Regularity of BSDEs with a convex constraint on the gains-process

Bruno Bouchard* Romuald Elie† Ludovic Moreau‡

August 13, 2021

Abstract

We consider the minimal super-solution of a backward stochastic differential equation with constraint on the gains-process. The terminal condition is given by a function of the terminal value of a forward stochastic differential equation. Under boundedness assumptions on the coefficients, we show that the first component of the solution is Lipschitz in space and $\frac{1}{2}$-Hölder in time with respect to the initial data of the forward process. Its path is continuous before the time horizon at which its left-limit is given by a face-lifted version of its natural boundary condition. This first component is actually equal to its own face-lift. We only use probabilistic arguments. In particular, our results can be extended to certain non-Markovian settings.

Key words: Backward stochastic differential equation with a constraint, stability, regularity.

MSC Classification (2010): 60H10, 60H30, 49L20.

1 Introduction

The aim of this paper is to establish new stability results for the minimal super-solution $(\hat{\mathcal{Y}}^\zeta, \hat{\mathcal{Z}}^\zeta)$ of a backward differential equation of the form

$$U_t = g(X_T^\zeta) + \int_t^T f(X_s^\zeta, U_s, V_s)ds - \int_t^T V_s dW_s,\ t \le T,$$

satisfying the constraint

$$\hat{\mathcal{Z}}^\zeta \sigma(X^\zeta)^{-1} \in K \ \ dt \otimes d\mathbb{P}-\text{a.e.}$$

*Université Paris-Dauphine CEREMADE UMR CNRS 7534 & CREST, email bouchard@ceremade.dauphine.fr

†Université Paris-Est, LAMA UMR CNRS 8050, email romuald.elie@univ-mlv.fr

‡ETH Zürich, Departement für Mathematik, email ludovic.moreau@math.ethz.ch. Partially supported by the Swiss National Science Foundation under the grant SNF 200021_152555 and by the ETH Foundation.

Research supported by ANR Liquirisk and Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047).

In the above, W is a d-dimensional Brownian motion and X^ζ solves a forward stochastic differential equation with volatility parameter σ, indexed by the initial conditions $\zeta = (t,x) \in [0,T] \times \mathbb{R}^d$: $X_t^\zeta = x$.

Estimates on the regularity can be of important use in many applications, in particular in the design of probabilistic numerical schemes which, to the best of our knowledge, are missing for such constrained backward differential equations.

When $K = \mathbb{R}^d$, i.e. there is no constraint, and the coefficients are Lipschitz continuous, it is well-known that $\hat{\mathcal{Y}}^\zeta$ has continuous path and that the (deterministic) map $(t,x) \mapsto \hat{\mathcal{Y}}_t^{(t,x)}$ is 1/2-Hölder in time and Lipschitz in space:

$$|\hat{\mathcal{Y}}_t^{(t,x)} - \hat{\mathcal{Y}}_{t'}^{(t',x')}| \le C\,(|t-t'|^{\frac{1}{2}} + |x-x'|). \tag{1.1}$$

See e.g. [9]. This basically follows from standard estimates using Itô's and Gronwall's Lemma.

In the general case, such a minimal super-solution solves an equation of the form

$$\hat{\mathcal{Y}}_t^\zeta \;=\; g(X_T^\zeta) + \int_t^T f(X_s^\zeta, \hat{\mathcal{Y}}_s^\zeta, \hat{\mathcal{Z}}_s^\zeta)ds - \int_t^T \hat{\mathcal{Z}}_s^\zeta dW_s + \hat{\mathcal{K}}_T^\zeta - \hat{\mathcal{K}}_t^\zeta,\ t \le T,$$

in which $\hat{\mathcal{K}}^\zeta$ is an adapted non-decreasing process, see [6, 10]. Because little is known on the regularity of this process, the technics used in the unconstrained case can not be reproduced.

Nevertheless, it is well-known that such a minimal super-solution can be approximated by a sequence of penalized unconstrained stochastic backward differential equations, see [6, 10]. It is therefore tempting to use the estimates associated to each element of the approximating sequence and to pass to the limit. Unfortunately, the Lipschitz continuity coefficients of the approximating sequence blow-up.

Another way to proceed consists in using the dual formulation of [5, 6]. In their representation, the component $\hat{\mathcal{Y}}^\zeta$ is identified to the value of the optimal control problem of a family of backward stochastic differential equations written under a suitable set of equivalent probability measures, see Section 4.1. The main difficulty is that it is singular: each of the controls is bounded, but the bound is not uniform.

In this paper, we essentially make use of this dual formulation, but we use a strong version: the controls are directly incorporated in the dynamics rather than through changes of measures. See Section 3. Space stability is essentially obvious for this strong version, while it is not in its original 'weak' form. Still, the singularity of the optimal control problem makes estimates on the stability in time quite delicate a-priori.

The key idea of this paper is to use the fact that the solution is automatically 'face-lifted' in the sense of Proposition 3.3 below. This 'face-lifting' phenomenon is well-known as far as the terminal condition is concerned, this goes back to [4] in the specific setting of mathematical finance, see also [3, 7] and the references therein. We use probabilistic arguments to show that it holds also in the parabolic interior $[0,T) \times \mathbb{R}^d$ of the domain, which can be guessed in the setting of [3, 7] from their pde characterization. This 'face-lifting' phenomenon allows to absorb the singular control, and to extend (1.1) to the constrained case. See Theorem 2.1.

The paper is organized as follows. The setting and the main results are stated in Section 2. Comments on our assumptions and possible extensions are discussed at the end in Section 5. Section 3 is dedicated to the strong version of the dual formulation for which our estimates are established. This part contains the main ideas of this paper. In Section 4, we show that the strong dual formulation coincides with its weak version, and that the latter actually provides (as well-known when the driver is convex) the first component $\hat{\mathcal{Y}}^{\zeta}$ of the constrained backward stochastic differential equation.

Notations: All over this paper, we let $\Omega = C([0,T], \mathbb{R}^d)$, $d \geq 1$, $T > 0$, be the canonical space of continuous d-dimensional functions ω on $[0,T]$ such that $\omega_0 = 0$. It is endowed with the Wiener measure $\mathbb{P}$. We let W be the coordinate process, $W_t(\omega) = \omega_t$, and we denote by $\mathbb{F} = (\mathcal{F}_t)_{t \leq T}$ the augmentation of its raw filtration under $\mathbb{P}$. Random variables are defined on $(\Omega, \mathcal{F}_T, \mathbb{P})$. For the expectation under $\mathbb{P}$, we simply use the symbol $\mathbb{E}$, while we write $\mathbb{E}^{\mathbb{Q}}$ if it is taken under a different measure $\mathbb{Q}$. Given a probability measure $\mathbb{Q}$, $\mathcal{G} \subset \mathcal{F}_T$, $p \geq 1$ and $A \subset \mathbb{R}^n$, we denote by $\mathbf{L}_p(A, \mathbb{Q}, \mathcal{G})$ the set of $\mathcal{G}$-measurable A-valued random variables whose p-moment under $\mathbb{Q}$ is finite. We let $\mathbf{S}_2(\mathbb{Q})$ (resp. $\mathbf{H}_2(\mathbb{Q})$)be the set of $\mathbb{R}^n$-valued progressively measurable processes V such that $\mathbb{E}^{\mathbb{Q}}[\sup_{t \leq T} |V_t|^2] < \infty$ (resp. $\mathbb{E}^{\mathbb{Q}}[\int_t^T |V_t|^2 dt] < \infty$), in which $|v|$ denotes the Euclydian norm of $v \in \mathbb{R}^n$ and n is given by the context. The set of stopping times with values in $[0,T]$ is $\mathcal{T}$, while $\mathcal{T}_\tau$ is the set of stopping times a.s. greater than $\tau \in \mathcal{T}$. Finally, $\mathbf{D}_2(\mathbb{Q})$ denotes the set of couples $(\tau, \xi) \in \mathcal{T} \times \mathbf{L}_2(\mathbb{Q}, \mathbb{R}^d, \mathcal{F}_T)$ such that ξ is $\mathcal{F}_\tau$-measurable. For $\zeta \in \mathbf{D}_2$, we write $(\tau_\zeta, \xi_\zeta) = \zeta$. In all these definitions, we omit the arguments that can be clearly identified by the context. When nothing else is specified, inequalities between random variables or convergence of sequences of random variables hold in the $\mathbb{P}$-a.s. sense.

2 Main regularity and stability results

As a first step of analysis, we concentrate on a Markovian setting with rather stringent boundedness assumptions. Possible extensions will be discussed in Section 5 at the end of this paper. They include another type of constraint, certain non-Markovian settings and optimal control problems.

The forward component is the unique strong solution X^ζ on $[0,T]$ of the stochastic differential equation

$$X^{\zeta}_{t \vee \tau_\zeta} = \xi_\zeta + \int_{\tau_\zeta}^{t \vee \tau_\zeta} b_s(X^\zeta_s) ds + \int_{\tau_\zeta}^{t \vee \tau_\zeta} \sigma_s(X^\zeta_s) dW_s, \tag{2.1}$$

in which the initial data $\zeta \in \mathbf{D}_2$, and the parameters $(b, \sigma) : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d \times \mathbb{R}^{d \times d}$ are measurable maps. They are assumed to be bounded, and Lipschitz in their space variable, uniformly in their time argument. We also assume that σ is invertible with bounded inverse. Namely, there exists $L > 0$ such that

$$|(b_t, \sigma_t)(x) - (b_t, \sigma_t)(x')| \leq L|x - x'| \quad \text{and} \quad (|b_t| + |\sigma_t| + |\sigma_t^{-1}|)(x) \leq L, \tag{2.2}$$

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

The backward equation is defined by two measurable maps $f : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ and $g : \mathbb{R}^d \mapsto \mathbb{R}$ such that, for all $(t,x,\theta),(t,x',\theta') \in [0,T] \times \mathbb{R}^d \times (\mathbb{R} \times \mathbb{R}^d)$,

$$|f_t(x,\theta) - f_t(x',\theta')| \leq L\left(|x-x'| + |\theta - \theta'|\right)\ ,\ |f_t(x,\theta)| \leq L(1+|\theta|)\ , \tag{2.3}$$

$$|g(x)| \leq L, \quad \text{and } g \text{ is lower-semicontinuous.} \tag{2.4}$$

A supersolution of BSDE(f,g,ζ) is a process $(U,V) \in \mathbf{S}_2 \times \mathbf{H}_2$ satisfying

$$\begin{array}{rcl} U_{t\vee\tau_\zeta} & \geq & U_{t'\vee\tau_\zeta} + \int_{t\vee\tau_\zeta}^{t'\vee\tau_\zeta} f_s(X_s^\zeta, U_s, V_s)ds - \int_{t\vee\tau_\zeta}^{t'\vee\tau_\zeta} V_s dW_s,\ t \leq t' \leq T. \\ U_T & = & g(X_T^\zeta). \end{array} \tag{2.5}$$

The constraint on the V coordinate is associated to a family $(K_t)_{t\leq T}$ of closed convex sets of $\mathbb{R}^d$:

$$V\sigma(X^\zeta)^{-1} \in K\ \ dt \otimes d\mathbb{P}\text{−a.e on } [\![\tau_\zeta, T]\!]. \tag{2.6}$$

When it is satisfied, we say that (U,V) is a super-solution of BSDE$_K(f,g,\zeta)$. This super-solution is said to be minimal if $U'_{s\vee\tau_\zeta} \geq U_{s\vee\tau_\zeta}$ for all $s \leq T$ and any other super-solution $(U',V') \in \mathbf{S}_2 \times \mathbf{H}_2$ of BSDE$_K(f,g,\zeta)$.

We require that

$$0 \in K_t \text{ for all } t \leq T \tag{2.7}$$

$$\cup_{t\leq T} K_t \text{ is bounded,} \tag{2.8}$$

and that, for all $u \in \mathbb{R}^d$,

$$t \in [0,T] \mapsto \delta_t(u) := \sup\{\ k^\top u, k \in K_t\}\ \text{ is left-continuous at } T, \tag{2.9}$$

$$\text{and non-increasing.} \tag{2.10}$$

Remark 2.1. *The conditions* (2.8)-(2.10) *are equivalent to : the family* $(K_t)_{t\leq T}$ *is non-increasing and* K_0 *is bounded.*

Note that our standing assumptions (2.3)-(2.4)-(2.7) ensure that BSDE$_K(f,g,\zeta)$ admits a trivial super-solution

$$(\mathrm{y}_t, \mathrm{z}_t) = ((1 + (T-t))L, 0), \tag{2.11}$$

which is bounded. In particular, [10, Theorem 4.2] implies that BSDE$_K(f,g,\zeta)$ admits a minimal super-solution. We denote it by $(\hat{\mathcal{Y}}^\zeta, \hat{\mathcal{Z}}^\zeta)$, and let $\hat{\mathcal{K}}^\zeta$ be the non-decreasing process defined on $[0,T]$ by

$$\hat{\mathcal{K}}^\zeta_{\tau_\zeta} = 0 \text{ and } \hat{\mathcal{Y}}^\zeta_{\cdot\vee\tau_\zeta} = g(X_T^\zeta) + \int_{\cdot\vee\tau_\zeta}^T f_s(X_s^\zeta, \hat{\mathcal{Y}}_s^\zeta, \hat{\mathcal{Z}}_s^\zeta)ds - \int_{\cdot\vee\tau_\zeta}^T \hat{\mathcal{Z}}_s^\zeta dW_s + \hat{\mathcal{K}}_T^\zeta - \hat{\mathcal{K}}^\zeta_{\cdot\vee\tau_\zeta}.$$

We do not impose any Lipschitz continuity assumption on g, although it is used in the unconstrained case $K = \mathbb{R}^d$ to obtain the Lipschitz continuity of the map $\xi \mapsto \hat{\mathcal{Y}}_\tau^{(\tau,\xi)}$. Instead, we assume that the map

$$\hat{g} : x \in \mathbb{R}^d \mapsto \sup_{u\in\mathbb{R}^d} \left(g(x+u) - \delta_T(u)\right),\ x \in \mathbb{R}^d \text{ is } L\text{-Lipschitz continuous.} \tag{2.12}$$

This map is usually referred to as the 'face-lift' of g for the constraint K_T, compare with e.g. [4, 3, 7]. We shall see below that it provides the correct time-T boundary condition for our constrained backward differential equation. Intuitively, this means that assuming that g is Lipschitz is useless, whenever $\hat{g}$ is, which is a weaker condition [1].

Our main result shows that the map $\zeta \in \mathbf{D}_2 \mapsto \hat{\mathcal{Y}}^{\zeta}_{\tau_\zeta}$ satisfies similar regularity properties in time and space as in the unconstrained case. It also shows that the non-decreasing process $\hat{\mathcal{K}}^{\zeta}$ is continuous on $[0, T)$ with a final jump of size $(\hat{g} - g)(X^{\zeta}_T)$. In particular, $\hat{\mathcal{Y}}^{\zeta}_{T-} = \hat{g}(X^{\zeta}_T)$ on $\{\tau_\zeta < T\}$.

From now on, we denote by C_L a generic constant which depends only on L, and may change from line to line.

Theorem 2.1. *The following holds for all $\zeta, \zeta' \in \mathbf{D}_2$:*

(a) *If $\tau_\zeta \leq \tau_{\zeta'} < T$, then*

$$|\hat{\mathcal{Y}}^{\zeta}_{\tau_\zeta} - \mathbb{E}_{\tau_\zeta}[\hat{\mathcal{Y}}^{\zeta'}_{\tau_{\zeta'}}]| \leq C_L \left(\mathbb{E}_{\tau_\zeta}[|\tau_{\zeta'} - \tau_\zeta|]^{\frac{1}{2}} + \mathbb{E}_{\tau_\zeta}[|\xi_{\zeta'} - \xi_\zeta|] \right). \tag{2.13}$$

(b) *If $\tau := \tau_\zeta = \tau_{\zeta'} < T$, then*

$$-\delta_\tau(\xi_{\zeta'} - \xi_\zeta) \leq \hat{\mathcal{Y}}^{\zeta}_\tau - \hat{\mathcal{Y}}^{\zeta'}_\tau \leq \delta_\tau(\xi_\zeta - \xi_{\zeta'}). \tag{2.14}$$

(c) *If $\tau_\zeta < T$, then $\hat{\mathcal{K}}^{\zeta}_{\cdot\wedge\vartheta}$ has continuous path for each stopping time $\vartheta < T$. Moreover, if $(\vartheta_n)_{n\geq 1}$ is a sequence of stopping times with values in $[\tau_\zeta, T)$ such that $\vartheta_n \to T$, then $\hat{\mathcal{Y}}^{\zeta}_{\vartheta_n} \to \hat{g}(X^{\zeta}_T)$.*

Sections 3 and 4 are devoted to the proof of these results. In view of Proposition 4.1 and Theorem 4.1: (a) is a consequence of Corollary 3.1 and Proposition 3.4, (b) follows from Proposition 3.3, and Proposition 3.5 implies (c).

3 Estimates via the strong dual formulation

As explained in the introduction, the constrained backward differential equation $\mathrm{BSDE}_K(f, g, \zeta)$ admits a dual representation which is formulated as an optimal control problem on a family of unconstrained backward stochastic differential equations written under a suitable family of equivalent laws, see Section 4 for a precise formulation. Although, each unconstrained backward stochastic differential equation satisfies the usual Hölder and Lipschitz regularity properties, this does not seem to allow one to obtain the estimates of Theorem 2.1. The reason is that the optimal control problem is of singular type: constants may blow up when passing to the supremum.

The main idea of this paper is to start with a strong version of this dual optimal control problem. Strong meaning that the probability measure is fixed, but we incorporate the control directly in the dynamics. It turns out to be much more flexible. In particular,

[1] The fact that $\hat{g}$ inherits the Lipschitz–continuity property from g is by construction, whereas the converse is not valid: for $d = 1$, $K = \mathbb{R}_+$ and $g : x \in \mathbb{R}_+ \longmapsto \mathbf{1}_{\{x<1\}}$, we have $\hat{g} : x \in \mathbb{R}_+ \longmapsto 1$.

space stability is essentially trivial in this setting, see Corollay 3.1. More importantly, we can show that the corresponding value process is itself automatically 'face-lifted', see Proposition 3.3 below. This will be the key result to obtain the time regularity estimates of Proposition 3.4.

3.1 The strong dual formulation

Let $\mathcal{U}$ denote the collection of $\mathbb{R}^d$-valued bounded predictable processes. Note that $\delta(\nu)$ is bounded for each $\nu \in \mathcal{U}$, see (2.8). To each $(\zeta, \nu) \in \mathbf{D}_2 \times \mathcal{U}$, we associate the stochastic driver

$$f^{\zeta,\nu} : (t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^d \mapsto \Big(f_t(X_t^{\zeta,\nu}, y, z) - \delta_t(\nu_t)\Big)\, \mathbf{1}_{[\![\tau_\zeta, T]\!]}(t)$$

where $X^{\zeta,\nu}$ is the solution of

$$X^{\zeta,\nu} \;=\; \xi_\zeta + \int_{\tau_\zeta}^{\cdot\vee\tau_\zeta} \Big(b_s(X_s^{\zeta,\nu}) + \nu_s\Big)\, ds + \int_{\tau_\zeta}^{\cdot\vee\tau_\zeta} \sigma_s(X_s^{\zeta,\nu}) dW_s. \tag{3.1}$$

Given $\tau \in \mathcal{T}_{\tau_\zeta}$, $\vartheta \in \mathcal{T}_\tau$ and $G \in \mathbf{L}_2(\mathcal{F}_\vartheta)$, we set

$$\mathcal{E}_{\tau,\vartheta}^{\zeta,\nu}[G] := U_\tau$$

where $(U, V) \in \mathbf{S}_2 \times \mathbf{H}_2$ is the solution of

$$U \;=\; G + \int_{\cdot\vee\tau_\zeta}^{\vartheta} f_s^{\zeta,\nu}(U_s, V_s) ds - \int_{\cdot\vee\tau_\zeta}^{\vartheta} V_s dW_s \;\text{ on } [0, T]. \tag{3.2}$$

In the special case where $\vartheta \equiv T$ and $G = g(X_T^{\zeta,\nu})$, the solution of (3.2) is denoted by $(Y^{\zeta,\nu}, Z^{\zeta,\nu})$. In particular,

$$Y_\cdot^{\zeta,\nu} = \mathcal{E}_{\cdot,T}^{\zeta,\nu}[g(X_T^{\zeta,\nu})].$$

We next define our optimal control problem

$$\mathcal{Y}_\tau^\zeta = \operatorname{ess\,sup}\{Y_\tau^{\zeta,\nu}, \nu \in \mathcal{U},\ \nu\mathbf{1}_{[\![0,\tau]\!]} \equiv 0\}\ ,\ \zeta \in \mathbf{D}_2,\ \tau \in \mathcal{T}_{\tau_\zeta}. \tag{3.3}$$

Note that $\mathcal{Y}_\tau^\zeta = \mathcal{Y}_\tau^{(\tau, X_\tau^\zeta)}$ since $X^{\zeta,0} = X^\zeta$.

Remark 3.1. The conditions (2.2)-(2.3)-(2.4) imply that $Y^{\zeta,\nu}$ is bounded in $\mathbf{L}_\infty$ uniformly in $\zeta \in \mathbf{D}_2$, for all $\nu \in \mathcal{U}$, see Lemma A.1. Moreover, $Y^{\zeta,\nu} \le \mathrm{y}$ defined in (2.11), for all $\nu \in \mathcal{U}$, see Lemma A.2. In particular, $Y^{\zeta,0} \le \mathcal{Y}^\zeta \le \mathrm{y}$, so that $\mathcal{Y}^\zeta$ is bounded in $\mathbf{L}_\infty$ uniformly in $\zeta \in \mathbf{D}_2$. The bound depends only on L.

3.2 Terminal face-lift and space stability

The first result of this section concerns the face-lift of the terminal condition. It shows that g can be replaced by $\hat{g}$ in (3.3). Apart from being of self-interest, this property will be used latter on in the proof of the space stability in our setting where $\hat{g}$ is assumed to be Lipschitz while g may not be. It will also be used to characterize the limit $\lim_{t\uparrow T} \mathcal{Y}_t^\zeta$.

Proposition 3.1. $\mathcal{Y}_{\tau_\zeta}^\zeta = \operatorname{ess\,sup}\{\mathcal{E}_{\tau_\zeta,T}^{\zeta,\nu}[\hat{g}(X_T^{\zeta,\nu})], \nu \in \mathcal{U}\}$ *on* $\{\tau_\zeta < T\}$*, for all* $\zeta \in \mathbf{D}_2$.

Proof. Since $\hat{g} \geq g$ by construction, one inequality is trivially deduced from Lemma A.2. We therefore concentrate on the difficult inequality. Fix $\zeta := (\tau, \xi) \in \mathbf{D}_2$. For sake of simplicity, we assume that $\tau < T$ a.s., the general case is handled by using the fact that $\{\tau < T\} = \cup_{n \geq 1}\{\tau \leq T - n^{-1}\}$. For $\nu \in \mathcal{U}$, and $u \in \mathbf{L}_\infty(\mathcal{F}_{T-\varepsilon_\circ})$, for some $\varepsilon_\circ > 0$, we then define

$$\tau_\varepsilon := (T - \varepsilon) \vee \tau \ , \ \nu^\varepsilon := \nu \mathbf{1}_{[\![\tau,\tau_\varepsilon]\!]} + \frac{u}{T - \tau_\varepsilon} \mathbf{1}_{]\!]\tau_\varepsilon, T]\!]}, \ 0 < \varepsilon < \varepsilon_\circ.$$

Then, (3.3) combined with the tower property for non linear expectations imply that

$$\mathcal{Y}^\zeta_\tau \ \geq \ \mathcal{E}^{\zeta,\nu}_{\tau,\tau_\varepsilon}\left[Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon}\right].$$

We claim that, after possibly considering a subsequence,

$$\liminf_{\varepsilon \to 0} \mathcal{E}^{\zeta,\nu}_{\tau,\tau_\varepsilon}\left[Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon}\right] \geq \mathcal{E}^{\zeta,\nu}_{\tau,T}\left[g(X^{\zeta,\nu}_T + u) - \delta_T(u)\right]. \tag{3.4}$$

By arbitrariness of $\varepsilon_\circ > 0$ and $u \in \mathbf{L}_\infty(\mathcal{F}_{T-\varepsilon_\circ})$, this implies that

$$\mathcal{Y}^\zeta_\tau \ \geq \ \operatorname*{ess\,sup}_{u \in \mathbf{L}_\infty(\mathcal{F}_{T-})} \mathcal{E}^{\zeta,\nu}_{\tau,T}\left[g(X^{\zeta,\nu}_T + u) - \delta_T(u)\right].$$

Since the map $(x, u) \in \mathbb{R}^d \times \mathbb{R}^d \mapsto g(x + u) - \delta_T(u)$ is Borel, it follows from [1, Proposition 7.40, p184], that, for each $\iota > 0$, we can find a universally measurable map $x \in \mathbb{R}^d \mapsto \tilde{u}_\iota(x)$ such that $\hat{g}(x) = \sup\{g(x+u) - \delta_T(u), \, u \in \mathbb{R}^d\} \leq g(x + \tilde{u}_\iota(x)) - \delta_T(\tilde{u}_\iota(x)) + \iota$ for all $x \in \mathbb{R}^d$. By [1, Lemma 7.27, p173], we can find a Borel measurable map $x \in \mathbb{R}^d \mapsto \hat{u}_\iota(x)$ such that $\hat{u}_\iota(X^{\zeta,\nu}_T) = \tilde{u}_\iota(X^{\zeta,\nu}_T)$ a.s. Since $X^{\zeta,\nu}_T \in \mathbf{L}_2(\mathcal{F}_{T-})$ by left-continuity of its path, this implies that

$$\mathcal{Y}^\zeta_\tau \ \geq \ \mathcal{E}^{\zeta,\nu}_{\tau,T}\left[\hat{g}(X^{\zeta,\nu}_T)\mathbf{1}_{\{|\hat{u}_\iota(X^{\zeta,\nu}_T)| \leq n\}} - \iota + g(X^{\zeta,\nu}_T)\mathbf{1}_{\{|\hat{u}_\iota(X^{\zeta,\nu}_T)| > n\}}\right],$$

for each $n \geq 1$ and $\iota \in (0, 1)$. The required result then follows from the stability principle, see Lemma A.2, by sending first $n \to \infty$ and then $\iota \to 0$.

It remains to prove our claim (3.4). We first deduce from Lemma 3.1 stated at the end of this section that, after possibly passing to a subsequence,

$$X^{\zeta,\nu^\varepsilon}_T \to X^{\zeta,\nu}_T + u \ \ a.s. \ \text{ as } \varepsilon \to 0,$$

and therefore

$$\liminf_{\varepsilon \to 0} g(X^{\zeta,\nu^\varepsilon}_T) \geq g(X^{\zeta,\nu}_T + u),$$

by lower-semicontinuity of g. Moreover, if $(H^\varepsilon)_{\varepsilon > 0}$ is a sequence of positive processes such that $\sup_{[\tau_\varepsilon, T]} H^\varepsilon \to 1$ a.s. as $\varepsilon \to 0$, then

$$\begin{aligned} \liminf_{\varepsilon \to 0} \int_{\tau_\varepsilon}^T H^\varepsilon_s \left(f_s(X^{\zeta,\nu^\varepsilon}_s, 0) - \delta_s(\nu^\varepsilon_s)\right) ds \ &\geq \ -\lim_{\varepsilon \to 0} (T - \tau_\varepsilon)^{-1} \int_{\tau_\varepsilon}^T H^\varepsilon_s \delta_s(u) ds \\ &\geq \ -\delta_T(u) \end{aligned}$$

since $t \mapsto \delta_t(u)$ is left-continuous at T and $f(\cdot,0)$ is bounded. We can then combine Lemma A.1 and Lemma A.3 to obtain

$$\liminf_{\varepsilon\to 0} Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon} \geq g(X^{\zeta,\nu}_T + u) - \delta_T(u) =: G,$$

after possibly passing to a subsequence. Now observe that $\mathcal{E}^{\zeta,\nu}_{\cdot,\tau_\varepsilon}[Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon}]$ coincides with the first component of the backward differential equation which driver is $f^{\zeta,\nu}\mathbf{1}_{[\![0,\tau_\varepsilon]\!]}$ and which terminal condition is $Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon}$ at T. It is constant on $[\![\tau_\varepsilon, T]\!]$. Then, by the stability and comparison principles in Lemma A.2, we obtain

$$\begin{aligned}
\mathcal{E}^{\zeta,\nu}_{\tau,\tau_\varepsilon}\left[Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon}\right] &\geq \mathcal{E}^{\zeta,\nu}_{\tau,T}\left[Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon}\right] - C_1\mathbb{E}_\tau\left[\int_{\tau_\varepsilon}^T |f^{\zeta,\nu}_s(Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon},0)|^2 ds\right]^{\frac{1}{2}} \\
&\geq \mathcal{E}^{\zeta,\nu}_{\tau,T}[G] - C_2\left(\mathbb{E}_\tau\left[\int_{\tau_\varepsilon}^T |f^{\zeta,\nu}_s(Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon},0)|^2 ds\right]^{\frac{1}{2}} + \mathbb{E}_\tau\left[\{(Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon} - G)^-\}^2\right]^{\frac{1}{2}}\right),
\end{aligned}$$

in which the constants C_1 and C_2 do not depend on (u,ε), see Remark 3.1. Since $\nu \in \mathcal{U}$, it follows from Lemma A.1 together with (2.2) and (2.3) that $(Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon}, f^{\zeta,\nu}(Y^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon},0)\mathbf{1}_{[\![\tau_\varepsilon,T]\!]})$ is bounded in $\mathbf{L}_2 \times \mathbf{S}_2$ uniformly in $\varepsilon > 0$. Then, combining the above shows that, along a subsequence if necessary, the right-hand side term in the last inequality converges to 0 as $\varepsilon \downarrow 0$. Hence, the claim (3.4) holds, which completes the proof. □

Since $\hat{g}$ is assumed to be Lipschitz continuous, the stability in space is now an easy consequence of the representation given in Proposition 3.1

Corollary 3.1. *$\left|\mathcal{Y}^{\zeta_1}_\tau - \mathcal{Y}^{\zeta_2}_\tau\right| \leq C_L|\xi_{\zeta_1} - \xi_{\zeta_2}|$, for all $\zeta_1,\zeta_2 \in \mathbf{D}_2$ with $\tau := \tau_{\zeta_1} = \tau_{\zeta_2} < T$.*

Proof. We simply use the fact that

$$\left|\mathcal{Y}^{\zeta_1}_\tau - \mathcal{Y}^{\zeta_2}_\tau\right| \leq \operatorname*{ess\,sup}_{\nu\in\mathcal{U}} \left|\mathcal{E}^{\zeta_1,\nu}_{\tau,T}[\hat{g}(X^{\zeta_1,\nu}_T)] - \mathcal{E}^{\zeta_2,\nu}_{\tau,T}[\hat{g}(X^{\zeta_2,\nu}_T)]\right|$$

by Proposition 3.1. The right-hand side is bounded by $|\xi_{\zeta_1} - \xi_{\zeta_2}|$ up to a multiplicative constant under our Lipschitz continuity assumptions (2.2)-(2.3)-(2.12), see Lemma A.2. □

We conclude this section with the technical lemma that was used in the proof of Proposition 3.1. The proof is trivial under (2.2) and we omit it. It is not difficult to see that it remains correct without the boundedness assumption on (b,σ), they only need to be Lipschitz continuous in space, uniformly in time (but then the constant appearing in the bound depends on (u,ζ) as well).

Lemma 3.1. *Fix $\zeta \in \mathbf{D}_2$, $\vartheta \in \mathcal{T}_{\tau_\zeta}$ and $u \in \mathbf{L}_\infty(\mathbb{R}^d, \mathcal{F}_{\tau_\zeta})$. Set $\nu := \varepsilon^{-1}u\mathbf{1}_{\{\varepsilon>0\}}\mathbf{1}_{[\![\tau_\zeta,\vartheta]\!]}$, with $\varepsilon := \vartheta - \tau_\zeta$. Then,*

$$\sup_{t\leq T}\mathbb{E}_{\tau_\zeta}\left[\left|X^{\zeta,\nu}_{t\vee\tau_\zeta\wedge\vartheta} - \xi_\zeta - \varepsilon^{-1}u(t\vee\tau_\zeta\wedge\vartheta - \tau_\zeta)\mathbf{1}_{\{\varepsilon>0\}}\right|^2\right] \leq C_L\mathbb{E}_{\tau_\zeta}[\varepsilon].$$

3.3 Dynamic programming and face-lifting on $[0,T)$

We first recall the dynamic programming principle for the optimal control problem (3.3). It will be used later on in this section to prove that the value process is automatically face-lifted, see Proposition 3.3.

Proposition 3.2. *For all $\zeta \in \mathbf{D}_2$ and $\vartheta \in \mathcal{T}_{\tau_\zeta}$,*

$$\mathcal{Y}^{\zeta}_{\tau_\zeta} = \operatorname*{ess\,sup}_{\nu \in \mathcal{U}} \mathcal{E}^{\zeta,\nu}_{\tau_\zeta,\vartheta}[\mathcal{Y}^{(\vartheta, X^{\zeta,\nu}_\vartheta)}_\vartheta].$$

Proof. The proof is standard. Since $Y^{\zeta,\nu}_{\tau_\zeta} = \mathcal{E}^{\zeta,\nu}_{\tau_\zeta}[Y^{(\vartheta,X^{\zeta,\nu}_\vartheta),\nu}_\vartheta] \le \mathcal{E}^{\zeta,\nu}_{\tau_\zeta}[\mathcal{Y}^{(\vartheta,X^{\zeta,\nu}_\vartheta)}_\vartheta]$ by the tower property for non-linear expectations and by the comparison principle, see Lemma A.2, one inequality is trivial. As for the reverse inequality, we observe that the family $\{Y^{(\vartheta,X^{\zeta,\nu}_\vartheta),\nu'}_\vartheta, \nu' \in \mathcal{U}\}$ is directed upward. Then [8, Proposition VI.1.1] ensures that we can find a sequence $(\nu'_n)_{n\ge1} \subset \mathcal{U}$ such that $Y^n_\vartheta := Y^{(\vartheta,X^{\zeta,\nu}_\vartheta),\nu'_n}_\vartheta \uparrow \mathcal{Y}^{(\vartheta,X^{\zeta,\nu}_\vartheta)}_\vartheta =: \mathcal{Y}_\vartheta$ a.s. as $n \to \infty$. Since, Y^1_ϑ and $\mathcal{Y}_\vartheta$ are bounded in $\mathbf{L}_2$, see Remark 3.1, the convergence holds in $\mathbf{L}_2$ as well. Moreover, we can find a constant $C > 0$ which does not depend on n and such that

$$\mathcal{E}^{\zeta,\nu}_{\tau_\zeta,\vartheta}[Y^n_\vartheta] \quad \ge \quad \mathcal{E}^{\zeta,\nu}_{\tau_\zeta,\vartheta}[\mathcal{Y}_\vartheta] - C\, \mathbb{E}_{\tau_\zeta}[|Y^n_\vartheta - \mathcal{Y}_\vartheta|^2]^{\frac{1}{2}},$$

see Lemma A.2. The latter combined with (3.3) implies that

$$\mathcal{Y}^{\zeta}_{\tau_\zeta} \ge \mathcal{E}^{\zeta,\nu}_{\tau_\zeta,\vartheta}[\mathcal{Y}_\vartheta] - \lim_{n\to\infty} C\, \mathbb{E}_{\tau_\zeta}[|Y^n_\vartheta - \mathcal{Y}_\vartheta|^2]^{\frac{1}{2}} = \mathcal{E}^{\zeta,\nu}_{\tau_\zeta,\vartheta}[\mathcal{Y}_\vartheta],$$

and we conclude by arbitrariness of $\nu \in \mathcal{U}$. □

We can now show that $\xi \in \mathbf{L}_2(\mathcal{F}_\tau) \mapsto \mathcal{Y}^{(\tau,\xi)}_\tau$ is itself automatically face-lifted in the following sense.

Proposition 3.3. *For all $\zeta \in \mathbf{D}_2$,*

$$\mathcal{Y}^{\zeta}_{\tau_\zeta} = \operatorname*{ess\,sup}_{u \in \mathbf{L}_\infty(\mathbb{R}^d, \mathcal{F}_{\tau_\zeta})} (\mathcal{Y}^{(\tau_\zeta, \xi_\zeta + u)}_{\tau_\zeta} - \delta_{\tau_\zeta}(u)) \quad a.s. \ on \ \{\tau_\zeta < T\}. \tag{3.5}$$

Proof. Take $\zeta := (\tau,\xi) \in \mathbf{D}_2$. One inequality follows from the fact that $\delta_\tau(0) = 0$. Fix $u \in \mathbf{L}_\infty(\mathcal{F}_\tau)$, $\varepsilon > 0$, and set $\tau_\varepsilon := (\tau+\varepsilon) \wedge T$ and $\nu^\varepsilon := \varepsilon^{-1} u \mathbf{1}_{[\![\tau,\tau_\varepsilon]\!]}$. It follows from Lemma 3.1 that

$$\mathbb{E}_\tau[|X^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon} - \xi - u|^2]^{\frac{1}{2}} \le C_L\, \varepsilon^{\frac{1}{2}}.$$

Then, by appealing to Proposition 3.2, Lemma A.1, (2.3) and Corollary 3.1 successively, we can find a family of non-negative continuous processes $(H^\varepsilon)_{\varepsilon>0}$, uniformly bounded in $\mathbf{S}_2$, such that $H^\varepsilon_{\tau_\varepsilon} \to 1$ in $\mathbf{L}_1$ as $\varepsilon \to 0$ and

$$\begin{aligned}
\mathcal{Y}^{\zeta}_{\tau} &\ge \mathbb{E}_\tau\left[H^\varepsilon_{\tau_\varepsilon} \mathcal{Y}^{(\tau_\varepsilon, X^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon})}_{\tau_\varepsilon} + \int_\tau^{\tau_\varepsilon} H^\varepsilon_s f^{\zeta,\nu^\varepsilon}_s(0) ds\right] \\
&\ge \mathbb{E}_\tau\left[H^\varepsilon_{\tau_\varepsilon} \mathcal{Y}^{(\tau_\varepsilon, X^{\zeta,\nu^\varepsilon}_{\tau_\varepsilon})}_{\tau_\varepsilon} - \int_\tau^{\tau_\varepsilon} H^\varepsilon_s \delta_s(\nu^\varepsilon_s) ds\right] - C_L \varepsilon \\
&\ge \mathbb{E}_\tau\left[H^\varepsilon_{\tau_\varepsilon} \mathcal{Y}^{(\tau_\varepsilon, \xi+u)}_{\tau_\varepsilon} - \varepsilon^{-1} \int_\tau^{\tau_\varepsilon} H^\varepsilon_s \delta_s(u) ds\right] - C_L \varepsilon^{\frac{1}{2}}.
\end{aligned}$$

Since $H^\varepsilon \geq 0$, we can use (2.10) and Remark 3.1 to obtain

$$
\begin{aligned}
\mathcal{Y}^\zeta_\tau &\geq \mathbb{E}_\tau\left[H^\varepsilon_{\tau_\varepsilon}\mathcal{Y}^{(\tau_\varepsilon,\xi+u)}_{\tau_\varepsilon} - \delta_\tau(u)\varepsilon^{-1}\int_\tau^{\tau_\varepsilon} H^\varepsilon_s ds\right] - C_L\varepsilon^{\frac{1}{2}} \\
&\geq \mathbb{E}_\tau\left[\mathcal{Y}^{(\tau_\varepsilon,\xi+u)}_{\tau_\varepsilon} - \delta_\tau(u)\varepsilon^{-1}\int_\tau^{\tau_\varepsilon} H^\varepsilon_s ds\right] - C_L(\varepsilon^{\frac{1}{2}} + \mathbb{E}_\tau[|H^\varepsilon_{\tau_\varepsilon} - 1|])
\end{aligned}
$$

where $\varepsilon^{-1}\int_\tau^{\tau_\varepsilon} H^\varepsilon_s ds$ and $H^\varepsilon_{\tau_\varepsilon}$ converge to 1 in $\mathbf{L}_1$, so that

$$
\mathcal{Y}^\zeta_\tau \geq \liminf_{\varepsilon\to 0}\mathbb{E}_\tau\left[\mathcal{Y}^{(\tau_\varepsilon,\xi+u)}_{\tau_\varepsilon}\right] - \delta_\tau(u).
$$

It remains to show that

$$
\liminf_{\varepsilon\to 0}\mathbb{E}_\tau\left[\mathcal{Y}^{(\tau_\varepsilon,\xi+u)}_{\tau_\varepsilon}\right] \geq \mathcal{E}^{(\tau,\xi+u),\nu}_{\tau,T}[\hat{g}(X^{(\tau,\xi+u),\nu}_T)]. \tag{3.6}
$$

Then, the arbitrariness of $\nu \in \mathcal{U}$ will allow one to conclude by appealing to Proposition 3.1. In order to alleviate notations, we set $\hat{Y}^{\tau_1,\xi_1}_{\tau_2} := \mathcal{E}^{(\tau_1,\xi_1),\nu}_{\tau_2,T}[\hat{g}(X^{(\tau_1,\xi_1),\nu}_T)]$ for any $\tau_1,\tau_2 \in \mathcal{T}$ and $\xi_1 \in \mathbf{L}_2(\mathcal{F}_{\tau_1})$. By Proposition 3.1,

$$
\mathbb{E}_\tau[\mathcal{Y}^{(\tau_\varepsilon,\xi+u)}_{\tau_\varepsilon}] \geq \mathbb{E}_\tau[\hat{Y}^{\tau_\varepsilon,\xi+u}_{\tau_\varepsilon}] = \hat{Y}^{\tau,\xi+u}_\tau + \mathbb{E}_\tau[\hat{Y}^{\tau_\varepsilon,\xi+u}_{\tau_\varepsilon}] - \hat{Y}^{\tau,\xi+u}_\tau.
$$

We now observe that

$$
\begin{aligned}
\mathbb{E}_\tau[\hat{Y}^{\tau_\varepsilon,\xi+u}_{\tau_\varepsilon}] - \hat{Y}^{\tau,\xi+u}_\tau &= \mathbb{E}_\tau[\hat{Y}^{\tau_\varepsilon,\xi+u}_{\tau_\varepsilon} - \hat{Y}^{\tau_\varepsilon,X^{(\tau,\xi+u),\nu}_{\tau_\varepsilon}}_{\tau_\varepsilon}] \\
&\quad + \mathbb{E}_\tau[\hat{Y}^{\tau_\varepsilon,X^{(\tau,\xi+u),\nu}_{\tau_\varepsilon}}_{\tau_\varepsilon}] - \mathcal{E}^{(\tau,\xi+u),\nu}_{\tau,\tau_\varepsilon}[\hat{Y}^{\tau_\varepsilon,X^{(\tau,\xi+u),\nu}_{\tau_\varepsilon}}_{\tau_\varepsilon}]
\end{aligned}
$$

in which, by Lemma A.2 combined with (2.2)-(2.3)-(2.12),

$$
\lim_{\varepsilon\to 0}\mathbb{E}_\tau[\hat{Y}^{\tau_\varepsilon,\xi+u}_{\tau_\varepsilon} - \hat{Y}^{\tau_\varepsilon,X^{(\tau,\xi+u),\nu}_{\tau_\varepsilon}}_{\tau_\varepsilon}] = 0.
$$

By the same assumptions combined with Lemma A.1,

$$
\lim_{\varepsilon\to 0}\mathbb{E}_\tau[\hat{Y}^{\tau_\varepsilon,X^{(\tau,\xi+u),\nu}_{\tau_\varepsilon}}_{\tau_\varepsilon}] - \mathcal{E}^{(\tau,\xi+u),\nu}_{\tau,\tau_\varepsilon}[\hat{Y}^{\tau_\varepsilon,X^{(\tau,\xi+u),\nu}_{\tau_\varepsilon}}_{\tau_\varepsilon}] = 0.
$$

This proves (3.6) and completes the proof. □

3.4 Stability in time

We now turn to the proof of the stability in time. The lower estimate trivially follows from the dynamic programming principle of Proposition 3.2. The second one is much more delicate. It is obtained by a suitable use of the face-lifting phenomenon observed in Proposition 3.3. It allows one to absorb the singularity due to the control when passing to the supremum over $\mathcal{U}$, see (3.8) below.

Proposition 3.4. *For all $\zeta \in \mathbf{D}_2$ and $\vartheta \in \mathcal{T}_{\tau_\zeta}$*

$$
\left|\mathcal{Y}^\zeta_{\tau_\zeta} - \mathbb{E}_{\tau_\zeta}[\mathcal{Y}^{(\vartheta,\xi_\zeta)}_\vartheta]\right| \leq C_L\,\mathbb{E}_{\tau_\zeta}[\vartheta - \tau_\zeta]^{\frac{1}{2}}.
$$

Proof. 1. By Proposition 3.2, Remark 3.1, (2.3) and Lemma A.1

$$\mathcal{Y}^{\zeta}_{\tau_\zeta} \geq \mathcal{E}^{\zeta,0}_{\tau_\zeta,\vartheta}[\mathcal{Y}^{(\vartheta,X^{\zeta,0}_\vartheta)}_\vartheta] \geq \mathbb{E}_{\tau_\zeta}[\mathcal{Y}^{(\vartheta,X^{\zeta,0}_\vartheta)}_\vartheta] - C_L\mathbb{E}_{\tau_\zeta}[\vartheta-\tau_\zeta]^{\frac{1}{2}},$$

while Corollary 3.1 and (2.2) imply

$$\mathbb{E}_{\tau_\zeta}[\mathcal{Y}^{(\vartheta,X^{\zeta,0}_\vartheta)}_\vartheta] \geq \mathbb{E}_{\tau_\zeta}[\mathcal{Y}^{(\vartheta,\xi_\zeta)}_\vartheta] - C_L\mathbb{E}_{\tau_\zeta}[\vartheta-\tau_\zeta]^{\frac{1}{2}}.$$

2. We now turn to the reverse inequality. Set $\zeta=(\tau,\xi)$ and let $U^{\zeta,\nu} := \beta Y^{\zeta,\nu}_{\cdot\vee\tau} - \int_\tau^{\cdot\vee\tau}\beta_s\delta_s(\nu_s)ds$, where $\beta_s := e^{L(s-\tau)^+}$, recall (2.3). Then,

$$U^{\zeta,\nu}_{t_1} \;=\; U^{\zeta,\nu}_{t_2} + \int_{\tau\vee t_1}^{\tau\vee t_2}\beta_s\left(f_s(X^{\zeta,\nu}_s,Y^{\zeta,\nu}_s,Z^{\zeta,\nu}_s) - LY^{\zeta,\nu}_s\right)ds - \int_{\tau\vee t_1}^{\tau\vee t_2}\beta_s Z^{\zeta,\nu}_s dW_s$$

for $t_2 \geq t_1$. Since $y\in\mathbb{R}\mapsto f(\cdot,y,\cdot)-Ly$ is non-increasing by (2.3), and $\delta\geq 0$ by (2.7), it follows that

$$U^{\zeta,\nu}_{t_1} \leq U^{\zeta,\nu}_{t_2} + \int_{\tau\vee t_1}^{\tau\vee t_2}\left(\beta_s f_s(X^{\zeta,\nu}_s,\beta_s^{-1}U^{\zeta,\nu}_s,Z^{\zeta,\nu}_s) - LU^{\zeta,\nu}_s\right)ds - \int_{\tau\vee t_1}^{\tau\vee t_2}\beta_s Z^{\zeta,\nu}_s dW_s. \quad (3.7)$$

On the other hand, we can use (3.3), the fact that $\delta\geq 0$ is sublinear and $\beta\geq 1$, (2.10), Remark 3.1 and Proposition 3.3 to deduce that

$$\begin{aligned} U^{\zeta,\nu}_\vartheta &\leq \beta_\vartheta\mathcal{Y}^{(\vartheta,X^{\zeta,\nu}_\vartheta)}_\vartheta - \int_\tau^\vartheta\beta_s\delta_s(\nu_s)ds \\ &\leq C_L(\beta_\vartheta-1) + \mathcal{Y}^{(\vartheta,X^{\zeta,\nu}_\vartheta)}_\vartheta - \delta_\vartheta\left(\int_\tau^\vartheta\nu_s ds\right) \\ &\leq C_L(\beta_\vartheta-1) + \mathcal{Y}^{(\vartheta,X^{\zeta,\nu}_\vartheta-\int_\tau^\vartheta\nu_s ds)}_\vartheta. \end{aligned} \quad (3.8)$$

The last inequality combined with (3.7), Lemma A.1 and (2.3) leads to

$$Y^{\zeta,\nu}_\tau = U^{\zeta,\nu}_\tau \;\leq\; \mathbb{E}_\tau\left[\hat{H}^\tau_\vartheta\left(C_L(\beta_\vartheta-1) + \mathcal{Y}^{(\vartheta,X^{\zeta,\nu}_\vartheta-\int_\tau^\vartheta\nu_s ds)}_\vartheta + C_L\,(\vartheta-\tau)\right)\right]$$

in which

$$\hat{H}^\tau_\vartheta := \sup_{[\tau,\vartheta]} e^{\int_\tau^\cdot(\kappa^1_s-2^{-1}|\kappa^2_s|^2)ds+\int_\tau^\cdot\kappa^2_s dW_s}$$

for some predictable processes κ^1,κ^2 that are uniformly bounded by a constant which only depends on L. We next use Corollary 3.1 together with standard estimates, recall (2.2), to deduce from the above that

$$\begin{aligned} Y^{\zeta,\nu}_\tau - \mathbb{E}_\tau\left[\hat{H}^\tau_\vartheta\mathcal{Y}^{(\vartheta,\xi)}_\vartheta\right] &\leq C_L\left(\mathbb{E}_\tau\left[\hat{H}^\tau_\vartheta\,|X^{\zeta,\nu}_\vartheta - \int_\tau^\vartheta\nu_s ds - \xi|\right] + \mathbb{E}_\tau\left[(\vartheta-\tau)\right]^{\frac{1}{2}}\right) \\ &\leq C_L\,\mathbb{E}_\tau\left[(\vartheta-\tau)\right]^{\frac{1}{2}}. \end{aligned}$$

Then, Remark 3.1 implies that

$$\mathbb{E}_\tau\left[\hat{H}^\tau_\vartheta\mathcal{Y}^{(\vartheta,\xi)}_\vartheta\right] - \mathbb{E}_\tau\left[\mathcal{Y}^{(\vartheta,\xi)}_\vartheta\right] \leq C_L\mathbb{E}_\tau\left[|\hat{H}^\tau_\vartheta-1|\right] \leq C_L\mathbb{E}_\tau\left[(\vartheta-\tau)\right]^{\frac{1}{2}}.$$

Combining the two last inequalities and using the arbitrariness of $\nu \in \mathcal{U}$ leads to the required result. $\square$

For later use, we state the following corollary of Proposition 3.4 and Corollary 3.1, recall (2.2).

Corollary 3.2. *For all $\zeta \in \mathbf{D}_2$ and $\vartheta \in \mathcal{T}_{\tau_\zeta}$*

$$\left| \mathcal{Y}^{\zeta}_{\tau_\zeta} - \mathbb{E}_{\tau_\zeta}[\mathcal{Y}^{(\vartheta, X^\zeta_\vartheta)}_{\vartheta}] \right| \leq C_L \, \mathbb{E}_{\tau_\zeta}[\vartheta - \tau_\zeta]^{\frac{1}{2}}.$$

3.5 Path continuity and boundary limit

The following is deduced from Proposition 3.1 and Corollary 3.2.

Proposition 3.5. *Fix $\zeta \in \mathbf{D}_2$ such that $\tau_\zeta < T$. Then, $\mathcal{Y}^\zeta$ is continuous on $[0, T)$ and satisfies*

$$\mathcal{Y}^\zeta_{T-} := \lim_{s \uparrow T, s < T} \mathcal{Y}^\zeta_s = \hat{g}(X^\zeta_T) \ .$$

In particular, the process $\mathcal{Y}^\zeta \mathbf{1}_{[0,T)} + \hat{g}(X^\zeta_T)\mathbf{1}_{\{T\}}$ is continuous.

Proof. 1. We first show that $\mathcal{Y}^\zeta$ is right-continuous. Fix $\vartheta \in \mathcal{T}_{\tau_\zeta}$ and let $(\vartheta_n)_{n\geq 1} \subset \mathcal{T}_\vartheta$ be such that $\vartheta_n \downarrow \vartheta$. Since $\mathcal{Y}^\zeta_\vartheta = \mathcal{Y}^{(\vartheta, X^\zeta_\vartheta)}_\vartheta$, $\mathcal{Y}^\zeta_{\vartheta_n} = \mathcal{Y}^{(\vartheta_n, X^\zeta_{\vartheta_n})}_{\vartheta_n}$, and $X^\zeta_{\vartheta_n} = X^{(\vartheta, X^\zeta_\vartheta)}_{\vartheta_n}$, it follows from Corollary 3.2 applied to the time intervalle $[\vartheta, \vartheta_n]$ that

$$\mathcal{Y}^\zeta_\vartheta = \lim_{n\to\infty} \mathbb{E}_\vartheta[\mathcal{Y}^\zeta_{\vartheta_n}]. \tag{3.9}$$

We claim that $\limsup_{n\to\infty} \mathcal{Y}^\zeta_{\vartheta_n} = \liminf_{n\to\infty} \mathcal{Y}^\zeta_{\vartheta_n}$ a.s. Then, the above combined with Remark 3.1 and the dominated converge theorem implies that

$$\mathcal{Y}^\zeta_\vartheta = \lim_{n\to\infty} \mathcal{Y}^\zeta_{\vartheta_n}.$$

It remains to prove our claim. Let us first define

$$\bar{\eta} := \limsup_{n\to\infty} \mathcal{Y}^\zeta_{\vartheta_n} \ \text{ and } \ \underline{\eta} := \liminf_{n\to\infty} \mathcal{Y}^\zeta_{\vartheta_n}.$$

Assume on the contrary that the set

$$A := \{\bar{\eta} > \underline{\eta}\} \ \text{ has positive probability.}$$

Note that $A \in \mathcal{F}_\vartheta$, by right-continuity of the filtration. Then, define $\bar{k}_1 = \underline{k}_1 = 1$ and

$$\begin{aligned} \bar{k}_{n+1} &:= \min\{k > \bar{k}_n : \mathcal{Y}^\zeta_{\vartheta_k} \geq 2\bar{\eta}/3 + \underline{\eta}/3\} \\ \underline{k}_{n+1} &:= \min\{k > \underline{k}_n : \mathcal{Y}^\zeta_{\vartheta_k} \leq \bar{\eta}/3 + 2\underline{\eta}/3\} \end{aligned}$$

for $k \geq 1$, and set $\bar{\vartheta}_n := \vartheta_{\bar{k}_n}\mathbf{1}_A + \vartheta_n \mathbf{1}_{A^c}$ and $\underline{\vartheta}_n := \vartheta_{\underline{k}_n}\mathbf{1}_A + \vartheta_n \mathbf{1}_{A^c}$ for $n \geq 1$. It follows from the definition of A that $\bar{\vartheta}_n, \underline{\vartheta}_n$ are well-defined. They decrease to ϑ. Applying (3.9) to the two sequences $(\bar\vartheta_n)_{n\geq 1}$ and $(\underline{\vartheta}_n)_{n\geq 1}$ and recalling the uniform bound of Remark 3.1 leads to

$$(2\bar\eta/3+\underline{\eta}/3)\mathbf{1}_A \leq \lim_{n\to\infty}\mathbb{E}_\vartheta[\mathcal{Y}^\zeta_{\bar\vartheta_n}]\mathbf{1}_A = \mathcal{Y}^\zeta_\vartheta\mathbf{1}_A = \lim_{n\to\infty}\mathbb{E}_\vartheta[\mathcal{Y}^\zeta_{\underline{\vartheta}_n}]\mathbf{1}_A \leq (\bar\eta/3+2\underline{\eta}/3)\mathbf{1}_A,$$

a contradiction.

2. It follows from Proposition 3.2 that $\mathcal{Y}^\zeta$ is a $f^{\zeta,0}$-supermartingale in the strong sense in the terminology of [10]. Since it is right-continuous, it follows from [10, Theorem 3.3] that it admits left-limits. It is clear from Corollary 3.2 that it can not have jumps on $[0,T)$.

3. We finally prove the limit behavior at T. It follows from Proposition 3.1, Lemma A.1, (2.12) and the fact that $\mathcal{Y}^\zeta$ is làg without positive jumps, see 1. and 2. above, that $\mathcal{Y}^\zeta_{T-}\geq \hat g(X^\zeta_T)$. On the other hand, since $g(X^\zeta_T)\leq \hat g(X^\zeta_T)$ and $\hat g(X^\zeta_T+u)-\delta_T(u)\leq \hat g(X^\zeta_T)$ by construction and the fact that δ_T is sub-linear, we can follow the arguments of the proof of Proposition 3.4 to deduce that

$$\mathcal{Y}^\zeta_{\tau_\zeta\vee(T-\varepsilon)} \quad\leq\quad \mathbb{E}_{\tau_\zeta\vee(T-\varepsilon)}\left[\hat g(X^\zeta_T)\right]+C_L\,\varepsilon^{\frac{1}{2}}.$$

By continuity of the filtration, this implies that $\mathcal{Y}^\zeta_{T-}\leq \hat g(X^\zeta_T)$. □

4 Weak versus strong formulation of the dual problem

The aim of this section is to prove that $\mathcal{Y}^\zeta$ as defined in (3.3) actually provides the minimal super-solution of $\mathrm{BSDE}_K(f,g,\zeta)$. Then, the statements of Theorem 2.1 will be a consequence of the results obtained in Section 3. To this purpose, we first introduce the weak formulation associated to the optimal control problem (3.3) and show that the value coincides. Then, we use standard arguments to show that this weak formulation actually provides the minimal super-solution of our constrained backward stochastic differential equation.

4.1 Weak formulation

Given $\nu\in\mathcal{U}$ and $\zeta\in\mathbf{D}_2$, we define the equivalent probability measure $\mathbb{P}^{\zeta,\nu}$ by

$$\frac{d\mathbb{P}^{\zeta,\nu}}{d\mathbb{P}} = e^{-\frac{1}{2}\int_{\tau_\zeta}^T|\sigma_s^{-1}(X^\zeta_s)\nu_s|^2ds+\int_{\tau_\zeta}^T\sigma_s^{-1}(X^\zeta_s)\nu_s dW_s}.$$

Recall that σ^{-1} is bounded by assumption. Then,

$$W^{\zeta,\nu}:=W-\int_{\tau_\zeta}^{\tau_\zeta\vee\cdot}\sigma_s^{-1}(X^\zeta_s)\nu_s ds \tag{4.1}$$

is a $\mathbb{P}^{\zeta,\nu}$-Brownian motion.

Given $\vartheta\in\mathcal{T}_{\tau_\zeta}$ and $G\in\mathbf{L}_2(\mathbb{P}^{\zeta,\nu},\mathcal{F}_\vartheta)$, we set

$$\tilde{\mathcal{E}}^{\zeta,\nu}_{t,\vartheta}[G]:=U_{t\vee\vartheta}$$

in which $(U, V) \in \mathbf{S}_2(\mathbb{P}^{\zeta,\nu}) \times \mathbf{H}_2(\mathbb{P}^{\zeta,\nu})$ satisfies

$$U_{t\vee\tau_\zeta} = G + \int_{t\vee\tau_\zeta}^{\vartheta} \Big[f_s(X^\zeta_s, U_s, V_s) - \delta_s(\nu_s) \Big] ds - \int_{t\vee\tau_\zeta}^{\vartheta} V_s dW^{\zeta,\nu}_s, \ t \le T. \tag{4.2}$$

We finally define for $\tau \in \mathcal{T}_{\tau_\zeta}$

$$\tilde{\mathcal{Y}}^\zeta_\tau := \operatorname{ess\,sup}\{\tilde{\mathcal{E}}^{\zeta,\nu}_{\tau,T}[g(X^\zeta_T)] : \ \nu \in \mathcal{U}, \ \nu \mathbf{1}_{[\![0,\tau]\!]} \equiv 0\} \ . \tag{4.3}$$

4.2 Equivalence of the strong and weak formulations

Proposition 4.1. $\tilde{\mathcal{Y}}^\zeta_{\tau_\zeta} = \mathcal{Y}^\zeta_{\tau_\zeta}$, *for each* $\zeta \in \mathbf{D}_2$.

Proof. We write $\zeta = (\tau, \xi)$. For sake of simplicity, we restrict to the situation $\tau = 0$ so that $x_0 := \xi \in \mathbb{R}^d$. The general case is obtained by a conditioning argument. Let $\mathcal{U}^{\rm simple}$ denote the set of processes $\nu \in \mathcal{U}$ of the form

$$\nu = \sigma(X^{\zeta,\nu}) \sum_{i=0}^{n-1} \phi_i \mathbf{1}_{(t_i,t_{i+1}]} \tag{4.4}$$

in which $0 = t_0 < \cdots < t_n = T$ and $\phi_i \in \mathbf{L}_\infty(\mathcal{F}_{t_i})$ for $i \le n$. Note that $(\nu, X^{\zeta,\nu})$ is well-defined for any $(\phi_i)_{i\le n} \subset \mathbf{L}_\infty$ satisfying the previous measurability condition. This follows from (2.2).

We define accordingly $\tilde{\mathcal{U}}^{\rm simple}$ as the set of processes $\nu \in \mathcal{U}$ of the form

$$\nu = \sigma(X^{\zeta}) \sum_{i=0}^{n-1} \phi_i \mathbf{1}_{(t_i,t_{i+1}]} \tag{4.5}$$

in which $0 = t_0 < \cdots < t_n = T$ and $\phi_i \in \mathbf{L}_\infty(\mathcal{F}_{t_i})$ for $i \le n$.

1. We first show that for each $\nu \in \mathcal{U}^{\rm simple}$ we can find $\tilde{\nu} \in \mathcal{U}$ such that

$$\mathcal{E}^{\zeta,\nu}_{\tau,T}[g(X^{\zeta,\nu}_T)] = \tilde{\mathcal{E}}^{\zeta,\tilde{\nu}}_{\tau,T}[g(X^\zeta_T)].$$

Let $\nu \in \mathcal{U}^{\rm simple}$ be as in (4.4) and note that we can identify ϕ_i to a Borel measurable map $\omega \in \Omega \mapsto \phi_i(\omega) = \phi_i(\omega_{\cdot\wedge t_i})$, up to $\mathbb{P}$-null sets. Let us define $\tilde{\nu}$ by

$$\tilde{\nu}(\omega) = \sigma(X^\zeta)\phi_i\left(\omega^\phi_{\cdot\wedge t_i}\right) \quad \text{on } (t_i, t_{i+1}]$$

where ω^ϕ is defined recursively by

$$\omega^\phi_s := \omega_s - \sum_{k=0}^{j-1}(t_{k+1} - t_k)\phi_k(\omega^\phi_{\cdot\wedge t_k}) - (s - t_j)\phi_j(\omega^\phi_{\cdot\wedge t_j}) \ \text{ for } s \in (t_j, t_{j+1}],$$

with $\omega^\phi_0 = 0$. Then, for $t \in (t_i, t_{i+1}]$,

$$X^{\zeta,\nu}_t = X^{\zeta,\nu}_{t_i} + \int_{t_i}^t (b_s(X^{\zeta,\nu}_s) + \sigma_s(X^{\zeta,\nu}_s)\phi_i(W_{\cdot\wedge t_i}))ds + \int_{t_i}^t \sigma_s(X^{\zeta,\nu}_s)dW_s$$

where W is a Brownian motion under $\mathbb{P}$, while

$$X_t^\zeta = X_{t_i}^\zeta + \int_{t_i}^t (b_s(X_s^\zeta) + \sigma_s(X_s^\zeta)\phi_i(W^{\tilde\nu}_{\cdot\wedge t_i}))ds + \int_{t_i}^t \sigma_s(X_s^\zeta)dW_s^{\tilde\nu}$$

where $W^{\tilde\nu}$ is a Brownian motion under $\mathbb{P}^{\zeta,\tilde\nu}$. This implies that the law of $(X^{\zeta,\nu},\nu)$ under $\mathbb{P}$ is the same as the law of $(X^\zeta,\tilde\nu)$ under $\mathbb{P}^{\zeta,\tilde\nu}$. In view of Lemma A.4, $\mathcal{E}^{\zeta,\nu}_{0,T}[g(X_T^{\zeta,\nu})]$ and $\tilde{\mathcal{E}}^{\zeta,\tilde\nu}_{0,T}[g(X_T^\zeta)]$ can be approximated by the same sequence of real numbers and are therefore equal.

2. The fact that for each $\tilde\nu \in \tilde{\mathcal{U}}^{\text{simple}}$ we can find $\nu \in \mathcal{U}$ such that

$$\mathcal{E}^{\zeta,\nu}_{\tau,T}[g(X_T^{\zeta,\nu})] = \tilde{\mathcal{E}}^{\zeta,\tilde\nu}_{\tau,T}[g(X_T^\zeta)]$$

follows from similar arguments.

3. To conclude the proof it remains to show that

$$\mathcal{Y}^\zeta_\tau = \sup\{\mathcal{E}^{\zeta,\nu}_{\tau,T}[g(X_T^{\zeta,\nu})],\ \nu \in \mathcal{U}^{\text{simple}}\} \ \text{ and } \ \tilde{\mathcal{Y}}^\zeta_{\tau_\zeta} = \sup\{\tilde{\mathcal{E}}^{\zeta,\tilde\nu}_{\tau,T}[g(X_T^\zeta)],\ \tilde\nu \in \tilde{\mathcal{U}}^{\text{simple}}\}.$$

We only prove the first identity, the second one being derived similarly. One inequality is trivial. Conversely, given any predictable and bounded process ϕ, we can find a bounded sequence of simple adapted processes $(\phi^n)_{n\ge1}$ such that $\mathbb{E}[\int_0^T |\phi_s^n - \phi_s|^2 ds] \to 0$. By (2.2)-(2.8), $\nu^n := \sigma(X^{\zeta,\nu^n})\phi^n \in \mathcal{U}$. In particular, it follows from (2.2) that X^{ζ,ν_n} converges in $\mathbf{S}_2$ to $X^{\zeta,\nu}$ in which $\nu := \sigma(X^{\zeta,\nu})\phi$. Hence, after possibly passing to a subsequence,

$$\liminf_{n\to\infty} g(X_T^{\zeta,\nu_n}) \ge g(X_T^{\zeta,\nu})$$

since g is assumed to be lower-semicontinuous. By the comparison principle, Lemma A.2, we have

$$\mathcal{E}^{\zeta,\nu^n}_{\tau,T}[g(X_T^{\zeta,\nu^n})] \ge \mathcal{E}^{\zeta,\nu^n}_{\tau,T}[g(X_T^{\zeta,\nu}) \wedge g(X_T^{\zeta,\nu_n})]$$

in which $g(X_T^{\zeta,\nu}) \wedge g(X_T^{\zeta,\nu_n}) \to g(X_T^{\zeta,\nu})$ a.s. and in $\mathbf{L}_2$ by dominated convergence, recall (2.4). We also have

$$\mathbb{E}\left[\int_\tau^T |\delta_s(\nu_s) - \delta_s(\nu_s^n)|^2 ds\right] \to 0$$

since $(\delta_t)_{t\le T}$ is equi-Lipschitz by (2.8). Then, Lemma A.2 implies that

$$\liminf_{n\to\infty} \mathcal{E}^{\zeta,\nu^n}_{\tau,T}[g(X_T^{\zeta,\nu^n})] \ge \lim_{n\to\infty} \mathcal{E}^{\zeta,\nu^n}_{\tau,T}[g(X_T^{\zeta,\nu}) \wedge g(X_T^{\zeta,\nu_n})] = \mathcal{E}^{\zeta,\nu}_{\tau,T}[g(X_T^{\zeta,\nu})].$$

□

4.3 Connection with the reflected backward stochastic differential equation

We now show that $\tilde{\mathcal{Y}}^\zeta$ identifies as the first component of the minimal super-solution of the backward stochastic differential equation with constraint $\mathrm{BSDE}_K(f,g,\zeta)$.

Theorem 4.1. *For all $\zeta \in \mathbf{D}_2$, there exists $\tilde{\mathcal{Z}}^\zeta \in \mathbf{H}_2$ such that $(\tilde{\mathcal{Y}}^\zeta, \tilde{\mathcal{Z}}^\zeta)$ is the minimal supersolution of* $\mathrm{BSDE}_K(f,g,\zeta)$.

Proof. The proof is standard and written in the spirit of [6, Proof of Proposition 2.5].
1. Similar arguments as in the proof of Proposition 3.2 show that $\tilde{\mathcal{Y}}$ satisfies a dynamic programming principle: for all $\vartheta_1 \le \vartheta_2 \in \mathcal{T}$, such that $\tau_\zeta \le \vartheta_1$, we have

$$\tilde{\mathcal{Y}}^{\zeta}_{\vartheta_1} = \operatorname*{ess\,sup}_{\nu\in\mathcal{U}} \tilde{\mathcal{E}}^{\zeta,\nu}_{\vartheta_1,\vartheta_2}[\tilde{\mathcal{Y}}^{\zeta}_{\vartheta_2}]. \tag{4.6}$$

We also observe that $\tilde{\mathcal{Y}}^{\zeta}$ is càd. This follows from Proposition 4.1 and Proposition 3.5. Then, the non linear Doob-Meyer decomposition of [10, Theorem 3.3] implies the existence of $\tilde{\mathcal{Z}}^{\zeta,\nu} \in \mathbf{H}_2(\mathbb{P}^{\nu,\zeta})$ and of a càdlàg non-decreasing adapted process $\tilde{\mathcal{K}}^{\zeta,\nu}$ such that

$$\tilde{\mathcal{Y}}^{\zeta}_{\vartheta} = g(X^{\zeta}_T) + \int_{\vartheta}^{T} \left(f_s(X^{\zeta}_s, \tilde{\mathcal{Y}}^{\zeta}_s, \tilde{\mathcal{Z}}^{\zeta,\nu}_s) - \delta_s(\nu_s)\right) ds - \int_{\vartheta}^{T} \tilde{\mathcal{Z}}^{\zeta,\nu}_s dW^{\zeta,\nu}_s + \tilde{\mathcal{K}}^{\zeta,\nu}_T - \tilde{\mathcal{K}}^{\zeta,\nu}_{\vartheta}\ ,\ \vartheta \ge \tau_\zeta\ .$$

By identification of the Itô decomposition under each $\mathbb{P}^{\zeta,\nu}$, we obtain $\tilde{\mathcal{Z}}^{\zeta,\nu} = \tilde{\mathcal{Z}}^{\zeta,0} =: \tilde{\mathcal{Z}}^{\zeta}$. Moreover, (4.1) implies that for any $\nu \in \mathcal{U}$ and $\vartheta \in \mathcal{T}_{\tau_\zeta}$

$$\tilde{\mathcal{K}}^{\zeta,0}_T - \tilde{\mathcal{K}}^{\zeta,0}_{\vartheta} = \tilde{\mathcal{K}}^{\zeta,\nu}_T - \tilde{\mathcal{K}}^{\zeta,\nu}_{\vartheta} + \int_{\vartheta}^{T} (\tilde{\mathcal{Z}}^{\zeta}_s \sigma^{-1}(X^{\zeta}_s)\nu_s - \delta_s(\nu_s))ds \ge \int_{\vartheta}^{T} (\tilde{\mathcal{Z}}^{\zeta}_s \sigma^{-1}(X^{\zeta}_s)\nu_s - \delta_s(\nu_s))ds\ ,$$

since $\tilde{\mathcal{K}}^{\zeta,\nu}$ is non-decreasing. By using a similar measurable selection argument as in the proof of Proposition 3.1, this shows that

$$\inf_{|u|=1} (\delta(u) - \tilde{\mathcal{Z}}^{\zeta}\sigma^{-1}(X^{\zeta})u) \ge 0 \ \ dt\otimes d\mathbb{P}\text{-a.e.} \ \Leftrightarrow\ \tilde{\mathcal{Z}}^{\zeta}\sigma^{-1}(X^{\zeta}) \in K \ \ dt\otimes d\mathbb{P}\text{-a.e.},$$

see e.g. [11]. Since $\tilde{\mathcal{K}}^{\zeta,0}$ is non-decreasing, writing the above for $\nu = 0$ implies that $(\tilde{\mathcal{Y}}^{\zeta}, \tilde{\mathcal{Z}}^{\zeta})$ is a super-solution of $\mathrm{BSDE}_K(f, g, \zeta)$.

2. We now prove the minimality property. If (U, V) is a super-solution of $\mathrm{BSDE}_K(f, g, \zeta)$, then it follows from the definition of δ and (4.1) that it is also a super-solution of (4.2) with $G = g(X^{\zeta}_T)$, for each $\nu \in \mathcal{U}$. In particular, $U \ge \tilde{\mathcal{E}}^{\zeta,\nu}_{\cdot,T}[g(X^{\zeta}_T)]$, and we conclude by arbitrariness of $\nu \in \mathcal{U}$. □

5 Possible extensions

In order to focus on the main ideas, we have restricted ourselves to a rather stringent framework. Some of the conditions used in this paper can certainly be weakened on a case by case basis. We discuss here some straightforward extensions or variations.

5.1 Invertibility condition

We have assumed that σ is invertible but all our arguments go through if we add a component X^o to X which has a dynamic of the form

$$dX^o_t = b^o_t(X^o_t, X_t)dt$$

with b^o Lipschitz and bounded in space, uniformly in time. Then $X^{\zeta,\nu}$ has to be replaced by $\bar{X}^{\zeta,\nu} = (X^{o,\zeta,\nu}, X^{\zeta,\nu})$ with dynamics

$$d\bar{X}^{\zeta,\nu}_t = \begin{pmatrix} b^o_t(\bar{X}^{\zeta,\nu}_t) \\ b_t(\bar{X}^{\zeta,\nu}_t) + \nu_t \end{pmatrix} dt + \begin{pmatrix} 0 \\ \sigma_t(\bar{X}^{\zeta,\nu}_t) \end{pmatrix} dW_t.$$

The face-lift of g is defined accordingly

$$\hat{g}(x^o, x) := \sup_{u \in \mathbb{R}^d} \left(g(x^o, x + u) - \delta_T(u) \right),$$

and so on.

The case of a general non-invertible coefficient σ can be treated along the lines of [3], in which it is explained how the face-lift should then be performed.

5.2 Direct constraint on the gains-process

The contraint (2.6) is motivated by financial applications in which the component V can be interpreted as the number of risky assets X held in an hedging portfolio for the contingent claim $g(X_T^\zeta)$, see [4]. It can be replaced by

$$V \in K \;\; dt \otimes d\mathbb{P}-\text{a.e} \;\; \text{on} \;\; [\![\tau_\zeta, T]\!], \tag{5.1}$$

when σ does not depend on x.

In this case, $X^{\zeta,\nu}$ must be taken of the form

$$dX_t^{\zeta,\nu} = \Big(b_t(X_t^{\zeta,\nu}) + \sigma_t \nu_t \Big) \, dt + \sigma_t dW_t.$$

If one assumes that $t \mapsto \sigma_t$ is right-continuous on $[0, T)$ and left-continuous at T, then (2.12)-(3.5) become

$$\begin{aligned} \hat{g}(x) &= \sup_{u \in \mathbb{R}^d} \left(g(x + \sigma_T u) - \delta_T(u) \right) \\ \mathcal{Y}_{\tau_\zeta}^\zeta &= \operatorname*{ess\,sup}_{u \in \mathbf{L}_\infty(\mathbb{R}^d, \mathcal{F}_{\tau_\zeta})} \left(\mathcal{Y}_{\tau_\zeta}^{(\tau_\zeta, \xi_\zeta + \sigma_{\tau_\zeta} u)} - \delta_{\tau_\zeta}(u) \right) \;\; \text{a.s. on } \{\tau_\zeta < T\}. \end{aligned}$$

The change of measure for the weak formulation is

$$\frac{d\mathbb{P}^{\zeta,\nu}}{d\mathbb{P}} = e^{-\frac{1}{2} \int_{\tau_\zeta}^T |\nu_s|^2 ds + \int_{\tau_\zeta}^T \nu_s dW_s}.$$

In particular, we do not need σ to be invertible anymore.

The results of Theorem 2.1 are obtained by following step by step the arguments used in this paper up to the modifications described above.

Moreover, the boundedness condition (2.8) can then be weakened. Indeed, it can be avoided by using (2.9)-(2.10) in all our proofs, except in the proof of Proposition 4.1 in which it is used twice. First to ensure that the controls ν constructed from the families $(\phi_i)_{i \le n}$ satisfy $\delta(\nu) < \infty$. But in the case (5.1), the ν's are of the form $\sum_{i=0}^{n-1} \phi_i \mathbf{1}_{(t_i, t_{i+1}]}$. Taking $\delta_{t_i}(\phi_i) < \infty$ is then enough. It is also used in the approximation argument of Step 3, as it implies that $(\delta_t)_{t \le T}$ is equi-Lipschitz, but for the constraint (5.1), it suffices to assume, for instance, that the domain of δ_t does not depend on t, which means that the directions in which K_t is bounded do not depend on t. This is not enough when ν is of the form used in the proof of Proposition 4.1 because of the transformation through the matrix σ, unless additional assumptions are made on it.

Our arguments are not valid if σ depends on x because the coefficients driving $X^{\zeta,\nu}$ are no more Lipschitz uniformly in the control. This is crucial for Corollary 3.1.

5.3 Optimal control of constrained BSDEs

One can allow the coefficients b, σ and f to depend on an additional control α in a set $\mathcal{A}$ of predictable processes with values in a compact set $A \subset \mathbb{R}^d$. Then, all our proofs go through whenever the conditions (2.2)-(2.3) are uniform with respect to this additional control, and the coefficients are continuous in this additional variable. The arguments used in Section 3 do not change. It is the same for Section 4.3, for $\alpha \in \mathcal{A}$ given. However, a continuity assumption on the coefficients with respect to the control will be required to prove the counterpart of Proposition 4.1: the approximation by step constant processes has to be applied to (ν, α) in place of ν.

5.4 Random coefficients with delay

One can also assume that the coefficients b, σ and f are random, satisfying the usual predictability condition, whenever the conditions (2.2)-(2.3) are uniform in ω. Again, the arguments of Section 3 and Section 4.3 do not change. However, the proof of Proposition 4.1 can not be adapted unless the dependence holds with a fixed delay: there exists $\iota > 0$ such that, for all $t \leq T$, b_t, σ_t, f_t depends on ω only through $(\omega_s)_{s \leq t-\iota}$. With this condition, Steps 1. and 2. remain correct for simple processes associated to a time grid $\{t_i, i \leq n\}$ such that $\max\{t_{i+1} - t_i, i \leq n-1\} \leq \iota$. As in the optimal control case, Step 3 also requires some continuity of the coefficient in ω, e.g. uniform continuity for the usual sup-norm topology.

A Auxiliary results

We collect here some standard results that have been used all over this paper.

In this section, we denote by $\mathcal{D}_b$ the set of measurable maps $\psi : \Omega \times [0, T] \times \mathbb{R}^d \times \mathbb{R} \mapsto \mathbb{R}$ such that $(\psi_t(y, z))_{t \leq T}$ is progressively measurable for all $(y, z) \in \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ and

$$|\psi(y, z) - \psi(0)| \leq L_\psi(|y| + |z|) \text{ for all } (y, z) \in \mathbb{R} \times \mathbb{R}^d,\ dt \otimes d\mathbb{P} - a.e.$$

for some constant $L_\psi > 0$. Given $(\vartheta, G) \in \mathbf{D}_2$ and $\tau \in \mathcal{T}$ such that $\tau \leq \vartheta$, we set $\mathcal{E}^\psi_{\tau,\vartheta}(G) := U_\tau$ where $(U, V) \in \mathbf{S}_2 \times \mathbf{H}_2$ is the solution of

$$U_{t \vee \tau} = G + \int_{t \vee \tau}^{\vartheta} \psi_s(U_s, V_s) ds - \int_{t \vee \tau}^{\vartheta} V_s dW_s,\ t \in [0, T]. \tag{A.1}$$

Lemma A.1. *Fix $\psi \in \mathcal{D}_b$ and $(\vartheta, G) \in \mathbf{D}_2$. Then, for all $\tau \in \mathcal{T}$ such that $\tau \leq \vartheta$:*

(a) *We have*

$$\mathcal{E}^\psi_{\tau,\vartheta}(G) = \mathbb{E}_\tau[H^\tau_\vartheta G + \int_\tau^\vartheta H^\tau_s \psi_s(0) ds]$$

where H^τ solves

$$H^\tau = 1 + \int_\tau^\cdot \kappa^Y_t H^\tau_t dt + \int_\tau^\cdot \kappa^Z_t H^\tau_t dW_t$$

for some predictable processes κ^Y and κ^Z that are bounded by a constant which only depends on L_ψ.

In particular, if there exists a constant $c > 0$ such that $\mathbb{E}[|G|^2] \le c$ and $|\psi(0)| \le c$ $dt \otimes d\mathbb{P}$, then

$$|\mathcal{E}^{\psi}_{\tau,\vartheta}(G) - \mathbb{E}_\tau[G]| \le C\,\mathbb{E}_\tau[\vartheta - \tau]^{\frac{1}{2}},$$

for some $C > 0$ which depends only on c and L_ψ.

(b) *If $(U, V) \in \mathbf{S}_2 \times \mathbf{H}_2$ satisfies*

$$U_{t\vee\tau} \le G + \int_{t\vee\tau}^{\vartheta} \psi_s(U_s, V_s)ds - \int_{t\vee\tau}^{\vartheta} V_s dW_s,\ t \in [0, T],$$

then

$$U_\tau \le \mathbb{E}_\tau[H^\tau_\vartheta G + \int_\tau^\vartheta H^\tau_s \psi_s(0)ds]$$

where H^τ is defined as in (a).

Proof. This follows from a standard linearization argument, see e.g. [9, Proof of Theorem 1.6]. □

Lemma A.2. *Fix $\psi^1, \psi^2 \in \mathcal{D}_b$ and $(\vartheta, G_1), (\vartheta, G_2) \in \mathbf{D}_2$.*

(a) *Assume that there exists a process κ such that*

$$|\psi^1 - \psi^2|(\theta) \le \kappa \quad dt \otimes d\mathbb{P}$$

for all $\theta \in \mathbb{R} \times \mathbb{R}^d$. Then, for all $\tau \in \mathcal{T}$ such that $\tau \le \vartheta$,

$$|\mathcal{E}^{\psi^1}_{\tau,\vartheta}(G_1) - \mathcal{E}^{\psi^2}_{\tau,\vartheta}(G_2)| \le C\mathbb{E}_\tau\left[|G_1 - G_2|^2 + \int_\tau^T |\kappa_s|^2 ds\right]^{\frac{1}{2}}$$

where $C > 0$ is a constant which depends only on L_{ψ^1} and L_{ψ^2}.

(b) *Assume that $G_1 \le G_2$ and $\psi^1(\theta) \le \psi^2(\theta)$ $dt \otimes d\mathbb{P}$ for all $\theta \in \mathbb{R} \times \mathbb{R}^d$. Then, $\mathcal{E}^{\psi^1}_{\tau,\vartheta}(G^1) \le \mathcal{E}^{\psi^2}_{\tau,\vartheta}(G^2)$ for all $\tau \in \mathcal{T}$ such that $\tau \le \vartheta$.*

Proof. The first assertion follows from [9, Theorem 1.5]. The second one is [9, Theorem 1.6]. □

Lemma A.3. *Let $(G_\varepsilon)_{\varepsilon>0}$ be a family of random variable, uniformly bounded in $\mathbf{L}_1$, and let $G \in \mathbf{L}_1$ be such that $\liminf_{\varepsilon\to 0} G_\varepsilon \ge G$. Let $(\tau_\varepsilon)_{\varepsilon>0}$ be a sequence of stopping times such that $\lim_{\varepsilon\to 0} \tau_\varepsilon = \tau \in \mathcal{T}$. Then, there exists a sequence $(\varepsilon_n)_{n\ge 1} \subset (0, 1)$ such that*

$$\liminf_{n\to\infty} \mathbb{E}_{\tau_{\varepsilon_n}}[G_{\varepsilon_n}] \ge \mathbb{E}_\tau[G] \text{ and } \lim_{n\to\infty} \varepsilon_n = 0.$$

Proof. We write

$$\mathbb{E}_{\tau_\varepsilon}[G_\varepsilon] = \mathbb{E}_{\tau_\varepsilon}[G] + \mathbb{E}_{\tau_\varepsilon}[G_\varepsilon - G] \ge \mathbb{E}_{\tau_\varepsilon}[G] - \mathbb{E}_{\tau_\varepsilon}[(G_\varepsilon - G)^-].$$

The first term on the right-hand side converges a.s. to $\mathbb{E}_\tau[G]$ by the continuity of the martingales in a Brownian filtration. The second term converges in $\mathbf{L}_1$ to 0, and therefore a.s. along a subsequence. □

Lemma A.4. *Fix $\psi \in \mathcal{D}_b, G \in \mathbf{L}_2(\mathcal{F}_T)$. Set $t_i^n := iT/n$ for $i \leq n$, $n \geq 1$, and define recursively for $i = n-1, \ldots, 0$*

$$U^n_{t_i^n} = \mathbb{E}_{t_i^n}[U^n_{t^n_{i+1}} + \int_{t_i^n}^{t^n_{i+1}} \psi_s(U^n_{t_i^n}, V^n_{t_i^n})ds] \ , \ V^n_{t_i^n} = (t^n_{i+1} - t^n_i)^{-1}\mathbb{E}_{t_i^n}[U^n_{t^n_{i+1}}(W_{t^n_{i+1}} - W_{t_i^n})]$$

in which $U^n_T := G$. Then, $U^n_0 \to \mathcal{E}^{\psi}_{0,T}[G]$ as $n \to \infty$.

Proof. It suffices to repeat the argument of [2, Proof of Theorem 3.1] and observe that their estimate contained in [2, Lemma 3.2] is not needed if we are not interested by the speed of convergence. Indeed, one can simply use the fact that, if (U, V) denotes the solution of (A.1) with $\tau = 0$, then

$$\max_{1 \leq i \leq n} \mathbb{E}[\sup_{t^n_{i-1} \leq t \leq t^n_i} |U_t - U_{t^n_{i-1}}|^2] + \sum_{i=1}^n \mathbb{E}[\int_{t^n_{i-1}}^{t^n_i} |V_t - \bar{V}^n_{t^n_{i-1}}|^2 dt] \to 0,$$

in which

$$\bar{V}^n_{t^n_{i-1}} := (t^n_i - t^n_{i-1})^{-1}\mathbb{E}_{t^n_{i-1}}[\int_{t^n_{i-1}}^{t^n_i} V_t dt].$$

The convergence of the left-hand side term is standard, it follows from the continuity of the path of U which belongs to $\mathbf{S}_2$. The convergence of the second term is also clear since $\bar{V}^n$ provides the best approximation of V in $\mathbf{L}_2(dt \otimes d\mathbb{P})$ by a step constant process on $\{t^n_i, i \leq n\}$. $\square$

References

[1] D.P. Bertsekas and S.E. Shreve. *Stochastic Optimal Control : The Discrete Time Case*, volume 139. Mathematics in Science and Engineering, 1978.

[2] B. Bouchard and N. Touzi. Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations. *Stochastic Processes and their applications*, 111(2):175–206, 2004.

[3] B. Bouchard and T.N. Vu. The obstacle version of the geometric dynamic programming principle: Application to the pricing of american options under constraints. *Applied Mathematics and Optimization*, 61(2):235–265, 2010.

[4] M. Broadie, J. Cvitanić, and H.M. Soner. Optimal replication of contingent claims under portfolio constraints. *Review of Financial Studies*, 11(1):59–79, 1998.

[5] J. Cvitanić and I. Karatzas. Hedging contingent claims with constrained portfolios. *The Annals of Applied Probability*, 3(3):652–681, 1993.

[6] J. Cvitanić, I. Karatzas, and H.M. Soner. Backward stochastic differential equations with constraints on the gains-process. *Ann. Probab.*, 26(4):1522–1551, 1998.

[7] J. Cvitanić, H. Pham, and N. Touzi. Super-replication in stochastic volatility models under portfolio constraints. *Journal of Applied Probability*, 36:523–545, 1999.

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

[9] E. Pardoux. Backward stochastic differential equations and viscosity solutions of systems of semilinear parabolic and elliptic pdes of second order. In *Stochastic Analysis and Related Topics VI: Proceedings of the Sixth Oslo Silivri Workshop, Geilo 1996*, volume 42, pages 79–128. Springer, 1998.

[10] S. Peng. Monotonic limit theorem of bsde and nonlinear decomposition theorem of doob–meyers type. *Probability theory and related fields*, 113(4):473–499, 1999.

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