ , \ (t,x) \in [0,T] \times \mathbb{R}^d.$$

We assume that, for each $x \in \mathbb{R}^d$, $y \in \mathbb{R} \mapsto \ell(x,y)$ admits a right-inverse

$$\ell^{-1}(x,m) := \inf\{y \in \mathbb{R} : \ \ell(x,y) \geq m\}$$

which is measurable and maps $[0,1]$ into itself. We then set $\Phi_{t,x}(\omega, m) := \ell^{-1}(X_T^{t,x}(\omega), m)$ for $m \in [0,1]$.

The set $\Gamma(t,x,m)$ is defined as $\Gamma(t,m)$ in (2.3) but for $\Psi_{t,x}$ and $g_{t,x}$ in place of Ψ and g. We shall also use the notations

$$\mathcal{Y}_{t,x}(m) := \operatorname{essinf} \Gamma(t,x,m)$$

and

$$Y^{t,x,m,\alpha} := \mathcal{E}_{\cdot,T}^{g_{t,x}}[\Phi_{t,x}(M_T^{t,m,\alpha})], \ \alpha \in \mathbf{A}_{t,m}.$$

All over this section, we assume that

$$(x,y,z) \in \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \mapsto g(x,y,z) \text{ is Lipschitz continuous.} \tag{4.2}$$

4.2 Link with stochastic target problems with controlled loss

We first relate $\mathcal{Y}_{t,x}(m)$ to the stochastic target with controlled loss problem of Bouchard, Elie and Touzi [3]. Given $Z \in \mathbf{H}_2$, we let $Y^{Z,(t,x,y)}$ denote the solution of

$$Y_s = y - \int_t^s g(X_r^{t,x}, Y_r, Z_r) ds + \int_t^s Z_r dW_r, \ s \in [t,T].$$

Proposition 4.1. *The map $(t,x,m) \in [0,T] \times \mathbb{R}^d \times [0,1] \mapsto \mathcal{Y}_{t,x}(m)$ admits a deterministic version and satisfies*

$$\mathcal{Y}_{t,x}(m) = v(t,x,m) := \inf\left\{y \in \mathbb{R} : \exists Z \in \mathbf{H}_2 \ s.t. \ E\left[\Psi_{t,x}(Y_T^{Z,(t,x,y)})\right] \geq m\right\}. \tag{4.3}$$

It has linear growth.

Proof. The linear growth property follows from the boundedness of Φ, (i) in Proposition 6.2 in the Appendix, (4.1) and (4.2):

$$|g(X^{t,x},0,0)| \leq C(1+\sup_{t\leq s\leq T}|X_s^{t,x}|), \text{ where } E\left[\sup_{t\leq s\leq T}|X_s^{t,x}|^2\right]^{\frac{1}{2}} \leq C(1+|x|)$$

for some $C>0$.

The rest of the proof is divided in several steps.

Step 1. *We first show that $(t,x,m)\in[0,T]\times\mathbb{R}^d\times[0,1]\mapsto\mathcal{Y}_{t,x}(m)$ is deterministic.*

We proceed as in [5]. Let H denote the Cameron-Martin space of absolutely continuous elements $h\in\Omega$ whose Radon-Nikodym derivative $\dot{h}$ is square integrable for the Lebesgue measure on $[0,T]$. One sets $H_t:=\{h\in H: h=h(\cdot\wedge t)\}$ and $\delta_h\omega=\omega+h$, $\omega\in\Omega$, $h\in H_t$. It suffices to show that $\mathcal{Y}_{t,x}(m)(\delta_h)$ is independent of $h\in H_t$, see [5, Lemma 4.1]. Let $\alpha\in\mathbf{H}_2$, $h\in H_t$, and set $\alpha^h:=\alpha(\delta_h\cdot)$. Note that

$$\alpha^h\in\mathbf{A}_{t,m} \text{ if and only if } \alpha\in\mathbf{A}_{t,m},\ h\in H_t. \tag{4.4}$$

Then, $X^{t,x}(\delta_h)=X^{t,x}$ and $M^{t,m,\alpha^h}=M^{t,m,\alpha}(\delta_h)$, and therefore

$$Y_t^{t,x,m,\alpha^h}=Y_t^{t,x,m,\alpha}(\delta_h). \tag{4.5}$$

Since, $\mathcal{Y}_{t,x}(m)\leq Y_t^{t,x,m,\alpha}$, one gets $\mathcal{Y}_{t,x}(m)(\delta_h)\leq Y_t^{t,x,m,\alpha}(\delta_h)$ for all $\alpha\in\mathbf{A}_{t,m}$, and therefore $\mathcal{Y}_{t,x}(m)(\delta_h)\leq \operatorname{ess}\inf_{\alpha\in\mathbf{A}_{t,m}}\{Y_t^{t,x,m,\alpha}(\delta_h)\}$. On the other hand, for a random variable ζ such that $\zeta\leq Y_t^{t,x,m,\alpha}(\delta_h)$, we have $\zeta(\delta_{-h})\leq Y_t^{t,x,m,\alpha}$. By arbitrariness of $\alpha\in\mathbf{A}_{t,m}$, this shows that $\zeta(\delta_{-h})\leq\mathcal{Y}_{t,x}(m)$. Hence, $\operatorname{ess}\inf_{\alpha\in\mathbf{A}_{t,m}}\{Y_t^{t,x,m,\alpha}(\delta_h)\}(\delta_{-h})\leq\mathcal{Y}_{t,x}(m)$, which, combined with the above, implies $\operatorname{ess}\inf_{\alpha\in\mathbf{A}_{t,m}}\{Y_t^{t,x,m,\alpha}(\delta_h)\}=\mathcal{Y}_{t,x}(m)(\delta_h)$. We now use the latter together with (4.5) and (4.4) to obtain

$$\mathcal{Y}_{t,x}(m)(\delta_h) = \operatorname{ess}\inf_{\alpha\in\mathbf{A}_{t,m}}\{Y_t^{t,x,m,\alpha}(\delta_h)\}=\operatorname{ess}\inf_{\alpha\in\mathbf{A}_{t,m}}Y_t^{t,x,m,\alpha^h}=\operatorname{ess}\inf_{\alpha\in\mathbf{A}_{t,m}}Y_t^{t,x,m,\alpha}=\mathcal{Y}_{t,x}(m),$$

which is the required result.

Step 2. *We now show that*

$$\mathcal{Y}_{t,x}(m)=\inf_{\alpha\in\bar{\mathbf{A}}_{t,m}}Y_t^{t,x,m,\alpha}. \tag{4.6}$$

where $\bar{\mathbf{A}}_{t,m}$ denotes the subset of elements $\alpha\in\mathbf{A}_{t,m}$ that are predictable with respect to the $\mathbb{P}$-augmented filtration generated by $W_{\cdot\vee t}-W_t$.

It follows from Step 1 above that $\mathcal{Y}_{t,x}(m)=E\left[\mathcal{Y}_{t,x}(m)\right]=E\left[\operatorname{ess}\inf_{\alpha\in\mathbf{A}_{t,m}}Y_t^{t,x,m,\alpha}\right]$. Thus, by Lemma 5.1, we can find a sequence $(\alpha_n)_n\subset\mathbf{A}_{t,m}$ such that Y_t^{t,x,m,α_n} decreases to $\mathcal{Y}_{t,x}(m)$ $\mathbb{P}-$a.s. The monotone convergence theorem then implies that $\mathcal{Y}_{t,x}(m)=\lim_n E\left[Y_t^{t,x,m,\alpha_n}\right]$ $\geq\inf_{\alpha\in\mathbf{A}_{t,m}}E\left[Y_t^{t,x,m,\alpha}\right]\geq E\left[\mathcal{Y}_{t,x}(m)\right]=\mathcal{Y}_{t,x}(m)$, in which the latter equality follows from Step 1. It remains to show that

$$\inf_{\alpha\in\mathbf{A}_{t,m}}E\left[Y_t^{t,x,m,\alpha}\right]=\inf_{\alpha\in\bar{\mathbf{A}}_{t,m}}Y_t^{t,x,m,\alpha}$$

The fact that the right-hand side is bigger than the left-hand side follows from the inclusion $\bar{\mathbf{A}}_{t,m} \subset \mathbf{A}_{t,m}$ and the fact that $Y_t^{t,x,m,\alpha}$ is deterministic for all $\alpha \in \bar{\mathbf{A}}_{t,m}$. Conversely, fix $\alpha \in \mathbf{A}_{t,m}$. It can be identified to a measurable map on the canonical space, see e.g. [22, Theorem 2.10] and [2, Lemma 1.3]. Let us denote $\alpha_\omega : \tilde{\omega} \in \Omega \mapsto \alpha(\omega_{\cdot \wedge t} + (\tilde{\omega}_{\cdot \vee t} - \tilde{\omega}_t))$. For $\omega \in \Omega$ fixed, this defines an element of $\bar{\mathbf{A}}_{t,m}$. Moreover, by independence of the increments of the Brownian motion,

$$E\left[Y_t^{t,x,m,\alpha}\right] = \int_\Omega E\left[Y_t^{t,x,m,\alpha_\omega}\right] d\mathbb{P}(\omega) = \int_\Omega Y_t^{t,x,m,\alpha_\omega} d\mathbb{P}(\omega).$$

This implies that

$$\inf_{\alpha \in \mathbf{A}_{t,m}} E\left[Y_t^{t,x,m,\alpha}\right] = \inf_{\alpha \in \mathbf{A}_{t,m}} \int_\Omega Y_t^{t,x,m,\alpha_\omega} d\mathbb{P}(\omega) \geq \int_\Omega \inf_{\alpha \in \bar{\mathbf{A}}_{t,m}} Y_t^{t,x,m,\alpha} d\mathbb{P}(\omega) = \inf_{\alpha \in \bar{\mathbf{A}}_{t,m}} Y_t^{t,x,m,\alpha},$$

which concludes the proof.

Step 3. Fix $y > v(t,x,m)$. Then, one can find $Z \in \mathbf{H}_2$ such that $E\left[\Psi(Y_T^{Z,(t,x,y)})\right] \geq m$. By the same argument as in Step 2, one can choose Z such that it is independent on $\mathcal{F}_t$. Hence, $E_t\left[\Psi(Y_T^{Z,(t,x,y)})\right] \geq m$. The process $(Y^{Z,(t,x,y)}, Z)$ is a supersolution of the BSDE with weak terminal condition BSDE$(g_{t,x}, \Psi_{t,x}, m, t)$ on $[t,T]$. Obviously it can be extended to $[0,T]$ by pasting it with the solution of BSDE$(g_{t,x} 1_{[0,t]}, Y_t^{Z,(t,x,y)})$ on $[0,t]$. This shows that $y \geq \mathcal{Y}_{t,x}(m)$. Hence, $v(t,x,m) \geq \mathcal{Y}_{t,x}(m)$. Conversely, for $\alpha \in \bar{\mathbf{A}}_{t,m}$, we can find $Z \in \mathbf{H}_2$ such that $Y^{Z,(t,x,y)} = Y^{t,x,m,\alpha}$, with $y := Y_t^{t,x,m,\alpha}$, satisfies $Y_T^{Z,(t,x,y)} = \Phi_{t,x}(M_T^{t,m,\alpha})$ and therefore $E\left[\Psi_{t,x}(Y_T^{Z,(t,x,y)})\right] \geq m$. This implies that $y \geq v(t,x,m)$. Hence, $\mathcal{Y}_{t,x}(m) \geq v(t,x,m)$ by Step 2. □

4.3 The dynamic programming equation

The PDE characterization of the value function v follows from Bouchard, Elie and Touzi [3]. For $(t,x,m,y,q,p,A) \in [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d \times d}$ and $a \in \mathbb{R}^d$, we set

$$F^a(t,x,m,y,q,p,A) := -g(x,y,p\bar{\sigma}(x,a)) - q - p^\top \bar{b}(x) - \frac{1}{2}\mathrm{Tr}[\bar{\sigma}\bar{\sigma}^\top(x,a)A]$$

with

$$\bar{b}(x) := \begin{pmatrix} b(x) \\ 0 \end{pmatrix}, \ \bar{\sigma}(x,a) := \begin{pmatrix} \sigma(x) \\ a \end{pmatrix}.$$

We then define

$$\hat{F} := \sup_{a \in \mathbb{R}^d} F^a.$$

Theorem 2.1 in [3] implies that v is a discontinuous viscosity solution of $\hat{F} = 0$. It is stated in terms of the lower- and upper-semicontinuous envelopes v_* and v^* of v defined by

$$v_*(t,x,m) := \liminf_{\substack{(t',x',m') \to (t,x,m) \\ (t',x',m') \in [0,T) \times \mathbb{R}^d \times (0,1)}} v(t',x',m')$$

and

$$v^*(t,x,m) := \limsup_{\substack{(t',x',m') \to (t,x,m) \\ (t',x',m') \in [0,T) \times \mathbb{R}^d \times (0,1)}} v(t',x',m').$$

Since $\hat{F}$ may only be lower-semicontinuous, we also need to consider its upper-semicontinuous envelope $\hat{F}^*$.

Proposition 4.2. *The function v_* is a viscosity supersolution on $[0,T)\times\mathbb{R}^d\times(0,1)$ of*

$$\hat{F}^*\varphi = 0.$$

The function v^* is a viscosity subsolution on $[0,T)\times\mathbb{R}^d\times(0,1)$ of

$$\hat{F}\varphi = 0.$$

We now discuss the boundary conditions, along the line of arguments suggested in [3] in a more abstract framework. We show that they can be fully characterized in our particular context.

We first consider the boundary as $t\to T$. In the following, we let $\widehat{\ell^{-1}}$ denote the convex envelope of $(x,m)\mapsto \ell^{-1}(x,m)$ with respect to m. We denote by $D^+\widehat{\ell^{-1}}$ its right-derivative with respect to m.

Proposition 4.3. *Assume that $\widehat{\ell^{-1}}$ and $D^+\widehat{\ell^{-1}}$ are continuous with polynomial growth. Then, $v_*(T,x,m)\geq\widehat{\ell^{-1}}(x,m)\geq v^*(T,x,m)$, for all $(x,m)\in\mathbb{R}^d\times[0,1]$.*

Proof. Let $(t_n,x_n,m_n)\to(T,x,m)$ be such that $v(t_n,x_n,m_n)\to v_*(T,x,m)$. It follows from Proposition 4.1 and Proposition 3.3 that $v(t_n,x_n,m_n)\leq\mathcal{E}_{t_n}^{g_{t_n,x_n}}\left[\widehat{\ell^{-1}}(X_T^{t_n,x_n},m_n)\right]$. Sending n to ∞ implies $v_*(T,x,m)\leq\limsup_{n\to\infty}\mathcal{E}_{t_n}^{g_{t_n,x_n}}\left[\widehat{\ell^{-1}}(X_T^{t_n,x_n},m_n)\right]=\widehat{\ell^{-1}}(x,m)$, where the latter follows from standard estimates and the continuity of $\widehat{\ell^{-1}}$. We conclude by proceeding as in [3, Proposition 3.2] and [12, Proposition 3.2]. We consider $(t_n,x_n,m_n)\to(T,x,m)$ such that $v(t_n,x_n,m_n)\to v^*(T,x,m)$, and set $y_n:=v(t_n,x_n,m_n)+n^{-1}$. Then, one can find $Z_n\in\mathbf{H}_2$ and $\alpha_n\in\mathbf{A}_{t_n,m_n}$ such that

$$\begin{aligned}
Y_T^{Z_n,(t_n,x_n,y_n)} &\geq \ell^{-1}(X_T^{t_n,x_n},M_T^{t_n,m_n,\alpha_n})\\
&\geq \widehat{\ell^{-1}}(X_T^{t_n,x_n},m)+D^+\widehat{\ell^{-1}}(X_T^{t_n,x_n},m)(M_T^{t_n,m_n,\alpha_n}-m)\\
&\geq \widehat{\ell^{-1}}(X_T^{t_n,x_n},m)+D^+\widehat{\ell^{-1}}(x,m)(M_T^{t_n,m_n,\alpha_n}-m)\\
&\quad -2\left|D^+\widehat{\ell^{-1}}(X_T^{t_n,x_n},m)-D^+\widehat{\ell^{-1}}(x,m)\right|,
\end{aligned}$$

where we used the convexity of $\widehat{\ell^{-1}}(X_T^{t_n,x_n},\cdot)$ and the fact that $M_T^{t_n,m_n,\alpha_n}$ and m have values in $[0,1]$. It follows from the uniform Lipschitz continuity assumption on g that we can find $(\rho^n,\gamma^n)_{n\geq1}\subset\mathbf{H}_2$ taking uniformly bounded values such that $L^nY^{Z_n,(t_n,x_n,y_n)}$ is a non-negative locale martingale (and therefore a super-martingale), for L^n defined as

$$L^n=e^{\int_{t_n}^{\cdot}\rho_s^nds}\exp\left(\int_{t_n}^{\cdot}\gamma_s^ndW_s-\frac{1}{2}\int_{t_n}^{\cdot}|\gamma_s^n|^2ds\right).$$

Note that $L^n\to1$ as $n\to\infty$ $\mathbb{P}-$a.s., after possibly passing to a subsequence, and that

$(L_T^n)_n$ is uniformly bounded in any $\mathbf{L}_q$, $q \geq 1$. The above implies that

$$\begin{aligned} y_n &\geq E\left[L_T^n Y_T^{Z_n,(t_n,x_n,y_n)}\right] \\ &\geq E\left[L_T^n\left(\widehat{\ell^{-1}}(X_T^{t_n,x_n},m) - 2\left|D^+\widehat{\ell^{-1}}(X_T^{t_n,x_n},m) - D^+\widehat{\ell^{-1}}(x,m)\right|\right)\right] \\ &\quad + D^+\widehat{\ell^{-1}}(x,m)E\left[M_T^{t_n,m_n,\alpha_n} - m\right] - 2D^+\widehat{\ell^{-1}}(x,m)E[|L_T^n - 1|] \\ &\geq E\left[L_T^n\left(\widehat{\ell^{-1}}(X_T^{t_n,x_n},m) - 2\left|D^+\widehat{\ell^{-1}}(X_T^{t_n,x_n},m) - D^+\widehat{\ell^{-1}}(x,m)\right|\right)\right] \\ &\quad - 2D^+\widehat{\ell^{-1}}(x,m)E\left[|L_T^n - 1|\right]. \end{aligned}$$

Passing to the limit and using standard estimates imply

$$v^*(T,x,m) = \lim_{n\to\infty} y_n \geq \widehat{\ell^{-1}}(x,m).$$

□

We now discuss the space boundary condition as $m \to \{0,1\}$.

We set $w_m : (t,x) \mapsto v(t,x,m)$ for $m \in \{0,1\}$. Since $m \in \{0,1\}$ trivially implies $\mathbf{A}_{t,m} = \{0\}$, one has

$$w_m(t,x) = \mathcal{E}_t^{g_{t,x}}[\ell^{-1}(X_T^{t,x},m)] \quad \text{for } m \in \{0,1\}.$$

Note that, ℓ being non-decreasing in y,

$$w_0 \leq v(\cdot,m) \leq w_1. \tag{4.7}$$

The following characterization is standard, see e.g. [15]. We denote by $\mathcal{L}_X$ the Dynkin operator associated to X.

Proposition 4.4. *Fix $m \in \{0,1\}$. Then, w_m is continuous on $[0,T]\times\mathbb{R}^d$ and is a viscosity solution on $[0,T)\times\mathbb{R}^d$ of*

$$-g(\cdot,\varphi,D\varphi\sigma) - \mathcal{L}_X\varphi = 0. \tag{4.8}$$

If $\ell^{-1}(\cdot,m)$ is continuous, then

$$w_m(T,\cdot) = \ell^{-1}(\cdot,m) \ \text{ on } \mathbb{R}^d. \tag{4.9}$$

We can now provide the space boundary condition.

Proposition 4.5. *Assume that the conditions of Proposition 4.3 hold. Then, $v^*(\cdot,m) = v_*(\cdot,m) = w_m$ on $[0,T]\times\mathbb{R}^d$, for $m \in \{0,1\}$.*

Proof. Let $(t_n,x_n,m_n) \to (t,x,0)$ be such that $v(t_n,x_n,m_n) \to v^*(t,x,0)$. Then, Proposition 4.1 and Proposition 3.3 imply that $v(t_n,x_n,m_n) \leq \mathcal{E}_{t_n}^{g_{t_n,x_n}}[\widehat{\ell^{-1}}(X_T^{t_n,x_n},m_n)]$. Sending $n \to \infty$ and using Proposition 6.1 below together with the continuity of $\widehat{\ell^{-1}}$ imply that $v^*(t,x,0) \leq \mathcal{E}_t^{g_{t,x}}[\widehat{\ell^{-1}}(X_T^{t,x},0)] \leq w_0(t,x)$. Recalling (4.7), this shows that $v^*(\cdot,0) = v_*(\cdot,0) = w_0$. Conversely, it follows from [3, Theorem 3.1] that $v_*(\cdot,1)$ is a viscosity supersolution of (4.8). Note that $\widehat{\ell^{-1}}(\cdot,1) = \ell^{-1}(\cdot,1)$ since ℓ^{-1} is non-decreasing.

Then, Proposition 4.3 implies that the boundary condition (4.9) with $m = 1$ holds for $v_*(\cdot, 1)$. Recall from Proposition 4.4 that w_1 is a subsolution of the same equations. Since $v_*(\cdot, 1)$ and w_1 have polynomial growth (Proposition 4.1), it follows from a standard comparison argument that $v_*(\cdot, 1) \geq w_1$, see e.g. [7]. In view of (4.7), this implies that $v^*(\cdot, 1) = v_*(\cdot, 1) = w_1$. □

Combining the above results shows that v^* is a viscosity subsolution of

$$\left\{ \begin{array}{cc} \hat{F}(\cdot, \varphi, \partial_t \varphi, D\varphi, D^2\varphi) = 0 & \text{on } [0,T) \times \mathbb{R}^d \times (0,1), \\ \varphi(\cdot, 1) = w_1 \text{ and } \varphi(\cdot, 0) = w_0 & \text{on } [0,T) \times \mathbb{R}^d, \\ \varphi(T, \cdot) = \widehat{\ell^{-1}} & \text{on } \mathbb{R}^d \times [0,1]. \end{array} \right. \tag{4.10}$$

However, v_* is only a supersolution of the same equation but with $\hat{F}^*$ in place of $\hat{F}$. Since, $\hat{F}^* \neq \hat{F}$ in general, this does not allow to characterize v as a unique viscosity solution of (4.10). One can however show that it is the biggest subsolution.

Theorem 4.1. *Let the conditions of Proposition 4.5 hold and assume that $\ell^{-1}(m, \cdot)$ is continuous for $m \in \{0, 1\}$. Then, the function v is upper-semicontinuous. Moreover, for any subsolution $\bar{v}$ of (4.10) with polynomial growth, one has $v \geq \bar{v}$ on $[0,T] \times \mathbb{R}^d \times [0,1]$.*

Proof. Step 1. Given $k \geq 1$, let us denote by $\bar{\mathbf{A}}^k_{t,m}$ the subset of elements $\alpha \in \bar{\mathbf{A}}_{t,m}$ such that $|\alpha| \leq k$ Leb$\times\mathbb{P}$ on $[0,T]$, where $\bar{\mathbf{A}}_{t,m}$ is defined in Step 2 of the proof of Proposition 4.1. We then set $v^k(t,x,m) := \inf_{\alpha \in \bar{\mathbf{A}}^k_{t,m}} Y_t^{t,x,m,\alpha}$. Let $\{\theta^\alpha, \alpha \in \bar{\mathbf{A}}_{t,m}\}$ be a family of stopping times with value in $[t,T]$ such that $\{X^{t,x}_{\theta^\alpha}, M^{t,m,\alpha}_{\theta^\alpha}, \alpha \in \bar{\mathbf{A}}_{t,m}\}$ takes values in a compact set $\mathcal{O}$ which is given and contains (x,m). One has

$$v^k(t,x,m) = \inf_{\alpha \in \bar{\mathbf{A}}^k_{t,m}} \mathcal{E}^{g_{t,x}}_{t,\theta^\alpha}\left[Y^{t,x,m,\alpha}_{\theta^\alpha} \right]$$

in which

$$Y^{t,x,m,\alpha}_{\theta^\alpha}(\omega) = Y^{\theta^\alpha, X^{t,x}_{\theta^\alpha}, M^{t,m,\alpha}_{\theta^\alpha}, \alpha}_{\theta^\alpha}(\omega) \geq \varphi(\theta^\alpha(\omega), X^{t,x}_{\theta^\alpha}(\omega), M^{t,m,\alpha}_{\theta^\alpha}(\omega))$$

for any continuous map φ lower than v^k on $\mathcal{O}$. This implies that v^k satisfies the weak dynamic programming principle, compare with [4],

$$v^k(t,x,m) \geq \inf_{\alpha \in \bar{\mathbf{A}}^k_{t,m}} \mathcal{E}^{g_{t,x}}_{t,\theta^\alpha}\left[\varphi(\theta^\alpha, X^{t,x}_{\theta^\alpha}, M^{t,m,\alpha}_{\theta^\alpha}) \right]$$

for any smooth function $\varphi \leq v^k$ on $\mathcal{O}$. Similar arguments as in [4] then implies that the lower semicontinuous envelope v^k_* of v^k is a viscosity supersolution of

$$\hat{F}_k(\cdot, \varphi, \partial_t \varphi, D\varphi, D^2\varphi) = 0 \text{ on } [0,T) \times \mathbb{R}^d \times (0,1) \tag{4.11}$$

where $\hat{F}_k := \sup_{|a| \leq k} F^a$.

Moreover, $v^k \geq v$ by Proposition 4.1 and Relation (4.6). It then follows from Propositions 4.3, 4.4 and 4.5 that v^k_* is a supersolution of

$$\begin{array}{c} \varphi(\cdot, 1) = w_1 \text{ and } \varphi(\cdot, 0) = w_0 \quad \text{on } [0,T) \times \mathbb{R}^d, \\ \varphi(T, \cdot) = \widehat{\ell^{-1}}(x,m) \quad \text{on } \mathbb{R}^d \times [0,1]. \end{array} \tag{4.12}$$

Step 2. Let $\bar{v}$ be a viscosity subsolution of (4.10) with polynomial growth. Then, it is a subsolution of (4.11)-(4.12). In view of Step 1, standard comparison results, see e.g. [7], imply that $v_*^k \geq \bar{v}$, the fact that v^k has polynomial growth falling from the same considerations as in the proof of Proposition 4.1.

Step 3. It remains to prove that $v^k \downarrow v$ pointwise. Clearly, $\{M_T^{t,m,\alpha},\ \alpha \in \mathbf{A}_{t,m}^k\}$ is dense in $\mathbf{L}_2$ in $\{M_T^{t,m,\alpha},\ \alpha \in \mathbf{A}_{t,m}\}$. The required result then follows from Proposition 6.1 in the Appendix, Proposition 4.1 and Relation (4.6).

□

5 Proof of Theorem 2.1

In all this section, we use the notations introduced at the beginning of Section 2.2. The first main result provides a dynamic programming principle for the family $\{\mathcal{Y}_\tau^\alpha,\ \tau \in \mathcal{T}, \alpha \in \mathbf{A}_0\}$.

Proposition 5.1. *For all $(\tau_1, \tau_2, \alpha) \in \mathcal{T} \times \mathcal{T} \times \mathbf{A}_0$ such that $\tau_1 \leq \tau_2$, we have*

$$\mathcal{Y}_{\tau_1}^\alpha \;=\; \operatorname{ess\,inf}_{\alpha' \in \mathbf{A}_{\tau_1}^\alpha} \mathcal{E}_{\tau_1,\tau_2}^g[\mathcal{Y}_{\tau_2}^{\alpha'}]\,.$$

Proof. We prove the two corresponding inequalities separately.
Step 1. $\mathcal{Y}_{\tau_1}^\alpha \geq \operatorname{ess\,inf}_{\alpha' \in \mathbf{A}_{\tau_1}^\alpha} \mathcal{E}_{\tau_1,\tau_2}^g\left[\mathcal{Y}_{\tau_2}^{\alpha'}\right]$.
It follows from Lemma 5.1 below that there exists $(\alpha^n)_n$ in $\mathbf{A}_{\tau_1}^\alpha$ such that the sequence $(\mathcal{E}_{\tau_1,T}^g[\Phi(M_T^{\alpha^n})])_n$ is non-increasing and

$$\lim_{n\to\infty} \mathcal{E}_{\tau_1,T}^g[\Phi(M_T^{\alpha^n})] \;=\; \mathcal{Y}_{\tau_1}^\alpha, \quad \mathbb{P}-\text{a.s.} \tag{5.1}$$

Since $\alpha^n \in \mathbf{A}_{\tau_2}^{\alpha^n}$ for every $n \geq 1$, we deduce that

$$\mathcal{Y}_{\tau_2}^{\alpha^n} \;\leq\; \mathcal{E}_{\tau_2,T}^g[\Phi(M_T^{\alpha^n})]\,.$$

By comparison for BSDEs with Lipschitz continuous drivers on the time interval $[\tau_1, \tau_2]$, this implies

$$\mathcal{E}_{\tau_1,\tau_2}^g[\mathcal{Y}_{\tau_2}^{\alpha^n}] \;\leq\; \mathcal{E}_{\tau_1,\tau_2}^g[\mathcal{E}_{\tau_2,T}^g[\Phi(M_T^{\alpha^n})]] = \mathcal{E}_{\tau_1,T}^g[\Phi(M_T^{\alpha^n})]\,,$$

leading to

$$\operatorname{ess\,inf}_{\alpha' \in \mathbf{A}_{\tau_1}^\alpha} \mathcal{E}_{\tau_1,\tau_2}^g\left[\mathcal{Y}_{\tau_2}^{\alpha'}\right] \;\leq\; \mathcal{E}_{\tau_1,T}^g[\Phi(M_T^{\alpha^n})]\,,$$

Letting n go to infinity in the above inequality, (5.1) provides directly

$$\operatorname{ess\,inf}_{\alpha' \in \mathbf{A}_{\tau_1}^\alpha} \mathcal{E}_{\tau_1,\tau_2}^g\left[\mathcal{Y}_{\tau_2}^{\alpha'}\right] \;\leq\; \mathcal{Y}_{\tau_1}^\alpha\,.$$

Step 2. $\mathcal{Y}_{\tau_1}^\alpha \leq \operatorname{ess\,inf}_{\alpha' \in \mathbf{A}_{\tau_1}^\alpha} \mathcal{E}_{\tau_1,\tau_2}^g\left[\mathcal{Y}_{\tau_2}^{\alpha'}\right]$.

Fix α' in $\mathbf{A}^{\alpha}_{\tau_1}$. Lemma 5.1 below ensures the existence of a sequence $(\alpha'_n)_n$ in $\mathbf{A}^{\alpha'}_{\tau_2}$ such that $(\mathcal{E}^g_{\tau_2,T}[\Phi(M_T^{\alpha'^n})])_n$ is non-increasing and

$$\lim_{n\to\infty} \mathcal{E}^g_{\tau_2,T}[\Phi(M_T^{\alpha'_n})] = \mathcal{Y}^{\alpha'}_{\tau_2}, \quad \mathbb{P}-\text{a.s.}$$

In view of Remark 2.1, the convergence holds in $\mathbf{L}_2$ as well. Thus the stability result of Proposition 6.1 below indicates that $\mathcal{E}^g_{\tau_1,T}[\Phi(M_T^{\alpha'_n})]$ converges to $\mathcal{E}^g_{\tau_1,\tau_2}[\mathcal{Y}^{\alpha'}_{\tau_2}]$ in $\mathbf{L}_2$. In addition, $\alpha'_n \in \mathbf{A}^{\alpha'}_{\tau_2} \subset \mathbf{A}^{\alpha}_{\tau_1}$ by construction. Combining the above leads to

$$\mathcal{E}^g_{\tau_1,\tau_2}[\mathcal{Y}^{\alpha'}_{\tau_2}] \;=\; \lim_{n\to\infty} \mathcal{E}^g_{\tau_1,T}[\Phi(M_T^{\alpha'_n})] \;\geq\; \mathcal{Y}^{\alpha}_{\tau_1}\,.$$

The arbitrariness of $\alpha' \in \mathbf{A}^{\alpha}_{\tau_1}$ allows one to conclude

$$\operatorname{ess}\inf_{\alpha'\in\mathbf{A}^{\alpha}_{\tau_1}} \mathcal{E}^g_{\tau_1,\tau_2}[\mathcal{Y}^{\alpha'}_{\tau_2}] \;\geq\; \mathcal{Y}^{\alpha}_{\tau_1}\,.$$

□

Lemma 5.1. *Fix $\theta,\tau \in \mathcal{T}$, with $\theta \geq \tau$, $\mu \in \mathbf{L}_0([0,1],\mathcal{F}_\tau)$ and $\alpha \in \mathbf{A}_{\tau,\mu}$. Then, there exists a sequence $(\alpha'_n) \subset \mathbf{A}^{\theta,\alpha}_{\tau,\mu}$:=$\{\alpha' \in \mathbf{A}_{\tau,\mu},\ \alpha'\mathbf{1}_{[0,\theta)} = \alpha\mathbf{1}_{[0,\theta)}\}$such that $\lim_n \downarrow \mathcal{E}^g_{\theta,T}[\Phi(M_T^{\tau,\mu,\alpha'_n})] = \mathcal{Y}^{\alpha}_{\theta}(M_\theta^{\tau,\mu,\alpha})$ $\mathbb{P}-$ a.s.*

Proof. It suffices to show that the family $\{J(\alpha') := \mathcal{E}^g_{\theta,T}[\Phi(M_T^{\tau,\mu,\alpha'})],\ \alpha' \in \mathbf{A}^{\theta,\alpha}_{\tau,\mu}\}$ is directed downward, see e.g. [13]. Fix α'_1,α'_2 in $\mathbf{A}^{\theta,\alpha}_{\tau,\mu}$ and set

$$\tilde{\alpha}' := \alpha\mathbf{1}_{[0,\theta)} + \mathbf{1}_{[\theta,T]}(\alpha'_1\mathbf{1}_A + \alpha'_2\mathbf{1}_{A^c})$$

where $A := \{J(\alpha'_1) \leq J(\alpha'_2)\} \in \mathcal{F}_\theta$, so that $\tilde{\alpha}' \in \mathbf{A}^{\theta,\alpha}_{\tau,\mu}$ and

$$J(\tilde{\alpha}') = \mathcal{E}^g_{\theta,T}[\Phi(M_T^{\tau,\mu,\alpha'_1})\mathbf{1}_A + \Phi(M_T^{\tau,\mu,\alpha'_2})\mathbf{1}_{A^c}] = \min\{J(\alpha'_1), J(\alpha'_2)\}.$$

□

We now observe that the family $(\mathcal{Y}^\alpha)_{\alpha\in\mathbf{H}_2}$ is undistinguishable from a làdlàg process[2], $(\bar{\mathcal{Y}}^\alpha)_{\alpha\in\mathbf{H}_2}$ hereafter, which also satisfies the preceding dynamic programming principle. If in addition Φ is assumed to be continuous, the process $(\mathcal{Y}^\alpha)_{\alpha\in\mathbf{H}_2}$ is even indistinguishable from a càdlàg[3] process.

Proposition 5.2. *Fix $\alpha \in \mathbf{A}_0$. Then, $\mathcal{Y}^\alpha$ is indistinguishable from a làdlàg process. Besides, if $m \in [0,1] \mapsto \Phi(\omega,m)$ is continuous for $\mathbb{P}$-a.e. $\omega \in \Omega$, then $\mathcal{Y}^\alpha$ is indistinguishable from a càdlàg process.*

Proof. Fix $\alpha \in \mathbf{A}_0$. Proposition 5.1 and Remark 2.1 imply that $-\mathcal{Y}^\alpha$ is a $-g(-\cdot)$-supermartingale in the sense of [6] (a g-submartingale in the sense of [16]). It follows from the non-linear up-crossing Lemma, see [6, Theorem 6][4], that the following limits

$$\lim_{s\in\mathbb{Q}\cap(t,T]\downarrow t} \mathcal{Y}^\alpha_s \text{ and } \lim_{s\in\mathbb{Q}\cap[0,t)\uparrow t} \mathcal{Y}^\alpha_s$$

[2]left and right-limited according to the french celebrated acronym

[3]right-continuous and left-limited

[4]Note that [6, Theorem 6] restricts to positive g-supermartingales. However, the proof can be reproduced without difficulty under the integrability condition of Remark 2.1. In addition, [6, Theorem 6] implies that $E^{\mathbb{Q}}[D^b_a(\mathcal{Y}^\alpha,n)] \leq \mathcal{Y}^\alpha_0 \wedge b \leq b$, where $D^b_a(\mathcal{Y}^\alpha,n)$ denotes the number of down crossing of $\mathcal{Y}^\alpha$ from an interval $[a,b]$ on a discrete time-grid $0 = t_0 \leq t_1 \leq \cdots \leq t_n = T$ and $\mathbb{Q}$ is a particular measure absolutely continuous with respect to $\mathbb{P}$. To conclude, it is enough to reproduce the proof of [8, Chapter VI Theorem (2) point 1)].

are well-defined for every t in $[0,T]$, $\mathbb{P}-$a.s., so is the process

$$\bar{\mathcal{Y}}^{\alpha}_t := \lim_{s\in\mathbb{Q}\cap(t,T]\downarrow t} \mathcal{Y}^{\alpha}_s\,, \qquad t\in[0,T]\,.$$

Hence, $\mathcal{Y}^{\alpha}$ is undistinguishable from a làdlàg process.

Besides, $\bar{\mathcal{Y}}^{\alpha}$ is by definition càd. Assuming that Φ is continuous, we will prove that, for every stopping time τ, it holds that:

$$\bar{\mathcal{Y}}^{\alpha}_{\tau} = \operatorname{ess\,inf}_{\alpha'\in\mathbf{A}^{\alpha}_{\tau}} \mathcal{E}^{g}_{\tau,T}\left[\Phi(M^{\alpha'}_T)\right] (= \mathcal{Y}^{\alpha}_{\tau}) \qquad \mathbb{P}-\text{a.s.} \tag{5.2}$$

By [8, Chapter IV. (86), p. 220], the relation (5.2) entails that $\mathcal{Y}^{\alpha}$ and $\bar{\mathcal{Y}}^{\alpha}$ are undistinguishable showing that $\mathcal{Y}^{\alpha}$ is undistinguishable from a càdlàg process. The rest of the proof is devoted to prove (5.2).

For this purpose, let us introduce $(\tau_n)_n$, a decreasing sequence of stopping times with values in $[0,T]\cap\mathbb{Q}$ such that $\tau\le\tau_n\le\tau+n^{-1}$ and $\bar{\mathcal{Y}}^{\alpha}_{\tau}=\lim_{n\to\infty}\mathcal{Y}^{\alpha}_{\tau_n}$.

Step 1. $\bar{\mathcal{Y}}^{\alpha}_{\tau} \le \operatorname{ess\,inf}_{\alpha'\in\mathbf{A}^{\alpha}_{\tau}} \mathcal{E}^{g}_{\tau,T}\left[\Phi(M^{\alpha'}_T)\right]$.

a. Fix $\alpha'\in\mathbf{A}^{\alpha}_{\tau}$ and set

$$\lambda_n := \left(\frac{M^{\alpha}_{\tau_n}}{M^{\alpha'}_{\tau_n}} \wedge \frac{1-M^{\alpha}_{\tau_n}}{1-M^{\alpha'}_{\tau_n}}\right)\mathbf{1}_{\{M^{\alpha}_{\tau_n}\notin\{0,1\}\}} \in [0,1],$$

with the convention $a/0=\infty$ for $a>0$. Using the fact that $M^{\alpha'}_{\tau_n}+\int_{\tau_n}^{T}\alpha'_s dW_s = M^{\alpha'}_T\in[0,1]$, direct computations lead to

$$0\le M^{\alpha}_{\tau_n}-\lambda_n M^{\alpha'}_{\tau_n} \le M^{\alpha}_{\tau_n}+\lambda_n\int_{\tau_n}^{T}\alpha'_s dW_s \le M^{\alpha}_{\tau_n}+\lambda_n(1-M^{\alpha'}_{\tau_n})\le 1.$$

We set $\alpha'_n := \alpha\mathbf{1}_{[0,\tau_n)}+\lambda_n\alpha'\mathbf{1}_{[\tau_n,T]}$. The above implies that α'_n belongs to $\mathbf{A}^{\alpha}_{\tau_n}$.

b. Now we prove that $M^{\alpha'_n}_T$ converges $M^{\alpha'}_T$ in $\mathbf{L}_2$ as n goes to infinity, possibly up to a subsequence. Since both have norms bounded by 1, it suffices to show the $\mathbb{P}-$a.s. convergence, possibly up to a subsequence. To see this, first note that

$$M^{\alpha'_n}_T - M^{\alpha'}_T = M^{\alpha}_{\tau_n}-M^{\alpha'}_{\tau_n}+\int_{\tau_n}^{T}(\lambda_n-1)\alpha'_s dW_s,$$

from which we deduce that

$$M^{\alpha'_n}_T - M^{\alpha'}_T = M^{\alpha}_{\tau_n}-M^{\alpha'}_{\tau_n}+(\lambda_n-\mathbf{1}_{\{M^{\alpha}_{\tau_n}\notin\{0,1\}\}})\int_{\tau_n}^{T}\alpha'_s dW_s - \mathbf{1}_{\{M^{\alpha}_{\tau_n}\in\{0,1\}\}}\int_{\tau_n}^{T}\alpha'_s dW_s.$$

Since $\tau_n\to\tau$ $\mathbb{P}-$a.s. and $\alpha'=\alpha$ on $[\![0,\tau]\!]$, the above construction implies that $\lim_{n\to\infty}M^{\alpha}_{\tau_n}-M^{\alpha'}_{\tau_n}=0$ $\mathbb{P}-$a.s. and $\lim_{n\to\infty}\lambda_n=\lim_{n\to\infty}\mathbf{1}_{\{M^{\alpha}_{\tau_n}\notin\{0,1\}\}}$ $\mathbb{P}-$a.s. It thus only remains to prove that $\mathbf{1}_{\{M^{\alpha}_{\tau_n}\in\{0,1\}\}}\int_{\tau_n}^{T}\alpha'_s dW_s\to 0$ $\mathbb{P}-$a.s. First note that $\alpha'\mathbf{1}_{[\tau_n,T]}=0$ on $\{M^{\alpha'}_{\tau_n}\in\{0,1\}\}$.

This follows from the martingale property of this process with values in $[0,1]$. Hence, it suffices to consider $\mathbf{1}_{\{M^{\alpha'}_{\tau_n} \neq M^{\alpha}_{\tau_n} \in \{0,1\}\}} \int_{\tau_n}^{T} \alpha'_s dW_s$. But, since $M^{\alpha'}_{\tau} = M^{\alpha}_{\tau}$,

$$\mathbb{P}[M^{\alpha'}_{\tau_n} \neq M^{\alpha}_{\tau_n} \in \{0,1\}] \leq \mathbb{P}[M^{\alpha'}_{\tau_n} \neq M^{\alpha}_{\tau_n}] = \mathbb{P}\left[\left|\int_{\tau}^{\tau_n} (\alpha_s - \alpha'_s) dW_s\right| > 0\right] \to_{n\to\infty} 0.$$

c. Now, since Φ is continuous and $M^{\alpha'_n}_T \in \mathbf{L}_0([0,1])$, we get that $\Phi(M^{\alpha'_n}_T) \to \Phi(M^{\alpha'}_T)$ in $\mathbf{L}_2$, after possibly passing to a subsequence. The stability property for Lipschitz BSDEs given in Proposition 6.1 implies that

$$\left\|\mathcal{E}^g_{\tau^n,T}\left[\Phi(M^{\alpha'_n}_T)\right] - \mathcal{E}^g_{\tau^n,T}\left[\Phi(M^{\alpha'}_T)\right]\right\|_{\mathbf{L}_2} \to_{n\to\infty} 0\,. \tag{5.3}$$

On the other hand, the bound of Remark 2.1 implies that

$$\left\|\mathcal{E}^g_{\tau^n,T}\left[\Phi(M^{\alpha'}_T)\right] - \mathcal{E}^g_{\tau,T}\left[\Phi(M^{\alpha'}_T)\right]\right\|_{\mathbf{L}_2} \to_{n\to\infty} 0\,, \tag{5.4}$$

by Lebesgue's dominated convergence Theorem and by continuity of the process $\mathcal{E}^g_{\cdot,T}\left[\Phi(M^{\alpha'}_T)\right]$. Combining (5.3) and (5.4) leads to

$$\bar{\mathcal{Y}}^{\alpha}_{\tau} = \lim_{n\to\infty} \mathcal{Y}^{\alpha}_{\tau^n} \leq \lim_{n\to\infty} \mathcal{E}^g_{\tau^n,T}\left[\Phi(M^{\alpha'_n}_T)\right] = \mathcal{E}^g_{\tau,T}\left[\Phi(M^{\alpha'}_T)\right].$$

We conclude by arbitrariness of $\alpha' \in \mathbf{A}^{\alpha}_{\tau}$.

Step 2. $\bar{\mathcal{Y}}^{\alpha}_{\tau} \geq \operatorname{ess\,inf}_{\alpha' \in \mathbf{A}^{\alpha}_{\tau}} \mathcal{E}^g_{\tau,T}\left[\Phi(M^{\alpha'}_T)\right]$.

Applying on $[\tau, \tau^n]$ the stability result of Proposition 6.1 for the BSDEs with parameters $(\bar{\mathcal{Y}}^{\alpha}_{\tau}, 0)$ and $(\mathcal{Y}^{\alpha}_{\tau^n}, g\mathbf{1}_{[0,\tau^n)})$, we get

$$\begin{aligned}
\left\|\bar{\mathcal{Y}}^{\alpha}_{\tau} - \mathcal{E}^g_{\tau,\tau^n}\left[\mathcal{Y}^{\alpha}_{\tau^n}\right]\right\|_{\mathbf{L}_2} &\leq C\left(\left\|\bar{\mathcal{Y}}^{\alpha}_{\tau} - \mathcal{Y}^{\alpha}_{\tau^n}\right\|_{\mathbf{L}_2} + E\left[\int_{\tau}^{\tau^n} \left|g(s, \bar{\mathcal{Y}}^{\alpha}_{\tau}, 0)\right|^2 ds\right]\right) \\
&\leq C\left\|\bar{\mathcal{Y}}^{\alpha}_{\tau} - \mathcal{Y}^{\alpha}_{\tau^n}\right\|_{\mathbf{L}_2} + \frac{C}{n}\,, \qquad n \in \mathbb{N}\,,
\end{aligned}$$

for some $C > 0$, since the bound of Remark 2.1 holds for $\bar{\mathcal{Y}}^{\alpha}_{\tau}$, recall that Assumption **$(\mathbf{H}_g)$** is in force. Therefore, $\mathcal{E}^g_{\tau,\tau^n}[\mathcal{Y}^{\alpha}_{\tau^n}]$ converges to $\bar{\mathcal{Y}}^{\alpha}_{\tau}$ as n goes to infinity. Proposition 5.1 implies $\mathcal{E}^g_{\tau,\tau^n}\left[\mathcal{Y}^{\alpha}_{\tau^n}\right] \geq \mathcal{Y}^{\alpha}_{\tau}$. Passing to the limit leads to the required inequality: $\bar{\mathcal{Y}}^{\alpha}_{\tau} \geq \mathcal{Y}^{\alpha}_{\tau} = \operatorname{ess\,inf}_{\alpha' \in \mathbf{A}^{\alpha}_{\tau}} \mathcal{E}^g_{\tau,T}\left[\Phi(M^{\alpha'}_T)\right]$. □

In the rest of this section, we complete the proof of Theorem 2.1.

Proof of Theorem 2.1. Items (i) and (ii) are already proved in Proposition 5.1 and Proposition 5.2, it remains to prove (iii) and (iv). For $\alpha \in \mathbf{A}_0$, it follows from Proposition 5.1, Proposition 5.2 and standard comparison results for BSDEs that $\mathcal{Y}^{\alpha}$ is a càdlàg strong g-submartingale in the sense of [16]. Hence, the existence of a process $(\mathcal{Z}^{\alpha}, \mathcal{K}^{\alpha}) \in \mathbf{H}_2 \times \mathbf{K}_2$ such that (2.9) holds follows from [16, Theorem 3.3]. We now verify successively that the family $(\mathcal{Y}^{\alpha}, \mathcal{Z}^{\alpha}, \mathcal{K}^{\alpha})_{\alpha \in \mathbf{H}_2}$ satisfies (2.8), (2.10), (2.11) and the uniqueness of solution for (2.8)-(2.9)-(2.10)-(2.11).

The bound (2.8) follows directly from Remark 2.1 and the representation Theorem 3.3 in [16], note that the driver function g does not depend on $\alpha \in \mathbf{A}_0$.

Step 1. The irrelevance of future property (2.11)
For $(\alpha, \tau) \in \mathbf{A}_0 \times \mathcal{T}$, observe that $\mathbf{A}_\cdot^{\alpha'} = \mathbf{A}_\cdot^{\alpha}$ on $[0, \tau]$ when $\alpha' \in \mathbf{A}_\tau^\alpha$. The definition of $\mathcal{Y}$ thus implies that $\mathcal{Y}^\alpha \mathbf{1}_{[0,\tau]} = \mathcal{Y}^{\alpha'} \mathbf{1}_{[0,\tau]}$ for $\alpha' \in \mathbf{A}_\tau^\alpha$. Hence (2.11) follows from the uniqueness of the representation provided in [16, Theorem 3.3].

Step 2. The minimality property (2.10)
We follow the arguments in the proof [21, Theorem 4.6]. We fix $(\alpha, \tau_1, \tau_2) \in \mathbf{H}_2 \times \mathcal{T} \times \mathcal{T}$ such that $\tau_1 \le \tau_2$. For any $\alpha' \in \mathbf{A}_{\tau_1}^\alpha$, we denote by $(Y^{\alpha'}, Z^{\alpha'})$ the solution of the classical BSDE

$$Y_t^{\alpha'} \;=\; \Phi(M_T^{\alpha'}) + \int_t^T g(s, Y_s^{\alpha'}, Z_s^{\alpha'}) ds - \int_t^T Z_s^{\alpha'} dW_s \,, \qquad 0 \le t \le T \,.$$

Let $L^{\alpha'}$ be the process whose dynamics is given by

$$L_t^{\alpha'} \;=\; \exp\left(\int_{\tau_1}^t \Lambda_s^z dW_s + \int_{\tau_1}^t \left(\Lambda_s^y - \frac{|\Lambda_s^z|^2}{2} \right) ds \right) \,, \qquad \tau_1 \le t \le T \,,$$

where (Λ^y, Λ^z) is the linearization process given by

$$\begin{aligned} \Lambda^y \;&:=\; \frac{g(\mathcal{Y}_s^{\alpha'}, \mathcal{Z}_s^{\alpha'}) - g(Y_s^{\alpha'}, Z_s^{\alpha'})}{\mathcal{Y}_s^{\alpha'} - Y_s^{\alpha'}} \mathbf{1}_{\{\mathcal{Y}^{\alpha'} \neq Y^{\alpha'}\}} \,, \\ \Lambda^z \;&:=\; \frac{g(\mathcal{Y}_s^{\alpha'}, \mathcal{Z}_s^{\alpha'}) - g(Y_s^{\alpha'}, Z_s^{\alpha'})}{|\mathcal{Z}_s^{\alpha'} - Z_s^{\alpha'}|^2} (\mathcal{Z}^{\alpha'} - Z^{\alpha'}) \mathbf{1}_{\{\mathcal{Z}^{\alpha'} \neq Z^{\alpha'}\}} \,. \end{aligned}$$

This linearization procedure implies that $Y_{\tau_1}^{\alpha'} - \mathcal{Y}_{\tau_1}^{\alpha'}$ rewrites as

$$\begin{aligned} Y_{\tau_1}^{\alpha'} - \mathcal{Y}_{\tau_1}^{\alpha'} \;&=\; E_{\tau_1}\left[L_{\tau_2}^{\alpha'} (Y_{\tau_2}^{\alpha'} - \mathcal{Y}_{\tau_2}^{\alpha'}) \right] + E_{\tau_1}\left[\int_{\tau_1}^{\tau_2} L_s^{\alpha'} d\mathcal{K}_s^{\alpha'} \right] \\ &\ge\; E_{\tau_1}\left[(\mathcal{K}_{\tau_2}^{\alpha'} - \mathcal{K}_{\tau_1}^{\alpha'}) \inf_{[\tau_1, \tau_2]} L^{\alpha'} \right] \,, \end{aligned} \tag{5.5}$$

where we used the fact that $Y^\alpha - \mathcal{Y}^\alpha \ge 0$. Using Hölder inequality, this implies

$$\begin{aligned} E_{\tau_1}\left[(\mathcal{K}_{\tau_2}^{\alpha'} - \mathcal{K}_{\tau_1}^{\alpha'}) \right]^3 \;&\le\; E_{\tau_1}\left[(\mathcal{K}_{\tau_2}^{\alpha'} - \mathcal{K}_{\tau_1}^{\alpha'}) \inf_{[\tau_1, \tau_2]} L^{\alpha'} \right] E_{\tau_1}\left[\sup_{[\tau_1, \tau_2]} (1/L^{\alpha'}) \right] E_{\tau_1}\left[(\mathcal{K}_{\tau_2}^{\alpha'} - \mathcal{K}_{\tau_1}^{\alpha'})^2 \right] \\ &\le\; C\, E_{\tau_1}\left[(\mathcal{K}_{\tau_2}^{\alpha'} - \mathcal{K}_{\tau_1}^{\alpha'})^2 \right] (Y_{\tau_1}^{\alpha'} - \mathcal{Y}_{\tau_1}^{\alpha'}) \,, \end{aligned}$$

for some $C > 0$ that depends on the uniform bounds on (Λ^y, Λ^z), recall $\mathbf{(H}_g\mathbf{)}$. Hence, the estimate (2.8) together with the monotonicity of $\mathcal{K}$ implies

$$0 \;\le\; E_{\tau_1}\left[(\mathcal{K}_{\tau_2}^{\alpha'} - \mathcal{K}_{\tau_1}^{\alpha'}) \right] \;\le\; C \eta'_{\tau_1} \;\; (Y_{\tau_1}^{\alpha'} - \mathcal{Y}_{\tau_1}^{\alpha'})^{1/3} \,, \qquad \alpha' \in \mathbf{A}_{\tau_1}^\alpha \,, \tag{5.6}$$

where

$$\eta'_{\tau_1} \;:=\; \operatorname*{ess\,sup}_{\bar\alpha \in \mathbf{A}_{\tau_1}^\alpha} E_{\tau_1}\left[(\mathcal{K}_{\tau_2}^{\bar\alpha} - \mathcal{K}_{\tau_1}^{\bar\alpha})^2 \right]^{1/3} .$$

By the same arguments as in Lemma 5.1, we can find a sequence $(\alpha'_n)_n \subset \mathbf{A}^\alpha_{\tau_1}$ such that

$$\eta'_{\tau_1} \;=\; \lim_{n\to\infty} \uparrow E_{\tau_1}\left[(\mathcal{K}^{\alpha'_n}_{\tau_2} - \mathcal{K}^{\alpha'_n}_{\tau_1})^2\right]^{1/3} .$$

The monotone convergence Theorem together with Jensen's inequality and Relation (2.8) imply that

$$E[\eta'_{\tau_1}] \;=\; \lim_{n\to\infty} \uparrow E\left[(\mathcal{K}^{\alpha'_n}_{\tau_2} - \mathcal{K}^{\alpha'_n}_{\tau_1})^2\right]^{1/3} \;<\; \infty .$$

Since η'_{τ_1} is in addition non-negative, it is a.s. bounded. Hence, combining (2.11) and (5.6), we obtain

$$0 \;\le\; E_{\tau_1}\left[\mathcal{K}^{\alpha'}_{\tau_2}\right] - \mathcal{K}^{\alpha'}_{\tau_1} \;\le\; C\,(\mathcal{E}^g_{\tau_1,\tau_2}[Y^{\alpha'}_{\tau_2}] - \mathcal{Y}^\alpha_{\tau_1})^{1/3} = C\,(\mathcal{E}^g_{\tau_1}[\Phi(M^{\alpha'}_T)] - \mathcal{Y}^\alpha_{\tau_1})^{1/3} \;,\; \alpha' \in \mathbf{A}^\alpha_{\tau_1} .$$

Taking the essential infimum in the above inequality and appealing to (2.6) leads to (2.10).

Step 3. The uniqueness property for (2.8)-(2.9)-(2.10)-(2.11)
Let us now consider a family $(\tilde{Y}^\alpha, \tilde{Z}^\alpha, \tilde{K}^\alpha)_{\alpha\in\mathbf{A}_0}$ of $\mathbf{S}_2 \times \mathbf{H}_2 \times \mathbf{K}_2$ satisfying (2.8)-(2.9)-(2.10)-(2.11). Then, (2.6) together with (2.9)-(2.11) applied to $(\tilde{Y}^\alpha, \tilde{Z}^\alpha, \tilde{K}^\alpha)_{\alpha\in\mathbf{A}_0}$ imply via a direct comparison argument that

$$\mathcal{Y}^\alpha_t \;=\; \operatorname{ess\,inf}_{\alpha'\in\mathbf{A}^\alpha_t} \mathcal{E}^g_t[\Phi(M^{\alpha'}_T)] \ge \tilde{Y}^\alpha_t \;, \qquad \alpha \in \mathbf{A}_0 \;, \quad 0 \le t \le T \;. \tag{5.7}$$

On the other hand, following the exact same line of arguments as the one developed in Step 2 in order to derive (5.5), one easily shows that there exists a $\mathbf{S}_2$-uniformly bounded family of processes $(\tilde{L}^\alpha)_{\alpha\in\mathbf{A}_0}$ such that

$$\mathcal{E}^g_t[\Phi(M^\alpha_T)] - \tilde{Y}^\alpha_t \;=\; E_t\left[\int_t^T \tilde{L}^\alpha_s d\tilde{K}^\alpha_s\right] \;\le\; CE_t\left[|\tilde{K}^\alpha_T - \tilde{K}^\alpha_t|^2\right]^{1/2} \;, \quad \alpha \in \mathbf{A}_0 \quad 0 \le t \le T \;,$$

for some $C > 0$.

Now observe that (2.10), applied to $\tilde{K}^\alpha$, and the same arguments as in Lemma 5.1 provide the existence of $(\hat{\alpha}^n)_n \subset \mathbf{A}^\alpha_t$ such that $E_t[\tilde{K}^{\hat{\alpha}^n}_T - \tilde{K}^\alpha_t] \to 0$, $\mathbb{P}-$a.s. Hence, (2.8) ensures that $E_t[|\tilde{K}^{\hat{\alpha}^n}_T - \tilde{K}^\alpha_t|^2] \to 0$. Since (2.11) implies $(\tilde{Y}^{\hat{\alpha}^n}_t, \tilde{K}^{\hat{\alpha}^n}_t) = (\tilde{Y}^\alpha_t, \tilde{K}^\alpha_t)$ for $n \in \mathbb{N}$, we deduce

$$\mathcal{E}^g_t[\Phi(M^{\hat{\alpha}^n}_T)] - \tilde{Y}^\alpha_t \;\le\; C\,E_t\left[|\tilde{K}^{\hat{\alpha}^n}_T - \tilde{K}^\alpha_t|^2\right]^{1/2} \;\to_{n\to\infty}\; 0 \;.$$

Combined with (5.7), this shows that

$$\tilde{Y}^\alpha_t \;=\; \operatorname{ess\,inf}_{\alpha'\in\mathbf{A}^\alpha_t} \mathcal{E}^g_t[\Phi(M^{\alpha'}_T)] \;=\; \mathcal{Y}^\alpha_t \;, \qquad \alpha \in \mathbf{H}_2 \;, \quad 0 \le t \le T \;.$$

The fact that $(\tilde{Z}^\alpha, \tilde{K}^\alpha)_{\alpha\in\mathbf{A}_0} = (\mathcal{Z}^\alpha, \mathcal{K}^\alpha)_{\alpha\in\mathbf{A}_0}$ then follows from the uniqueness of the non-linear Doob-Meyer decomposition of [16, Theorem 3.3]. □

6 Appendix

We report here some standard results for Lipschitz BSDEs. The first one can be found in, e.g., Theorem 1.5 in [14]. The second one is proved for completeness, and by lack of a good reference.

Proposition 6.1. *(Stability for Lipschitz BSDEs) Let (Y^1, Z^1) and (Y^2, Z^2) in $\mathbf{S}_2 \times \mathbf{H}_2$ be solutions on $[0,T]$ of Lipschitz BSDEs associated to parameters (ξ^1, g^1) ad (ξ^2, g^2). Then the following stability result holds:*

$$\left\|Y^1 - Y^2\right\|^2_{\mathbf{S}_2} + \left\|Z^1 - Z^2\right\|^2_{\mathbf{H}_2} \;\leq\; C\left(\left\|\xi^1 - \xi^2\right\|^2_{\mathbf{L}_2} + \int_0^T E\left|g^1 - g^2\right|^2(t, Y^1_t, Z^1_t)dt\right) ,$$

for some constant $C > 0$ depending only on T and on the Lipschitz constants of g^1 and g^2.

Proposition 6.2. *Let the conditions $(\mathbf{H}_g)$ hold. Then:*

(i) There exists $C > 0$ which only depends on K_g and T such that

$$\text{esssup}_{\xi \in \mathbf{L}_0([0,1])} |\mathcal{E}^g_t[\xi]| \leq C(1 + E_t\left[|\chi_g|^2\right]^{\frac{1}{2}}) , \qquad 0 \leq t \leq T .$$

(ii) For some $\xi \in \mathbf{L}_2$ and $t \in [0,T]$, consider a family $(\xi^\varepsilon)_{\varepsilon \geq 0} \subset \mathbf{L}_0(\mathbb{R}^d)$ satisfying $|\xi^\varepsilon| \leq \xi$ and $\xi^\varepsilon \in L^0(\mathcal{F}_{(t+\varepsilon)\wedge T})$, for any $\varepsilon > 0$. Then, there exists a family $(\eta_\varepsilon)_{\varepsilon > 0} \subset \mathbf{L}_0(\mathbb{R})$ which converges to 0 $\mathbb{P}-$ a.s. as $\varepsilon \to 0$ such that

$$|\mathcal{E}^g_{t,t+\varepsilon}[\xi^\varepsilon] - E_t\left[\xi^\varepsilon\right]| \leq \eta_\varepsilon, \quad \forall \varepsilon \in [0, T-t].$$

(iii) Let $(\xi^\varepsilon)_{\varepsilon > 0}$ and $t \in [0,T]$ be as in (ii). Then, there exists a family $(\eta_\varepsilon)_{\varepsilon > 0} \subset \mathbf{L}_0(\mathbb{R})$ which converges to 0 $\mathbb{P}-$ a.s. as $\varepsilon \to 0$ such that

$$|\mathcal{E}^g_{t-\varepsilon,t}[\xi^\varepsilon] - E_t\left[\xi^\varepsilon\right]| \leq \eta_\varepsilon, \quad \forall \varepsilon \in [0, t].$$

Proof. a. We first prove (ii) (property (iii) being similar) using the standard linearization argument. Fix $t \in [0,T]$ and set $Y^\varepsilon := \mathcal{E}^g_{\cdot,t+\varepsilon}[\xi^\varepsilon]$. Assumption $(\mathbf{H}_g)$ implies that we can find a family of predictable processes $(\rho^\varepsilon, \gamma^\varepsilon)$ with values in $[-K_g, K_g]^{d+1}$ such that

$$L^\varepsilon Y^\varepsilon + \int_t^\cdot L^\varepsilon_r g(r,0,0)dr$$

is a martingale on $[t, t+\varepsilon]$, with

$$L^\varepsilon_s = 1 + \int_t^s \rho^\varepsilon_r L^\varepsilon_r dr + \int_t^s \gamma^\varepsilon_r L^\varepsilon_r dW_r, \; t \leq s \leq t+\varepsilon.$$

In particular,

$$\mathcal{E}^g_{t,t+\varepsilon}[\xi^\varepsilon] = L^\varepsilon_t Y^\varepsilon_t = E_t\left[L^\varepsilon_{t+\varepsilon}\xi^\varepsilon + \int_t^{t+\varepsilon} L^\varepsilon_r g(r,0,0)dr\right].$$

Condition $(\mathbf{H}_g)$ and the assumption on $(\xi^\varepsilon)_{\varepsilon>0}$ thus leads to

$$|\mathcal{E}^g_{t,t+\varepsilon}[\xi^\varepsilon] - E_t\left[\xi^\varepsilon\right]| \leq \eta_\varepsilon,$$

in which

$$\eta_\varepsilon := E_t\left[\xi|L^\varepsilon_{t+\varepsilon} - L^\varepsilon_t| + \chi_g \int_t^{t+\varepsilon} L^\varepsilon_r dr\right].$$

We have:

$$\begin{aligned}|\eta_\varepsilon| &\le E_t[|\xi|^2]^{1/2} E_t[|L^\varepsilon_{t+\varepsilon} - L^\varepsilon_t|^2]^{1/2} + E_t[|\chi_g|^2]^{1/2} E_t\left[\left|\int_t^{t+\varepsilon} L^\varepsilon_r dr\right|^2\right]^{1/2} \\ &\le E_t[|\xi|^2]^{1/2} E_t[|L^\varepsilon_{t+\varepsilon} - L^\varepsilon_t|^2]^{1/2} + \varepsilon E_t[|\chi_g|^2]^{1/2} E_t\left[\sup_{t\le s\le t+\varepsilon} |L^\varepsilon_s|^2\right]^{1/2}. \qquad (6.1)\end{aligned}$$

In addition,

$$\begin{aligned}E_t[|L^\varepsilon_{t+\varepsilon} - L^\varepsilon_t|^2] &\le C E_t\left[\int_t^{t+\varepsilon} |L^\varepsilon_r|^2 dr\right] \\ &\le \varepsilon C E_t\left[\sup_{t\le r\le t+\varepsilon} |L^\varepsilon_r|^2\right]\end{aligned}$$

Hence,

$$E_t[|L^\varepsilon_{t+\varepsilon} - L^\varepsilon_t|^2] \le \varepsilon C\left(1 + E_t\left[\sup_{t\le r\le t+\varepsilon} |L^\varepsilon_r - L^\varepsilon_t|^2\right]\right).$$

Since γ^ε and ρ^ε are bounded, the quantity $\sup_{t\le\tau\le t+\varepsilon} E_\tau[|L^\varepsilon_{t+\varepsilon} - L^\varepsilon_\tau|^2]$ is uniformly bounded. Plugging back this estimate in (6.1) and recalling that $\sup_{t\in[0,T]} E_t[\xi^2]$ is finite $\mathbb{P}-$a.s. we get that $E_t[|\xi|^2]^{1/2} E_t[|L^\varepsilon_{t+\varepsilon} - L^\varepsilon_t|^2]^{1/2}$ tends to 0 uniformly in t, $\mathbb{P}-$a.s. as ε goes to 0. The second term of (6.1) can be estimated in the same way.

b. We now prove (i). Pick any $t \in [0,T]$ and $\xi \in \mathbf{L}_0([0,1])$. The same arguments as above yield

$$|\mathcal{E}^g_t[\xi]| \;\le\; \left|E_t\left[L^\xi_T \xi + \int_t^T L^\xi_r g(r,0,0) dr\right]\right| \le E_t\left[|L^\xi_T| + T|\chi_g| \sup_{r\le T} |L^\xi_r| dr\right],$$

where L^ξ solves

$$L^\xi_s = 1 + \int_t^s \rho^\xi_r L^\xi_r dr + \int_t^s \gamma^\xi_r L^\xi_r dW_r,\ t \le s \le T,$$

for some predictable processes (ρ^ξ, γ^ξ) with values in $[-K_g, K_g]^{d+1}$. Hence,

$$|\mathcal{E}^g_t[\xi]| \le E_t\left[|L^\xi_T| + T|\chi_g| \sup_{t\le r\le T} |L^\xi_r| dr\right].$$

Since (ρ^ξ, γ^ξ) are valued in $[-K_g, K_g]^{d+1}$, standard estimates imply that we can find $C > 0$, which only depends on K_g such that $E_t\left[\sup_{t\le r\le T} |L^\xi_r|^2\right] \le C^2$ $\mathbb{P}-$a.s. The above leads to

$$|\mathcal{E}^g_t[\xi]| \le \left(C + TCE_t\left[|\chi_g|^2\right]^{\frac{1}{2}}\right),$$

and the arbitrariness of $\xi \in \mathbf{L}_0([0,1])$ concludes the proof. □

References

[1] C.D. Aliprantis and K.C. Border. *Infinite dimensional analysis: a hitchhiker's guide*. Springer Verlag, 2006.

[2] E. Bayraktar and S. Yao. On zero-sum stochastic differential games. Technical report, University of Michigan, 2012.

[3] B. Bouchard, R. Elie, and N. Touzi. Stochastic target problems with controlled loss. *SIAM Journal on Control and Optimization*, 48(5):3123–3150, 2009.

[4] B. Bouchard and N. Touzi. Weak dynamic programming principle for viscosity solutions. *SIAM Journal on Control and Optimization*, 49(3):948–962, 2011.

[5] R. Buckdahn and J. Li. Stochastic differential games and viscosity solutions of Hamilton-Jacobi-Bellman-isaacs equations. *SIAM Journal on Control and Optimization*, 47(1):444–475, 2008.

[6] Z. Chen and S. Peng. A general downcrossing inequality for g-martingales. *Statistics & probability letters*, 46(2):169–175, 2000.

[7] M.G. Crandall, H. Ishii, and P.L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Amer. Math. Soc.*, 27:1–67, 1992.

[8] C. Delacherie and PA Meyer. Probabilities and potential. *Math. Studies*, 29, 1978.

[9] N. El Karoui, S. Peng, and M.C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.

[10] H. Föllmer and P. Leukert. Quantile hedging. *Finance and Stochastics*, 3:251–273, 1999.

[11] H. Föllmer and P. Leukert. Efficient hedging: cost versus shortfall risk. *Finance and Stochastics*, 4:117–146, 2000.

[12] L. Moreau. Stochastic target problems with controlled expected loss in jump diffusion models. *SIAM Journal on Control and Optimization*, 49:2577–2607, 2011.

[13] J. Neveu. *Discrete-parameter martingales*, volume 10. Elsevier, 1975.

[14] E. Pardoux. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic pdes of second order. *Progress in Probability*, pages 79–128, 1998.

[15] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Stochastic partial differential equations and their applications*, pages 200–217, 1992.

[16] S. Peng. Monotonic limit theorem of BSDE and nonlinear decomposition theorem of doob-meyer's type. *Probability theory and related fields*, 113:473–499, 1999.

[17] S. Peng. Nonlinear expectations, nonlinear evaluations and risk measures. *Stochastic methods in finance*, pages 243–256, 2004.

[18] R.T. Rockafellar. *Convex analysis*, volume 28. Princeton Univ Pr, 1997.

[19] E. Rosazza Gianin. Risk measures via g-expectations. *Insurance: Mathematics and Economics*, 39(1):19–34, 2006.

[20] H.M. Soner and N. Touzi. Stochastic target problems, dynamic programming and viscosity solutions. *SIAM Journal on Control and Optimization*, 41:404–424, 2002.

[21] H.M. Soner, N. Touzi, and J. Zhang. Wellposedness of second order backward SDEs. *Probability Theory and Related Fields*, 153(1-2):149–190, 2011.

[22] J. Yong and X.Y. Zhou. *Stochastic controls: Hamiltonian systems and HJB equations*, volume 43. Springer Verlag, 1999.