Minimal Supersolutions of Convex BSDEs under Constraints

Gregor Heyne[a,1,*], Michael Kupper[b,2], Christoph Mainberger[a,3], Ludovic Tangpi[b,4]

November 14, 2013

ABSTRACT

We study supersolutions of a backward stochastic differential equation, the control processes of which are constrained to be continuous semimartingales of the form $dZ = \Delta dt + \Gamma dW$. The generator may depend on the decomposition (Δ, Γ) and is assumed to be positive, jointly convex and lower semicontinuous, and to satisfy a superquadratic growth condition in Δ and Γ. We prove the existence of a supersolution that is minimal at time zero and derive stability properties of the non-linear operator that maps terminal conditions to the time zero value of this minimal supersolution such as monotone convergence, Fatou's lemma and L^1-lower semicontinuity. Furthermore, we provide duality results within the present framework and thereby give conditions for the existence of solutions under constraints.

KEYWORDS: Supersolutions of Backward Stochastic Differential Equations; Gamma Constraints; Minimality under Constraints; Duality

AUTHORS INFO

[a] *Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany*
[b] *University of Konstanz, Universitätsstr. 10, 78457 Konstanz, Germany*
[1] heyne@math.hu-berlin.de
[2] kupper@uni-konstanz.de
[3] mainberg@math.hu-berlin.de
[4] ludovic.tangpi@uni-konstanz.de

[*] *Financial support: MATHEON project E.2*

PAPER INFO

AMS CLASSIFICATION: *60H20; 60H30*

1 Introduction

On a filtered probability space, the filtration of which is generated by a d-dimensional Brownian motion, we are interested in quadruplets (Y, Z, Δ, Γ) of processes such that, for all $0 \leq s \leq t \leq T$, the system

$$\begin{aligned} Y_s - \int_s^t g_u(Y_u, Z_u, \Delta_u, \Gamma_u)du + \int_s^t Z_u dW_u \geq Y_t\,, \qquad Y_T \geq \xi\,, \\ Z_t = z + \int_0^t \Delta_u du + \int_0^t \Gamma_u dW_u \end{aligned} \tag{1.1}$$

is satisfied. Here, for ξ a terminal condition, Y is the càdlàg value process and Z the continuous control process with decomposition (Δ, Γ). The generator g is assumed to be jointly convex and may depend on the decomposition of the continuous semimartingale Z. It is our objective to give conditions ensuring that the set $\mathcal{A}(\xi, g, z)$, consisting of all admissible pairs (Y, Z) satisfying (1.1), henceforth called supersolution of the backward stochastic differential equation (BSDE) under gamma and delta constraints, contains elements $(\hat{Y}, \hat{Z})$ that are minimal at time zero. Furthermore, we give conditions relying on BSDE duality for the existence of solutions

under constraints.

Finding the minimal initial value of a supersolution under constraints is closely related to the superreplication problem in a financial market under gamma constraints, first studied in Soner and Touzi [21]. Indeed, the classical gamma constraints can be incorporated into our more general framework by setting the generator to $+\infty$ whenever the diffusion part Γ is outside a predetermined interval.

In a nutshell, inspired by the methods first used in Drapeau et al. [9] and then later in Heyne et al. [13], we begin by considering the operator $\mathcal{E}_0^g(\xi, z) := \inf\{Y_0 : (Y, Z) \in \mathcal{A}(\xi, g, z)\}$ where $z \in \mathbb{R}^{1\times d}$ is the initial value of controls. We then show that the set of supersolutions $(\hat{Y}, \hat{Z})$ satisfying $\hat{Y}_0 = \mathcal{E}_0^g(\xi, z)$ is non-empty. In order to do so, we impose a superquadratic growth condition in the decomposition parts (Δ, Γ) of controls on the generator g, reflecting a penalization of rapid changes in control values and accounting for the expression "Delta- and Gamma-Constraints". The consequence is twofold. First, it ensures that the sequence of stochastic integrals $(\int Z^n dW)$ corresponding to the minimizing sequence $Y_0^n \downarrow \mathcal{E}_0^g(\xi, z)$ is bounded in $\mathcal{H}^2$. Drawing from compactness results for the space of martingales $\mathcal{H}^2$ given in Delbaen and Schachermayer [7], we obtain our candidate control process $\hat{Z}$ as the limit of a sequence in the asymptotic convex hull of (Z^n). At this point it is crucial to preserve the continuous semimartingale structure of the limit object, that is $\hat{Z} = z + \int \hat{\Delta} du + \int \hat{\Gamma} dW$, possible by using once more the aforementioned growth condition on g.

In a next step, we provide stability results of $\xi \mapsto \mathcal{E}_0^g(\xi, z)$, the non-linear operator that maps a terminal condition to the value of the minimal supersolution at time zero, such as monotone convergence, Fatou's lemma or L^1-lower semicontinuity. This, together with convexity, gives way to a dual representation of $\mathcal{E}_0^g$ as a consequence of the Fenchel-Moreau theorem. We characterize the conjugate $\mathcal{E}_0^*$ in terms of the decomposition parts of the controls and show that $\mathcal{E}_0^*$ is always attained. In particular, for the case of a quadratic generator we show that it is possible to explicitly compute the conjugate by means of classical calculus of variations methods, giving additional structural insight into the problem. If we assume in turn the existence of an optimal subgradient such that $\mathcal{E}_0^g(\xi, z)$ is attained in its dual representation, we can prove that the associated BSDE with parameters (ξ, g) admits a solution under constraints. This extends results of Delbaen et al. [8] and Drapeau et al. [10] obtained in the unconstrained case.

Before we continue, let us briefly discuss the existing literature on the subject. Ever since the seminal paper Pardoux and Peng [18], an extensive amount of work has been done in the field of BSDEs, resulting in such important contributions as for instance El Karoui et al. [11], Kobylanski [17] or Briand and Hu [1]. We refer the reader to Peng [19] or Drapeau et al. [9] for a more thourough treatment of the literature concerning solutions and in particular supersolutions of BSDEs. There are many works dealing with optimization or (super-)replication under constraints, see for instance Cvitanic and Karatzas [5], Jouini and Kallal [14] or Broadie et al. [2] and references therein, but the notion of gamma constraints in the context of superhedging was introduced in Soner and Touzi [21]. Therein, the authors identify the superreplication cost as the solution to a variational inequality. In Cheridito et al. [3], the aforementioned problem is solved in a multi-dimensional setting and the superreplication price is characterized as the unique viscosity solution of a nonstandard partial differential equation, whereas in Cheridito et al. [4] the

authors treat the related system of BSDEs and SDEs in a more abstract fashion. Compare also the more recent work Soner et al. [22] for a dual characterization of the superreplication problem.

The remainder of this paper is organized as follows. Setting and notations are specified in Section 2. A precise definition of supersolutions under gamma and delta constraints is then given in Section 3, along with existence and stability results. We conclude this work with duality results in Section 4.

2 Setting and notations

We consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$, where the filtration $(\mathcal{F}_t)$ is generated by a d-dimensional Brownian motion W and is assumed to satisfy the usual conditions. For some fixed time horizon $T > 0$ the set of $\mathcal{F}_T$-measurable random variables is denoted by L^0, where random variables are identified in the P-almost sure sense. Let furthermore denote L^p the set of random variables in L^0 with finite p-norm, for $p \in [1, +\infty]$. Inequalities and strict inequalities between any two random variables or processes X^1 and X^2 are understood in the P-almost sure or in the $P \otimes dt$-almost everywhere sense, respectively. We denote by $\mathcal{T}$ the set of stopping times with values in $[0, T]$ and hereby call an increasing sequence of stopping times (τ^n) such that $P[\bigcup_n \{\tau^n = T\}] = 1$ a localizing sequence of stopping times. For $m, n \in \mathbb{N}$ we denote by $|\cdot|$ the Euclidean norm on $\mathbb{R}^{m\times n}$, that is $|x| = (\sum_{i,j} x_{ij}^2)^{\frac{1}{2}}$. By $\mathcal{S} := \mathcal{S}(\mathbb{R})$ we denote the set of càdlàg progressively measurable processes Y with values in $\mathbb{R}$. For $p \in [1, +\infty[$, we further denote by $\mathcal{H}^p$ the set of càdlàg local martingales M with finite $\mathcal{H}^p$-norm on $[0, T]$, that is $\|M\|_{\mathcal{H}^p} := E[\langle M, M\rangle_T^{p/2}]^{1/p} < \infty$. By $\mathcal{L}^p := \mathcal{L}^p(W)$ we denote the set of $\mathbb{R}^{1\times d}$-valued, progressively measurable processes Z such that $\int ZdW \in \mathcal{H}^p$, that is, $\|Z\|_{\mathcal{L}^p} := E[(\int_0^T |Z_s|^2\, ds)^{p/2}]^{1/p}$ is finite. For $Z \in \mathcal{L}^p$, the stochastic integral $\int ZdW$ is well defined, see [20], and is by means of the Burkholder-Davis-Gundy inequality [20, Theorem 48] a continuous martingale. We further denote by $\mathcal{L} := \mathcal{L}(W)$ the set of $\mathbb{R}^{1\times d}$-valued, progressively measurable processes Z such that there exists a localizing sequence of stopping times (τ^n) with $Z1_{[0,\tau^n]} \in \mathcal{L}^1$, for all $n \in \mathbb{N}$. For $Z \in \mathcal{L}$, the stochastic integral $\int ZdW$ is well defined and is a continuous local martingale. Finally, for a given sequence (x_n) in some convex set, we say that a sequence $(\tilde{x}_n)$ is in the asymptotic convex hull of (x_n) if $\tilde{x}_n \in conv\{x_n, x_{n+1}, \dots\}$, for all $n \in \mathbb{N}$.

3 Minimal supersolutions of BSDEs under delta and gamma constraints

3.1 Definitions

Throughout this work, a generator is a jointly measurable function g from $\Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^{1\times d} \times \mathbb{R}^{1\times d} \times \mathbb{R}^{d\times d}$ to $\mathbb{R} \cup \{+\infty\}$ where $\Omega \times [0, T]$ is endowed with the progressive σ-field. A control $Z \in \mathcal{L}$ with initial value $z \in \mathbb{R}^{1\times d}$ is said to have the decomposition (Δ, Γ) if it is of the form $Z = z + \int \Delta du + \int \Gamma dW$, for progressively measurable (Δ, Γ) taking values in

$\mathbb{R}^{1\times d} \times \mathbb{R}^{d\times d}$.[1] A control is said to be admissible if the continuous local martingale $\int Z dW$ is a supermartingale. Let us collect all these processes in the set Θ defined by

$$\Theta := \left\{ Z \in \mathcal{L} \,:\, \begin{array}{l} \text{there exists } z \in \mathbb{R}^{1\times d} \text{ and progressively measurable } (\Delta, \Gamma) \text{ such that} \\ Z = z + \int \Delta du + \int \Gamma dW \text{ and } \int Z dW \text{ is a supermartingale} \end{array} \right\}.$$

Whenever we want to stress the dependence of controls on a fixed initial value $z \in \mathbb{R}^{1\times d}$, we make use of the set $\Theta(z) := \{Z \in \Theta : Z_0 = z\}$. Given a generator g and a terminal condition $\xi \in L^0$, a pair $(Y, Z) \in \mathcal{S} \times \Theta$ is a supersolution of a BSDE under gamma and delta constraints if, for $0 \le s \le t \le T$, it holds

$$Y_s - \int_s^t g_u(Y_u, Z_u, \Delta_u, \Gamma_u) du + \int_s^t Z_u dW_u \ge Y_t \quad \text{and} \quad Y_T \ge \xi \,. \tag{3.1}$$

For a supersolution (Y, Z), we call Y the value process and Z its corresponding control process.[2] Given $z \in \mathbb{R}^{1\times d}$, we are now interested in the set

$$\mathcal{A}(\xi, g, z) := \{(Y, Z) \in \mathcal{S} \times \Theta(z) \,:\, (3.1) \text{ holds}\} \,.$$

Throughout this paper a generator g is said to be

(LSC) if $(y, z, \delta, \gamma) \mapsto g(y, z, \delta, \gamma)$ is lower semicontinuous.

(POS) positive, if $g(y, z, \delta, \gamma) \ge 0$, for all $(y, z, \delta, \gamma) \in \mathbb{R} \times \mathbb{R}^{1\times d} \times \mathbb{R}^{1\times d} \times \mathbb{R}^{d\times d}$.

(CON) convex, if $(y, z, \delta, \gamma) \mapsto g(y, z, \delta, \gamma)$ is jointly convex.

(DGC) delta- and gamma-compatible, if there exist $c_1 \in \mathbb{R}$ and $c_2 > 0$ such that, for all $(\delta, \gamma) \in \mathbb{R}^{1\times d} \times \mathbb{R}^{d\times d}$,

$$g(y, z, \delta, \gamma) \ge c_1 + c_2 \left(|\delta|^2 + |\gamma|^2 \right)$$

holds for all $(y, z) \in \mathbb{R} \times \mathbb{R}^{1\times d}$.

Remark 3.1.

(i) Note that (DGC) reflects a penalization of rapid changes in control values. In contrast to [3] or [4], where the single decomposition parts Δ and Γ were demanded to satisfy certain boundedness, continuity or growth properties, we embed this in (DGC) so that suitable $\mathcal{L}^2$-bounds emerge naturally from the problem (3.1).

[1] In order to be compatible with the dimension of Z, actually the transpose $(\int \Gamma dW)^T$ of $\int \Gamma dW$ needs to be considered. However, we suppress this operation for the remainder in order to keep the notation simple.

[2] Note that the formulation in (3.1) is equivalent to the existence of a càdlàg increasing process K, with $K_0 = 0$, such that $Y_t = \xi + \int_t^T g_u(Y_u, Z_u, \Delta_u, \Gamma_u) du + (K_T - K_t) - \int_t^T Z_u dW_u$ for all $t \in [0, T]$, see for example [11, 19].

(ii) An example of a generator that excludes values of Γ exceeding a certain level by penalization and fits into our setting is given by

$$g(y,z,\delta,\gamma) = \begin{cases} \tilde{g}(y,z,\delta) & \text{if } |\gamma| \leq M \\ +\infty & \text{else} \end{cases},$$

where $M > 0$ and $\tilde{g}$ is any positive, jointly convex and lower semicontinuous generator satisfying $\tilde{g}(y,z,\delta) \geq c_1 + c_2|\delta|^2$ for constants $c_1 \in \mathbb{R}$ and $c_2 > 0$. This particular choice of g is closely related to the kind of gamma constraints studied in [3].

(iii) Setting the generator $g(\cdot,z,\cdot,\cdot)$ equal to $+\infty$ outside a desired subset of $\mathbb{R}^{1\times d}$ shows for instance that our framework is flexible enough to comprise shortselling constraints. ⧫

3.2 General properties

The proof of the ensuing Lemma 3.2 can be found in [9, Lemma 3.2].

Lemma 3.2. *Let g be a generator satisfying (POS). Assume further that $\mathcal{A}(\xi,g) \neq \emptyset$ and that for the terminal condition ξ holds $\xi^- \in L^1$. Then $\xi \in L^1$ and, for any $(Y,Z) \in \mathcal{A}(\xi,g)$, the control Z is unique and the value process Y is a supermartingale such that $Y_t \geq E[\xi|\mathcal{F}_t]$. Moreover, the unique canonical decomposition of Y is given by*

$$Y = Y_0 + M - A\,, \tag{3.2}$$

where $M = \int Z dW$ and A is an increasing, predictable, càdlàg process with $A_0 = 0$.

The joint convexity of the generator g immediately yields the following lemma.

Lemma 3.3. *Let g be a generator satisfying (CON). Then, for each $z \in \mathbb{R}^{1\times d}$, the set $\mathcal{A}(\xi,g,z)$ is convex. Furthermore, from $\mathcal{A}(\xi,g,z^1) \neq \emptyset$ and $\mathcal{A}(\xi,g,z^2) \neq \emptyset$ follows $\mathcal{A}(\xi^\lambda,g,z^\lambda) \neq \emptyset$, for $z^\lambda := \lambda z^1 + (1-\lambda)z^2$ and $\xi^\lambda := \lambda\xi^1 + (1-\lambda)\xi^2$ where $\lambda \in [0,1]$.*

Proof. The first assertion is a direct implication of (CON). As to the latter, it follows from (CON) that $\lambda(Y^1,Z^1) + (1-\lambda)(Y^2,Z^2) \in \mathcal{A}(\xi^\lambda,g,z^\lambda)$ whenever (Y^1,Z^1) and (Y^2,Z^2) belong to $\mathcal{A}(\xi^1,g,z^1)$ and $\mathcal{A}(\xi^2,g,z^2)$, respectively. □

For the proof of our main existence theorem we will need an auxiliary result concerning the stability of the set $\Theta(z)$ under convergence in $\mathcal{L}^2$, given that the decomposition parts can be uniformly bounded in $\mathcal{L}^2$.

Lemma 3.4. *For any $M > 0$ and $z \in \mathbb{R}^{1\times d}$, the set*

$$\Theta_M(z) = \{Z \in \Theta(z) \,:\, \max\{\|\Delta\|_{\mathcal{L}^2}, \|\Gamma\|_{\mathcal{L}^2}\} \leq M\}$$

is closed under convergence in $\mathcal{L}^2$. If a sequence $(Z^n) \subset \Theta_M(z)$ with $Z^n = z + \int \Delta^n dt + \int \Gamma^n dW$ converges in $\mathcal{L}^2$ to some $Z = z + \int \Delta dt + \int \Gamma dW$, then there is a sequence $((\tilde{\Delta}^n, \tilde{\Gamma}^n))$ in the asymptotic convex hull of $((\Delta^n,\Gamma^n))$ converging in $\mathcal{L}^2 \times \mathcal{L}^2$ to (Δ,Γ).

Proof. First observe that for $Z \in \Theta(z)$ we have

$$|Z_t|^2 \leq 4\left(|z|^2 + \int_0^t |\Delta_s|^2\, ds + \left|\int_0^t \Gamma_s dW_s\right|^2\right).$$

Hence, for $Z \in \Theta_M(z)$, this in turn yields $E[|Z_t|^2] \leq 4(|z|^2 + \|\Delta\|^2_{\mathcal{L}^2} + \|\Gamma\|^2_{\mathcal{L}^2}) \leq 4(|z|^2 + 2M^2) := C < \infty$, and hence $\Theta_M(z)$ is a bounded subset of $\mathcal{L}^2$, since by Fubini's theorem we obtain that $\|Z\|_{\mathcal{L}^2} \leq \sqrt{TC}$. Consider a sequence $Z^n = z + \int \Delta^n du + \int \Gamma^n dW$ in $\Theta_M(z)$ converging in $\mathcal{L}^2$ to some process Z. Since $((\Delta^n, \Gamma^n))$ are bounded in $\mathcal{L}^2 \times \mathcal{L}^2$, we can find a sequence $(\tilde{\Delta}^n, \tilde{\Gamma}^n) \in conv\{(\Delta^n, \Gamma^n), (\Delta^{n+1}, \Gamma^{n+1}), \dots\}$ converging in $\mathcal{L}^2 \times \mathcal{L}^2$ to $(\Delta, \Gamma) \in \mathcal{L}^2 \times \mathcal{L}^2$. Furthermore, it holds that $\|\Delta\|_{\mathcal{L}^2} \vee \|\Gamma\|_{\mathcal{L}^2} \leq M$. Let us denote by $(\tilde{Z}^n)$ the respective sequence in the asymptotic convex hull of (Z^n). From Jensen's inequality we deduce that

$$E\left[\int_0^T \left|\int_0^t (\tilde{\Delta}^n_s - \Delta_s) ds\right|^2 dt\right] \leq TE\left[\int_0^T \left|\tilde{\Delta}^n_s - \Delta_s\right|^2 ds\right] \to 0\,,$$

and thus $(\int \tilde{\Delta}^n ds)$ converges to $\int \Delta ds$ in $\mathcal{L}^2$. Applying Fubini's theorem and using the Itô isometry yield that

$$E\left[\int_0^T \left|\int_0^t \tilde{\Gamma}^n_s dW_s - \int_0^t \Gamma_s dW_s\right|^2 dt\right] \leq TE\left[\int_0^T \left|\tilde{\Gamma}^n_s - \Gamma_s\right|^2 ds\right],$$

where the term on the right-hand side tends to zero by means of the $\mathcal{L}^2$-convergence of $(\tilde{\Gamma}^n)$ to Γ. Hence, $(\int \tilde{\Gamma}^n dW)$ converges to $\int \Gamma dW$ in $\mathcal{L}^2$. $(\tilde{Z}^n)$ inheriting the $\mathcal{L}^2$-convergence to Z from (Z^n) together with the $P \otimes dt$-uniqueness of $\mathcal{L}^2$-limits finally allows us to write the process Z as $Z = z + \int \Delta ds + \int \Gamma dW$, we are done. □

Lemma 3.4 yields the following compactness result.

Lemma 3.5. *Assume that $\mathcal{A}(\xi, g, z)$ is non-empty for some $z \in \mathbb{R}^{1\times d}$. Let ξ^- be in L^1 and g satisfy (POS), (CON) and (DGC). Then, for any sequence $((Y^n, Z^n)) \subset \mathcal{A}(\xi, g, z)$ of supersolutions satisfying $\sup_n Y^n_0 < \infty$, the following holds: There is a sequence $(\tilde{Z}^n)$ in the asymptotic convex hull of (Z^n) that converges in $\mathcal{L}^2$ to some process $\hat{Z} \in \Theta(z)$.*

Proof. Step 1: Existence of $((\tilde{Y}^n, \tilde{Z}^n))$. $\mathcal{L}^2$-convergence of $(\tilde{Z}^n)$ to $\hat{Z}$. First observe that (3.1) and the supermartingale property of all $\int Z^n dW$ imply that

$$E\left[\int_0^T g_t(Y^n_t, Z^n_t, \Delta^n_t, \Gamma^n_t) dt\right] \leq Y^n_0 + E\left[\xi^-\right] \leq C + E\left[\xi^-\right] < \infty\,, \tag{3.3}$$

where we put $C := \sup_n Y_0^n$. Now, using (3.3) together with (DGC) we estimate

$$\|\Delta^n\|_{\mathcal{L}^2}^2 + \|\Gamma^n\|_{\mathcal{L}^2}^2 = E\left[\int_0^T |\Delta_t^n|^2\, dt\right] + E\left[\int_0^T |\Gamma_t^n|^2\, dt\right]$$
$$\leq \frac{1}{c_2} E\left[\int_0^T g_t(Y_t^n, Z_t^n, \Delta_t^n, \Gamma_t^n) dt\right] - \frac{c_1}{c_2} T \leq \frac{1}{c_2}\left(C + E\left[\xi^-\right] - c_1 T\right) < \infty\,.$$

Since the right-hand above is independent of n, we obtain that $(Z^n) \subset \Theta_M(z)$ with $M := [\frac{1}{c_2}(C + E[\xi^-] - c_1 T)]^{\frac{1}{2}}$ and the arguments within the proof of Lemma 3.4 show that the sequence (Z^n) is uniformly bounded in $\mathcal{L}^2$. This in turn guarantees the existence of a sequence $(\tilde{Z}^n)$ in the asymptotic convex hull of (Z^n) that converges to some process $\hat{Z}$ in $\mathcal{L}^2$ and, up to a subsequence, $P \otimes dt$-almost everywhere.

Step 2: The process $\hat{Z}$ belongs to $\Theta(z)$. The sequence $((\tilde{Y}^n, \tilde{Z}^n))$ lies in $\mathcal{A}(\xi, g, z)$, due to (CON). Moreover, the linearity of the integrals within the Itô decompositions of (Z^n) yields that $\tilde{Z}^n = z + \int \tilde{\Delta}^n du + \int \tilde{\Gamma}^n dW$ where $((\tilde{\Delta}^n, \tilde{\Gamma}^n))$ denotes the corresponding convex combinations of the decomposition parts. In addition, $((\tilde{\Delta}^n, \tilde{\Gamma}^n))$ inherits the uniform bound from $((\Delta^n, \Gamma^n))$, that is $\max\{\sup_n \|\tilde{\Delta}^n\|_{\mathcal{L}^2}, \sup_n \|\tilde{\Gamma}^n\|_{\mathcal{L}^2}\} \leq M$. Hence, Lemma 3.4 ensures that $\hat{Z}$ is of the form

$$\hat{Z} = z + \int \hat{\Delta} du + \int \hat{\Gamma} dW\,,$$

with suitable $\mathcal{L}^2$-convergence of the decomposition parts by possibly passing to yet another subsequence in the respective asymptotic convex hull. This finishes the proof. □

3.3 Minimality under constraints

Within the current setup of admissible controls constrained to follow certain dynamics, we are interested in supersolutions $(\hat{Y}, \hat{Z}) \in \mathcal{A}(\xi, g, z)$ minimal at time zero, that is $\hat{Y}_0 \leq Y_0$ for all $(Y, Z) \in \mathcal{A}(\xi, g, z)$. In the remainder of this work, a major role is thus played by the operator

$$\mathcal{E}_0^g(\xi, z) := \inf\{Y_0 : (Y, Z) \in \mathcal{A}(\xi, g, z)\}\,, \tag{3.4}$$

since any $(\hat{Y}, \hat{Z})$ satisfying $\hat{Y}_0 = \mathcal{E}_0^g(\xi, z)$ naturally exhibits the property of being minimal at time zero. Note that the definition of a supersolution directly yields that $\mathcal{A}(\xi^1, g, z) \subseteq \mathcal{A}(\xi^2, g, z)$ whenever $\xi^1 \geq \xi^2$. Hence, we immediately obtain monotonicity of the operator $\mathcal{E}_0(\cdot, z)$, that is $\xi^1 \geq \xi^2$ implies $\mathcal{E}_0^g(\xi^1, z) \geq \mathcal{E}_0^g(\xi^2, z)$. The ensuing Theorem 3.7 provides existence of supersolutions minimal at time zero making use of the fact that the set $\{Y_0 : (Y, Z) \in \mathcal{A}(\xi, g, z)\}$ is directed downwards. Parts of it rely on a version of Helly's theorem which we state here for the sake of completeness. In order to keep this work self-contained, we include the proof given in [12, Lemma 1.25].

Lemma 3.6. *Let (A^n) be a sequence of increasing positive processes such that the sequence (A^n_T) is bounded in L^1. Then, there is a sequence $(\tilde{A}^n)$ in the asymptotic convex hull of (A^n) and an increasing positive integrable process $\tilde{A}$ such that*

$$\lim_{n\to\infty} \tilde{A}^n_t = \tilde{A}_t\,, \quad \textit{for all } t \in [0,T],\ P\textit{-almost surely}\,.$$

Proof. Let (t_j) be a sequence running through $I := ([0,T] \cap \mathbb{Q}) \cup \{T\}$. Since $(A^n_{t_1})$ is an L^1-bounded sequence of positive random variables, due to [6, Lemma A1.1] there exists a sequence $(\tilde{A}^{1,k})$ in the asymptotic convex hull of (A^n) and a random variable $\tilde{A}_{t_1}$ such that $(A^{1,k}_{t_1})$ converges P-almost surely to $\tilde{A}_{t_1}$. Moreover, Fatou's lemma yields $\tilde{A}_{t_1} \in L^1$. Let $(\tilde{A}^{2,k})$ be a sequence in the asymptotic convex hull of $(\tilde{A}^{1,k})$ such that $(\tilde{A}^{2,k}_{t_2})$ converges P-almost surely to $\tilde{A}_{t_2} \in L^1$ and so on. Then, for $s \in I$, it holds $\tilde{A}^{k,k}_s \to \tilde{A}_s$ on a set $\hat{\Omega} \subset \Omega$ satisfying $P(\hat{\Omega}) = 1$. The process $\tilde{A}$ is positive, increasing and integrable on I. Thus we may define

$$\hat{A}_t := \lim_{r\downarrow t, r\in I} \tilde{A}_r\,, \quad t \in [0,T)\,, \quad \hat{A}_T := \tilde{A}_T\,.$$

We now show that $(\tilde{A}^{k,k})$, henceforth named $(\tilde{A}^k)$, converges P-almost surely on the continuity points of $\hat{A}$. To this end, fix $\omega \in \hat{\Omega}$ and a continuity point $t \in [0,T)$ of $\hat{A}(\omega)$. We show that $(\tilde{A}^k_t(\omega))$ is a Cauchy sequence in $\mathbb{R}$. Fix $\varepsilon > 0$ and set $\delta = \frac{\varepsilon}{11}$. Since t is a continuity point of $\hat{A}(\omega)$, we may choose $p_1, p_2 \in I$ such that $p_1 < t < p_2$ and $\hat{A}_{p_1}(\omega) - \hat{A}_{p_2}(\omega) < \delta$. By definition of $\hat{A}$, we may choose $r_1, r_2 \in I$ such that $p_1 < r_1 < t < p_2 < r_2$ and $|\hat{A}_{p_1}(\omega) - \tilde{A}_{r_1}(\omega)| < \delta$ and $|\hat{A}_{p_2}(\omega) - \tilde{A}_{r_2}(\omega)| < \delta$. Now choose $N \in \mathbb{N}$ such that $|\tilde{A}^m_{r_1}(\omega) - \tilde{A}^n_{r_1}(\omega)| < \delta$, for all $m, n \in \mathbb{N}$ with $m, n \geq N$, and $|\tilde{A}^j_{r_2}(\omega) - \tilde{A}_{r_2}(\omega)| < \delta$ and $|\tilde{A}_{r_1}(\omega) - \tilde{A}^j_{r_1}(\omega)| < \delta$ for $j = m, n$. We estimate

$$|\tilde{A}^m_t(\omega) - \tilde{A}^n_t(\omega)| \leq |\tilde{A}^m_t(\omega) - \tilde{A}^m_{r_1}(\omega)| + |\tilde{A}^m_{r_1}(\omega) - \tilde{A}^n_{r_1}(\omega)| + |\tilde{A}^n_{r_1}(\omega) - \tilde{A}^n_t(\omega)|\,.$$

For the first and the third term on the right hand side, since A and A are increasing, we deduce that $|\tilde{A}^m_t(\omega) - \tilde{A}^m_{r_1}(\omega)| \leq |\tilde{A}^m_{r_2}(\omega) - \tilde{A}^m_{r_1}(\omega)|$ and $|\tilde{A}^n_t(\omega) - \tilde{A}^n_{r_1}(\omega)| \leq |\tilde{A}^n_{r_2}(\omega) - \tilde{A}^n_{r_1}(\omega)|$. Furthermore,

$$\begin{aligned}|\tilde{A}^j_{r_2}(\omega) - \tilde{A}^j_{r_1}(\omega)| \leq\ & |\tilde{A}^j_{r_2}(\omega) - \tilde{A}_{r_2}(\omega)| + |\tilde{A}_{r_2}(\omega) - \hat{A}_{p_2}(\omega)| \\ & + |\hat{A}_{p_2}(\omega) - \hat{A}_{p_1}(\omega)| + |\hat{A}_{p_1}(\omega) - \tilde{A}_{r_1}(\omega)| + |\tilde{A}_{r_1}(\omega) - \tilde{A}^j_{r_1}(\omega)|\,,\end{aligned}$$

for $j = m, n$. Combining the previous inequalities yields $|\tilde{A}^m_t(\omega) - \tilde{A}^n_t(\omega)| \leq \varepsilon$, for all $m, n \geq N$. Hence, $(\tilde{A}^k(\omega))$ converges for all continuity points $t \in [0,T)$ of $\hat{A}(\omega)$, for all $\omega \in \hat{\Omega}$. We denote the limit by $\tilde{A}$.

It remains to be shown that $(\tilde{A}^k)$ also converges for the discontinuity points of $\hat{A}$. To this end, note that $\hat{A}$ is càdlàg and adapted to our filtration which fulfills the usual conditions. By a well-known result, see for example [15, Proposition 1.2.26], this implies that the jumps of $\hat{A}$ may be exhausted by a sequence of stopping times (ρ^j). Applying once more [6, Lemma A1.1] iteratively on the sequences $(\tilde{A}^k_{\rho^j})_{k\in\mathbb{N}}$, $j = 1, 2, 3 \ldots$, and diagonalizing yields the result. □

Theorem 3.7. *Assume that $\mathcal{A}(\xi, g, z) \neq \emptyset$ for some $\xi^- \in L^1$ and $z \in \mathbb{R}^{1\times d}$ and let g satisfy* (LSC), (POS), (CON) *and* (DGC). *Then, the set $\{(\hat{Y}, \hat{Z}) \in \mathcal{A}(\xi, g, z) : \hat{Y}_0 = \mathcal{E}_0^g(\xi, z)\}$ is non-empty.*

Proof. Step 1: The candidate control $\hat{Z}$. We extract a sequence $((Y^n, Z^n)) \subset \mathcal{A}(\xi, g, z)$ such that

$$\lim_{n\to\infty} Y_0^n = \mathcal{E}_0^g(\xi, z)\,.$$

Because $\sup_n Y_0^n \leq Y_0^1 < \infty$, Lemma 3.5 assures the existence of a sequence $(\tilde{Z}^n)$ in the asymptotic convex hull of (Z^n) that converges in $\mathcal{L}^2$ to some admissible process $\hat{Z} \in \Theta(z)$, including $\mathcal{L}^2$-convergence of the corresponding decomposition parts. In particular, we obtain that

$$\int_0^t \tilde{Z}_u^n dW_u \underset{n\to\infty}{\longrightarrow} \int_0^t \hat{Z}_u dW_u\,, \quad \text{for all } t \in [0,T]\,,\ P\text{-almost surely}\,. \tag{3.5}$$

Moreover, up to a subsequence, $((\tilde{Z}^n, \tilde{\Delta}^n, \tilde{\Gamma}^n))$ converges $P \otimes dt$-almost everywhere towards $(\hat{Z}, \hat{\Delta}, \hat{\Gamma})$.

Step 2: The candidate value process $\hat{Y}$. If we denote by $(\tilde{Y}^n)$ the sequence in the asymptotic convex hull of (Y^n) corresponding to $(\tilde{Z}^n)$, then all $(\tilde{Y}^n, \tilde{Z}^n)$ satisfy (3.1) due to (CON). Let $\tilde{A}^n$ denote the increasing, predicable process of finite variation stemming from the decomposition of $\tilde{Y}^n = \tilde{Y}_0^n + \tilde{M}^n - \tilde{A}^n$ given in Lemma 3.2. Since $(\tilde{Z}^n)$ is uniformly bounded in $\mathcal{L}^2$ and thus all $\int \tilde{Z}^n dW$ are true martingales, and g satisfies (POS), the decomposition (3.2) yields

$$E\left[\tilde{A}_T^n\right] \leq Y_0^1 + E\left[\xi^-\right] < \infty\,,$$

as we assumed ξ^- to be an element of L^1. Now a version of Helly's theorem, see Lemma 3.6, yields the existence of a sequence in the asymptotic convex hull of $(\tilde{A}^n)$, again denoted by the previous expression, and of an increasing positive integrable process $\tilde{A}$ such that $\lim_{n\to\infty} \tilde{A}_t^n = \tilde{A}_t$, for all $t \in [0,T]$, P-almost surely. We pass to the corresponding sequence on the side of $(\tilde{Y}^n)$ and $(\tilde{Z}^n)$, define the process $\tilde{Y}$ pointwise for all $t \in [0,T]$ by $\tilde{Y}_t := \lim_{n\to\infty} \tilde{Y}_t^n = \mathcal{E}_0^g(\xi, z) + \int_0^t \hat{Z}_u dW_u - \tilde{A}_t$, and observe that it fulfills $\tilde{Y}_0 = \mathcal{E}_0^g(\xi, z)$ by construction. However, since $\tilde{Y}$ is not necessarily càdlàg, we define our candidate value process $\hat{Y}$ by $\hat{Y}_t := \lim_{s\downarrow t, s\in\mathbb{Q}} \tilde{Y}_s$, for all $t \in [0,T)$ and $\hat{Y}_T := \xi$. The continuity of $\int \hat{Z} dW$ yields that

$$\hat{Y}_t = \mathcal{E}_0^g(\xi, z) + \int_0^t \hat{Z}_u dW_u - \lim_{s\downarrow t, s\in\mathbb{Q}} \tilde{A}_s\,. \tag{3.6}$$

Since jump times of càdlàg processes[3] can be exhausted by a sequence of stopping times $(\sigma_j) \subset \mathcal{T}$, compare [15, Proposition 1.2.26], which coincide with the jump times of $\tilde{A}$, we conclude that

$$\hat{Y} = \tilde{Y}\,, \quad P \otimes dt\text{-almost everywhere}\,. \tag{3.7}$$

[3]Note that as an increasing process, $\tilde{A}$ is in particular a submartingale and thus its right- and left-hand limits exist, compare [15, Proposition 1.3.14]. Consequently, the process $\lim_{s\downarrow\cdot, s\in\mathbb{Q}} \tilde{A}_s$ is càdlàg.

Furthermore, $\tilde{A}$ increasing implies that $\hat{A}_t := \lim_{s \downarrow t, s \in \mathbb{Q}} \tilde{A}_s \geq \tilde{A}_t$, for all $t \in [0,T]$ which, together with (3.6), in turn yields that

$$\hat{Y}_t \leq \tilde{Y}_t\,, \quad \text{for all } t \in [0,T]\,. \tag{3.8}$$

Given that $(\hat{Y}, \hat{Z})$ satisfies (3.1), we could conclude that $(\hat{Y}, \hat{Z}) \in \mathcal{A}(\xi, g, z)$ and thus $\hat{Y}_0 \geq \mathcal{E}_0^g(\xi, z) = \tilde{Y}_0$ which, combined with (3.8), would imply $\hat{Y}_0 = \mathcal{E}_0^g(\xi, z)$ and thereby finish the proof.

Step 3: Verification. As to the remaining verification, we deduce from (3.7) the existence of a set $A \in \mathcal{F}_T$, $P(A) = 1$ with the following property. For all $\omega \in A$, there exists a Lebesgue measurable set $\mathcal{I}(\omega) \subset [0,T]$ of measure T such that $\tilde{Y}_t^n(\omega) \longrightarrow \hat{Y}_t(\omega)$, for all $t \in \mathcal{I}(\omega)$. We suppress the dependence of $\mathcal{I}$ on ω and recall however that in the following s and t may depend on ω. For $s, t \in \mathcal{I}$ with $s \leq t$ holds

$$\begin{aligned} \hat{Y}_s - \int_s^t g_u(\hat{Y}_u, \hat{Z}_u, \hat{\Delta}_u, \hat{\Gamma}_u)du + \int_s^t \hat{Z}_u dW_u \\ \geq \limsup_n \left(\tilde{Y}_s^n - \int_s^t g_u(\tilde{Y}_u^n, \tilde{Z}_u^n, \tilde{\Delta}_u^n, \tilde{\Gamma}_u^n)du + \int_s^t \tilde{Z}_u^n dW_u \right) \end{aligned} \tag{3.9}$$

by means of (3.5), the $P \otimes dt$-almost-everywhere convergence of $((\tilde{Y}^n, \tilde{Z}^n, \tilde{\Delta}^n, \tilde{\Gamma}^n))$ towards $(\hat{Y}, \hat{Z}, \hat{\Delta}, \hat{\Gamma})$, the property (LSC) and Fatou's lemma. Using $((\tilde{Y}^n, \tilde{Z}^n)) \subset \mathcal{A}(\xi, g, z)$, for all $n \in \mathbb{N}$, (3.9) can be further estimated by

$$\hat{Y}_s - \int_s^t g_u(\hat{Y}_u, \hat{Z}_u, \hat{\Delta}_u, \hat{\Gamma}_u)du + \int_s^t \hat{Z}_u dW_u \geq \limsup_n \tilde{Y}_t^n = \hat{Y}_t\,. \tag{3.10}$$

Whenever $s, t \in \mathcal{I}^c$ with $s \leq t$, we approximate both times from the right by sequences $(s^n) \subset \mathcal{I}$ and $(t^n) \subset \mathcal{I}$, respectively, such that $s^n \leq t^n$. Since (3.10) holds for all s^n and t^n, the claim follows from the right-continuity of $\hat{Y}$ and the continuity of all appearing integrals, which finally concludes the proof. □

Convexity of the mapping $(\xi, z) \mapsto \mathcal{E}_0^g(\xi, z)$ is provided by the following lemma.

Lemma 3.8. *Under the assumptions of Theorem 3.7, the operator $\mathcal{E}_0^g(\cdot, \cdot)$ is jointly convex.*

Proof. For $z^1, z^2 \in \mathbb{R}^{1\times d}$ and $\xi^1, \xi^2 \in L^0$, the negative parts of which are integrable, assume that $\mathcal{A}(\xi^1, g, z^1) \neq \emptyset$ and $\mathcal{A}(\xi^2, g, z^2) \neq \emptyset$, as otherwise convexity trivially holds. For $\lambda \in [0,1]$ we set $z^\lambda := \lambda z^1 + (1-\lambda) z^2$ and $\xi^\lambda := \lambda \xi^1 + (1-\lambda)\xi^2$ so that Lemma 3.3 implies $\mathcal{A}(\xi^\lambda, g, z^\lambda) \neq \emptyset$. By Theorem 3.7, there exist (Y^1, Z^1) and (Y^2, Z^2) in $\mathcal{A}(\xi^1, g, z^1)$ and $\mathcal{A}(\xi^2, g, z^2)$, respectively, such that $Y_0^1 = \mathcal{E}_0^g(\xi^1, z^1)$ and $Y_0^2 = \mathcal{E}_0^g(\xi^2, z^2)$. Since $(\bar{Y}, \bar{Z}) := \lambda(Y^1, Z^1) + (1-\lambda)(Y^2, Z^2)$ is an element of $\mathcal{A}(\xi^\lambda, g, z^\lambda)$ due to (CON), it holds $\mathcal{E}_0^g(\xi^\lambda, z^\lambda) \leq \bar{Y}_0$ by definition of the operator $\mathcal{E}_0^g$. □

3.4 Stability results

Next, we show that the non-linear operator $\xi \mapsto \mathcal{E}_0^g(\xi, z)$ exhibits stability properties such as monotone convergence, the Fatou property or L^1-lower semicontinuity. The following theorem establishes monotone convergence and the Fatou property of $\mathcal{E}_0^g(\cdot, z)$. Similar results in the unconstrained case have been obtained in [9, Theorem 4.7].

Theorem 3.9. *For $z \in \mathbb{R}^{1\times d}$ and g a generator fulfilling (LSC), (POS), (CON) and (DGC), and (ξ_n) a sequence in L^0 such that $(\xi_n^-) \subset L^1$, the following holds.*

- *Monotone convergence: If (ξ_n) is increasing P-almost surely to $\xi \in L^0$, then it holds $\lim_{n\to\infty} \mathcal{E}_0^g(\xi_n, z) = \mathcal{E}_0^g(\xi, z)$.*
- *Fatou's lemma: If $\xi_n \geq \eta$, for all $n \in \mathbb{N}$, where $\eta \in L^1$, then it holds $\mathcal{E}_0^g(\liminf_n \xi_n, z) \leq \liminf_n \mathcal{E}_0^g(\xi_n, z)$.*

Proof. Monotone convergence: First, note that by monotonicity of the operator under consideration the limit $\bar{Y}_0 := \lim_n \mathcal{E}_0^g(\xi_n, z)$ exists and satisfies $\bar{Y}_0 \leq \mathcal{E}_0^g(\xi, z)$. Other than in the trivial case of $+\infty = \bar{Y}_0 \leq \mathcal{E}_0^g(\xi, z)$ we have $\mathcal{A}(\xi_n, g, z) \neq \emptyset$, for all $n \in \mathbb{N}$, which, together with $(\xi_n^-) \subset L^1$ implies $(\xi_n) \subset L^1$. Furthermore, Theorem 3.7 yields the existence of supersolutions $(Y^n, Z^n) \in \mathcal{A}(\xi_n, g, z)$ fulfilling $Y_0^n = \mathcal{E}_0^g(\xi_n, z)$, for all $n \in \mathbb{N}$. In particular, we have that $Y_0^n \leq \bar{Y}_0$ and $\xi_n^- \leq \xi_1^-$, for all $n \in \mathbb{N}$. Arguments analogous to the ones used in Lemma 3.5 and the proof of Theorem 3.7 directly translate to the present setting and provide both a candidate control $\hat{Z} \in \Theta(z)$ to which $(\tilde{Z}^n)$ converges and a corresponding $\tilde{Y}_t := \lim_n \tilde{Y}_t^n$, and ensure that $(\hat{Y}, \hat{Z})$ belongs to $\mathcal{A}(\xi, g, z)$, where $\hat{Y} := \lim_{s\in\mathbb{Q}, s\downarrow \cdot} \tilde{Y}_s$ on $[0, T)$ and $\hat{Y}_T := \xi$. In particular, we obtain $\hat{Y}_0 \leq \tilde{Y}_0 = \bar{Y}_0$. Hence, as $\mathcal{A}(\xi, g, z) \neq \emptyset$ and $\xi^- \in L^1$, there exists $(Y, Z) \in \mathcal{A}(\xi, g, z)$ such that $Y_0 = \mathcal{E}_0^g(\xi, z)$. By minimality of (Y, Z) at time zero, however, this entails $Y_0 \leq \hat{Y}_0 \leq \bar{Y}_0$ and we conclude that $\lim_{n\to\infty} \mathcal{E}_0^g(\xi_n, z) = \mathcal{E}_0^g(\xi, z)$.

Fatou's lemma: If we define $\zeta_n := \inf_{k\geq n} \xi_k$, then $\xi_k \geq \eta$ for all $k \in \mathbb{N}$ implies $\zeta_n \geq \eta$ for all $n \in \mathbb{N}$ which in turn gives $(\zeta_n^-) \subset L^1$, and thus the monotone convergence established above can be used exactly as in [9, Theorem 4.7] to obtain the assertion. □

As a consequence of the monotone convergence property we obtain the ensuing theorem providing L^1-lower semicontinuity of the operator $\mathcal{E}_0^g(\cdot, z)$. The proof goes along the lines of [9, Theorem 4.9] and is thus omitted here.

Theorem 3.10. *Let $z \in \mathbb{R}^{1\times d}$ and g be a generator fulfilling (LSC), (POS), (CON) and (DGC). Then $\mathcal{E}_0^g(\cdot, z)$ is L^1-lower semicontinuous.*

4 Duality under constraints

The objective of this section is to construct a solution of constrained BSDEs via duality and, for the case of a quadratic generator, to obtain an explicit form for $\mathcal{E}_0^*$, the Fenchel-Legendre transform of $\mathcal{E}_0^g$. Let us assume for the rest of this section that our generator g is independent of y, that is $g_u(y, z, \delta, \gamma) = g_u(z, \delta, \gamma)$, and that it satisfies (LSC), (POS), (CON) and (DGC). Let us further fix some $z \in \mathbb{R}^{1\times d}$ as initial value of the controls and set $\mathcal{E}_0^g(\cdot) := \mathcal{E}_0^g(\cdot, z)$ for the remainder

of this section. Whenever we say that the BSDE(ξ, g) has a solution (Y, Z), we mean that there exists $(Y, Z) \in \mathcal{A}(\xi, g, z)$ such that (3.1) is satisfied with equalities instead of inequalities. Observe that $\mathcal{E}_0^g(\cdot)$, being convex and L^1-lower semicontinuous, is in particular $\sigma(L^1, L^\infty)$-lower semicontinuous, and thus, by classical duality results admits the Fenchel-Moreau representation

$$\mathcal{E}_0^g(\xi) = \sup_{v \in L^\infty} \{E[v\xi] - \mathcal{E}_0^*(v)\}\,, \quad \xi \in L^1\,, \tag{4.1}$$

where for $v \in L^\infty$ the convex conjugate is given by

$$\mathcal{E}_0^*(v) := \sup_{\xi \in L^1} \{E[v\xi] - \mathcal{E}_0^g(\xi)\}\,.$$

It is proved in the next lemma that the domain of $\mathcal{E}_0^*$ is concentrated on non-negative $v \in L^\infty_+$ satisfying $E[v] = 1$.

Lemma 4.1. *Within the representation (4.1), that is $\mathcal{E}_0^g(\xi) = \sup_{v \in L^\infty}\{E[v\xi] - \mathcal{E}_0^*(v)\}$, the supremum might be restricted to those $v \in L^\infty_+$ satisfying $E[v] = 1$.*

Proof. First, we assume without loss of generality that $\mathcal{E}_0^g(0) < +\infty$. Indeed, a slight modification of the argumentation below remains valid using any $\xi \in L^1$ such that $\mathcal{E}_0^g(\xi) < +\infty$.[4] We show that $\mathcal{E}_0^*(v) = +\infty$ as soon as $v \in L^\infty \backslash L^\infty_+$ or $E[v] \neq 1$. For $v \in L^\infty \backslash L^\infty_+$, L^1_+ being the polar of L^∞_+ yields the existence of $\bar{\xi} \in L^1_+$ such that $E[v\bar{\xi}] < 0$. Monotonicity of $\mathcal{E}_0^g$ then gives $\mathcal{E}_0^g(-n\bar{\xi}) \leq \mathcal{E}_0^g(0)$ for all $n \in \mathbb{N}$. Hence,

$$\mathcal{E}_0^*(v) \geq \sup_n \left\{nE[-v\bar{\xi}] - \mathcal{E}_0^g(-n\bar{\xi})\right\} \geq \sup_n \left\{nE[-v\bar{\xi}]\right\} - \mathcal{E}_0^g(0) = +\infty\,.$$

Furthermore, since the generator does not depend on y, the function $\mathcal{E}_0^g$ is cash additive, compare [9, Proposition 3.3.5], and we deduce that, for all $n \in \mathbb{N}$ it holds

$$\mathcal{E}_0^*(v) \geq E[vn] - \mathcal{E}_0^g(0) - n = n(E[v] - 1) - \mathcal{E}_0^g(0)\,.$$

Thus, if $E[v] > 1$, then $\mathcal{E}_0^*(v) = +\infty$. A reciprocal argument with $\xi = -n$ finally gives $\mathcal{E}_0^*(v) = +\infty$ whenever $E[v] < 1$. □

By the previous result, we may use $v \in L^\infty_+$, $E[v] = 1$, in order to define a measure Q that is absolutely continuous with respect to P by setting $\frac{dQ}{dP} := v$. Thereby (4.1) may be reformulated as

$$\mathcal{E}_0^g(\xi) = \sup_{Q \ll P} \{E_Q[\xi] - \mathcal{E}_0^*(Q)\}\,, \quad \xi \in L^1\,, \tag{4.2}$$

where

$$\mathcal{E}_0^*(Q) := \sup_{\xi \in L^1} \{E_Q[\xi] - \mathcal{E}_0^g(\xi)\}\,. \tag{4.3}$$

Note that $\mathcal{A}(\xi, g, z) = \emptyset$ implies $\mathcal{E}_0^g(\xi) = +\infty$ and hence such terminal conditions are irrelevant for the supremum in (4.3). Let us denote by $\mathcal{Q}$ the set of all probability measures equivalent

[4] Note that the case $\mathcal{E}_0^g \equiv +\infty$ on L^1 immediately yields $\mathcal{E}_0^* \equiv -\infty$ on L^∞ and is thus neglected.

to P with bounded Radon-Nikodym derivative. For each $Q \in \mathcal{Q}$, there exists a progressively measurable process q taking values in $\mathbb{R}^{1\times d}$ such that for all $t \in [0,T]$

$$\frac{dQ}{dP}|_{\mathcal{F}_t} = \exp\left(\int_0^t q_u dW_u - \frac{1}{2}\int_0^t |q_u|^2 du\right).$$

By Girsanov's theorem, the process $W_t^Q := W_t - \int_0^t q_u du$ is a Q-Brownian motion. The following lemma is a valuable tool regarding the characterization of $\mathcal{E}_0^*$.

Lemma 4.2. *The supremum in (4.3) can be restricted to random variables $\xi \in L^1$ for which the BSDE with parameters (ξ, g) has a solution with value process starting in $\mathcal{E}_0^g(\xi)$. More precisely, for any $Q \in \mathcal{Q}$ holds*

$$\mathcal{E}_0^*(Q) = \sup_{\xi\in L^1}\{E_Q[\xi] - \mathcal{E}_0^g(\xi) \,:\, \textit{BSDE}(\xi,g) \textit{ has a solution } (Y,Z) \textit{ with } Y_0 = \mathcal{E}_0^g(\xi)\}.$$

Proof. It suffices to show that

$$\begin{aligned}&\mathcal{E}_0^*(Q)\\ &\quad\le \sup_{\xi\in L^1}\{E_Q[\xi] - \mathcal{E}_0^g(\xi) \,:\, \text{BSDE}(\xi,g) \text{ has a solution } (Y,Z) \text{ with } Y_0 = \mathcal{E}_0^g(\xi)\},\end{aligned} \tag{4.4}$$

since the reverse inequality is satisfied by definition of $\mathcal{E}_0^*(\cdot)$. Consider to this end a terminal condition $\xi \in L^1$ with associated minimal supersolution $(Y,Z) \in \mathcal{A}(\xi,g,z)$, that is $Y_0 = \mathcal{E}_0^g(\xi)$. Put, for all $t \in [0,T]$,

$$Y_t^1 = \mathcal{E}_0^g(\xi) - \int_0^t g_u\left(Z_u, \Delta_u, \Gamma_u\right) du + \int_0^t Z_u dW_u.$$

Relation (3.1) implies $Y_T^1 \ge Y_T \ge \xi$ and thus $\mathcal{E}_0^g(Y_T^1) \ge \mathcal{E}_0^g(\xi)$ and $(Y_T^1)^- \in L^1$. Furthermore, observe that

$$(Y_T^1)^+ = \left(\mathcal{E}_0^g(\xi) - \int_0^T g_u\left(Z_u, \Delta_u, \Gamma_u\right) du + \int_0^T Z_u dW_u\right)^+ \le \left(\mathcal{E}_0^g(\xi) + \int_0^T Z_u dW_u\right)^+$$

due to the positivity of the generator. But since the right-hand side is in L^1 by means of the martingale property of $\int Z dW$, we deduce that $(Y_T^1)^+ \in L^1$, allowing us to conclude that $Y_T^1 \in L^1$. On the other hand, $(Y^1, Z) \in \mathcal{A}(Y_T^1, g, z)$ holds by definition of Y^1. Hence, we conclude $\mathcal{E}_0^g(Y_T^1) \le Y_0^1 = \mathcal{E}_0^g(\xi)$. Thus, $\mathcal{E}_0^g(Y_T^1) = \mathcal{E}_0^g(\xi)$, and (Y^1, Z) is a solution of the BSDE with parameters (Y_T^1, g). Observe further that $\mathcal{E}_0^g(Y_T^1) - \mathcal{E}_0^g(\xi) = 0 \le Y_T^1 - \xi$ which, by taking expectation under Q, implies

$$E_Q[\xi] - \mathcal{E}_0^g(\xi) \le E_Q[Y_T^1] - \mathcal{E}_0^g(Y_T^1).$$

Taking the supremum yields (4.4), the proof is done. □

By means of the preceding lemma it holds

$$\begin{aligned}\mathcal{E}_0^*(Q) &= \sup_{\xi\in L^1}\{E_Q[\xi]-\mathcal{E}_0^g(\xi)\}\\ &= \sup_{\xi\in L^1}\left\{E_Q\left[\mathcal{E}_0^g(\xi)-\int_0^T g_u(Z_u,\Delta_u,\Gamma_u)du+\int_0^T Z_u dW_u\right]-\mathcal{E}_0^g(\xi)\right\}\\ &= \sup_{(\Delta,\Gamma)\in\Pi}\left\{E_Q\left[-\int_0^T g_u(Z_u,\Delta_u,\Gamma_u)du+\int_0^T Z_u dW_u\right]\right\} \qquad (4.5)\end{aligned}$$

where

$$\Pi := \left\{(\Delta,\Gamma)\in\mathcal{L}^2\times\mathcal{L}^2 : \begin{array}{l}\exists\xi\in L^1: \text{ BSDE}(\xi,g) \text{ has a solution } (Y,Z)\\ \text{with } Y_0=\mathcal{E}_0^g(\xi) \text{ and } Z=z+\int\Delta du+\int\Gamma dW\end{array}\right\}. \qquad (4.6)$$

Whenever $Q\in\mathcal{Q}$, Girsanov's theorem applies and we may exploit the decomposition of Z and use that $\int ZdW^Q$ and $\int\Gamma dW^Q$are Q-martingales in order to express the right-hand side of (4.5) without Brownian integrals. More precisely,

$$\begin{aligned}\mathcal{E}_0^*(Q) = \sup_{(\Delta,\Gamma)\in\Pi}\left\{E_Q\left[\int_0^T\left(-g_u(Z_u,\Delta_u,\Gamma_u)+q_u\int_0^u(\Delta_s+q_s\Gamma_s)ds\right)du\right]\right\}\\ + zE_Q\left[\int_0^T q_u du\right]. \qquad (4.7)\end{aligned}$$

We continue with two lemmata that allow us to restrict the set of measures in the representation (4.2) to a sufficiently nice subset of $\mathcal{Q}$ on the one hand, and to change the set Π appearing in (4.5) to the whole space $\mathcal{L}^2\times\mathcal{L}^2$ on the other hand.

Lemma 4.3. *Assume there exists some $\xi\in L^1$ such that $\mathcal{A}(\xi,g,z)\neq\emptyset$. Then it is sufficient to consider measures with densities that are bounded away from zero, that is*

$$\mathcal{E}_0^g(\xi) = \sup_{v\in L_b^\infty}\{E[v\xi]-\mathcal{E}_0^*(v)\} \qquad (4.8)$$

where $L_b^\infty := \{v\in L^\infty : v>0 \quad \text{and} \quad \|\frac{1}{v}\|_\infty<\infty\}$.

Proof. The assumption of $\mathcal{A}(\xi,g,z)$ being non-empty for some $\xi\in L^1$ implies the existence of $(\Delta,\Gamma)\in\Pi$ and corresponding Z such that $E_P[\int_0^T g_u(Z_u,\Delta_u,\Gamma_u)du]<\infty$ which together with (POS), (4.5) and the martingale property of all occurring $\int ZdW$ under P immediately yields that $\mathcal{E}_0^*(P)<\infty$. For any $Q\ll P$ with $\frac{dQ}{dP}=v\in L_+^\infty$ and $\lambda\in(0,1)$ we define a measure Q^λ by its Radon-Nikodym derivative $v_\lambda := (1-\lambda)v+\lambda$ where naturally $\frac{dP}{dP}=1$. Observe that $\lambda>0$ implies $v_\lambda\in L_b^\infty$. Next, we show that $\lim_{\lambda\downarrow 0}\mathcal{E}_0^*(v_\lambda)=\mathcal{E}_0^*(v)$. Indeed, convexity of $\mathcal{E}_0^*(\cdot)$ together with $\mathcal{E}_0^*(P)=\mathcal{E}_0^*(1)<\infty$ yields $\liminf_{\lambda\downarrow 0}\mathcal{E}_0^*(v_\lambda)\leq\mathcal{E}_0^*(v)$,

whereas the reverse inequality is satisfied by means of the lower semicontinuity. On the other hand, dominated convergence gives $\lim_{\lambda \downarrow 0} E[v_\lambda \xi] = E[v\xi]$, since $|v_\lambda \xi| \leq |v\xi| + |\xi|$ which is integrable. Consequently, the expression $\{E[v\xi] - \mathcal{E}_0^*(v)\}$ is the limit of a sequence $(E[v_{\lambda_n}\xi] - \mathcal{E}_0^*(v_{\lambda_n}))_n$ where $(v_{\lambda_n}) \subset L_b^\infty$ and $\lambda_n \downarrow 0$. Since $\mathcal{E}_0^g(\xi)$ can be expressed as the supremum of $\{E[v\xi] - \mathcal{E}_0^*(v)\}$ over all v, it suffices to consider the supremum over $v \in L_b^\infty$, the proof is done. □

Lemma 4.4. *For each $Q \in \mathcal{Q}$ such that $\frac{dQ}{dP} \in L_b^\infty$ it holds*

$$\mathcal{E}_0^*(Q) = \sup_{(\Delta,\Gamma)\in\mathcal{L}^2\times\mathcal{L}^2} \left\{ E_Q\left[-\int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du + \int_0^T Z_u dW_u \right] \right\}. \tag{4.9}$$

Proof. Since Π defined in (4.6) is a subset of $\mathcal{L}^2 \times \mathcal{L}^2$, "$\leq$" certainly holds in (4.9). As to the reverse inequality, observe first that, since we consider a supremum in (4.9) and $Z \in \mathcal{L}^2$ whenever $(\Delta, \Gamma) \in \mathcal{L}^2 \times \mathcal{L}^2$, those (Δ, Γ) such that $E_Q[\int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du] = +\infty$ can be neglected in the following. In particular, since $v = \frac{dQ}{dP} \in L_b^\infty$, we can restrict our focus to those elements satisfying $E[\int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du] \leq \|\frac{1}{v}\|_{L^\infty} E_Q[\int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du] < +\infty$. Thus, given such a pair (Δ, Γ), the terminal condition $\xi := -\int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du + \int_0^T Z_u dW_u$ fulfills $\xi^- \in L^1$ due to the martingale property of $\int Z dW$. Furthermore, the pair $(-\int_0^\cdot g_u(Z_u, \Delta_u, \Gamma_u)du + \int_0^\cdot Z_u dW_u, Z)$ is an element of $\mathcal{A}(\xi, g, z)$ by construction and hence Theorem 3.7 yields the existence of $(\bar{Y}, \bar{Z}) \in \mathcal{A}(\xi, g, z)$ satisfying $\bar{Y}_0 = \mathcal{E}_0^g(\xi) \leq 0$. Now, using the same techniques as in the proof of Lemma 4.2, we define Y^1 by $Y_t^1 := \mathcal{E}_0^g(\xi) - \int_0^t g_u(\bar{Z}_u, \bar{\Delta}_u, \bar{\Gamma}_u)du + \int_0^t \bar{Z}_u dW_u$, for all $t \in [0, T]$, where $(\bar{\Delta}, \bar{\Gamma})$ is the decomposition of $\bar{Z}$, and obtain that $Y_T^1 \geq \xi$ as well as $\mathcal{E}_0^g(Y_T^1) = \mathcal{E}_0^g(\xi)$. Consequently,

$$\begin{aligned} -\int_0^T g_u(\bar{Z}_u, \bar{\Delta}_u, \bar{\Gamma}_u)du + \int_0^T \bar{Z}_u dW_u \\ = Y_T^1 - \mathcal{E}_0^g(\xi) \geq Y_T^1 \geq \xi = -\int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du + \int_0^T Z_u dW_u\,, \end{aligned} \tag{4.10}$$

which, by taking expectation under Q in (4.10) and using $(\bar{\Delta}, \bar{\Gamma}) \in \Pi$, implies

$$\mathcal{E}_0^*(Q) \geq E_Q\left[-\int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du + \int_0^T Z_u dW_u \right]. \tag{4.11}$$

Since (Δ, Γ) was arbitrary, we have finally shown that $\mathcal{E}_0^*(Q)$ is greater or equal to the supremum over $(\Delta, \Gamma) \in \mathcal{L}^2 \times \mathcal{L}^2$ of the right-hand side of (4.11), which finishes the proof. □

The ensuing proposition provides, for a given measure $Q \in \mathcal{Q}$ with $\frac{dQ}{dP} \in L_b^\infty$, the existence of a pair of processes attaining the supremum in (4.5).

Proposition 4.5. *For each $Q \in \mathcal{Q}$ with $\frac{dQ}{dP} \in L_b^\infty$ there exist $(\Delta^Q, \Gamma^Q) \in \Pi$ and a corresponding control Z^Q of the form $Z^Q = z + \int \Delta^Q du + \int \Gamma^Q dW$ such that*

$$\mathcal{E}_0^*(Q) = E_Q\left[-\int_0^T g_u(Z_u^Q, \Delta_u^Q, \Gamma_u^Q)du + \int_0^T Z_u^Q dW_u\right]. \tag{4.12}$$

Furthermore, if the convexity of g is strict, then the triple $(Z^Q, \Delta^Q, \Gamma^Q)$ is unique.

Proof. Step 1: The integral $\int qdW$ is an element of BMO. We begin by proving that, for $Q \in \mathcal{Q}$ the density $\frac{dQ}{dP} = \exp(\int_0^T q_u dW_u - \frac{1}{2}\int_0^T |q_u|^2 du)$ of which belongs to L_b^∞, the process $(\int_0^t q_u dW_u)_{t\in[0,T]}$ is an element of BMO. Indeed, since the process $v_t := E[\frac{dQ}{dP} \,|\, \mathcal{F}_t]$ is uniformly bounded away from zero, it satisfies the Muckenhaupt (A_1) condition, see [16, Definition 2.2], and therefore $\int qdW \in BMO$ by means of [16, Theorem 2.4].

Step 2: $\mathcal{L}^2$-boundedness of a minimizing sequence and the candidate (Δ^Q, Γ^Q). Since the generator g satisfies (DGC), it holds for all (Δ, Γ, Z) that

$$\|\Delta\|^2_{\mathcal{L}^2(Q)} + \|\Gamma\|^2_{\mathcal{L}^2(Q)} \leq \frac{1}{c_2}\left(E_Q\left[\int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du\right] - c_1 T\right). \tag{4.13}$$

If we put $F(Z, \Delta, \Gamma) := E_Q[\int_0^T (g_u(Z_u, \Delta_u, \Gamma_u)du - q_u \int_0^u (\Delta_s + q_s\Gamma_s)ds)du]$, then (4.7) in combination with Lemma 4.4 implies that the conjugate can be expressed by $\mathcal{E}_0^*(Q) = -\inf_{(\Delta,\Gamma)\in\mathcal{L}^2\times\mathcal{L}^2} F(Z, \Delta, \Gamma) + zE_Q[\int_0^T q_u du]$. We claim that, for $(Z^n, \Delta^n, \Gamma^n)$ a minimizing sequence of F, both (Δ^n) and (Γ^n) are bounded in $\mathcal{L}^2(Q)$. Since in our case the $\mathcal{L}^2$-norms with respect to P and Q are equivalent, we suppress the dependence on the measure in the notation to follow. Assume now contrary to our assertion that $\|\Delta^n\|^2_{\mathcal{L}^2} \to \infty$ and $\|\Gamma^n\|^2_{\mathcal{L}^2} \to \infty$ as n tends to infinity. This in turn would imply either

$$E_Q\left[\int_0^T q_u \int_0^u \Delta_s^n ds\, du\right] \to \infty \quad \text{and} \quad \limsup_n \frac{\|\Delta^n\|^2_{\mathcal{L}^2}}{E_Q\left[\int_0^T q_u \int_0^u \Delta_s^n ds\, du\right]} = K \tag{4.14}$$

or

$$E_Q\left[\int_0^T q_u \int_0^u q_s\Gamma_s^n ds\, du\right] \to \infty \quad \text{and} \quad \limsup_n \frac{\|\Gamma^n\|^2_{\mathcal{L}^2}}{E_Q\left[\int_0^T q_u \int_0^u q_s\Gamma_s^n ds\, du\right]} = L \tag{4.15}$$

or both, for $K, L \in \mathbb{R}$. Indeed, (4.13) would otherwise lead to $\lim_n F(Z^n, \Delta^n, \Gamma^n) = \infty$ and thereby contradict $((Z^n, \Delta^n, \Gamma^n))$ being a minimizing sequence of F. On the other hand

however, an application of Hölders inequality yields

$$\left| E_Q\left[\int_0^T q_u \int_0^u \Delta_s^n ds\, du\right]\right| \leq \left(E_Q\left[\int_0^T |q_u|^2 du\right] E_Q\left[\int_0^T \left(\int_0^u \Delta_s^n ds\right)^2 du\right]\right)^{\frac{1}{2}}$$
$$\leq \|q\|_{\mathcal{L}^2} \left(E_Q\left[\int_0^T \int_0^u |\Delta_s^n|^2 ds du\right]\right)^{\frac{1}{2}} \leq T^{\frac{1}{2}} \|q\|_{\mathcal{L}^2} \|\Delta^n\|_{\mathcal{L}^2} . \quad (4.16)$$

Taking the square on both sides above we obtain

$$\left(E_Q\left[\int_0^T q_u \int_0^u \Delta_s^n ds\, du\right]\right)^2 \leq T \|q\|_{\mathcal{L}^2}^2 \|\Delta^n\|_{\mathcal{L}^2}^2$$

which in turn implies $\|\Delta^n\|_{\mathcal{L}^2}^2 \, (E_Q[\int_0^T q_u \int_0^u \Delta_s^n ds\, du])^{-1} \to \infty$, a contradiction to (4.14). As to (Γ^n), we argue similarly and, for $(Q_u)_{u\in[0,T]}$ defined by $Q_u := \int_u^T q_s ds$ estimate

$$\left| E_Q\left[\int_0^T q_u \int_0^u q_s \Gamma_s^n ds\, du\right]\right| = \left| E_Q\left[\int_0^T q_u \Gamma_u^n Q_u du\right]\right| \leq E_Q\left[\int_0^T |q_u||\Gamma_u^n||Q_u| du\right]$$
$$\leq \left(E_Q\left[\int_0^T |q_u|^2 |Q_u|^2 du\right]\right)^{\frac{1}{2}} \|\Gamma^n\|_{\mathcal{L}^2} = \left(E_Q\left[\int_0^T |q_u|^2 \left|\int_u^T q_s ds\right|^2 du\right]\right)^{\frac{1}{2}} \|\Gamma^n\|_{\mathcal{L}^2}$$
$$\leq \left(E_Q\left[\left(\int_0^T |q_u|^2 du\right)^2\right]\right)^{\frac{1}{2}} \|\Gamma^n\|_{\mathcal{L}^2} = \|q\|_{\mathcal{L}^4}^2 \|\Gamma^n\|_{\mathcal{L}^2} ,$$

where we used Fubini's theorem in the first equality above. Since $\int q dW \in BMO$, the $\mathcal{L}^4$-norm of q is finite[5] and the contradiction to (4.15) is derived analogously to the argumentation following (4.16). Consequently, there exists a sequence $((\tilde{\Delta}^n, \tilde{\Gamma}^n))$ in the asymptotic convex hull of $((\Delta^n, \Gamma^n))$ and $(\Delta^Q, \Gamma^Q) \in \mathcal{L}^2 \times \mathcal{L}^2$ such that $((\tilde{\Delta}^n, \Gamma^n))$ converges in $\mathcal{L}^2 \times \mathcal{L}^2$ to (Δ^Q, Γ^Q). On the side of (Z^n) we pass to the corresponding sequence $(\tilde{Z}^n)$ and recall from the proof of Lemma 3.4 that it is bounded in $\mathcal{L}^2$. Hence, there is a sequence in the asymptotic convex hull of $(\tilde{Z}^n)$, denoted likewise, that converges in $\mathcal{L}^2$ to some $Z^Q = z + \int \Delta^Q du + \int \Gamma^Q dW$. Of course, we pass the corresponding sequence on the side of $((\tilde{\Delta}^n, \Gamma^n))$ without violating the convergence to (Δ^Q, Γ^Q).

Step 3: Lower Semicontinuity and convexity of F. In a next step we show that the earlier defined function $F(Z, \Delta, \Gamma) = E_Q[\int_0^T g_u(Z_u, \Delta_u, \Gamma_u) du - \int_0^T Z_u dW_u]$ is lower semicontinuous and convex on $\mathcal{L}^2 \times \mathcal{L}^2 \times \mathcal{L}^2$ where $Z = z + \int \Delta du + \int \Gamma dW$. Indeed, the part

[5] Recall that BMO can be embedded into any $\mathcal{H}^p$-space , compare [16, Section 2.1, p. 26].

$E_Q[\int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du]$ is lower semicontinuous by (POS), (LSC) and Fatou's lemma. As to the second part, first observe that $\mathcal{L}^2$-convergence of $(\tilde{Z}^n)$ towards Z^Q implies

$$\left| E_Q \left[\int_0^T (\tilde{Z}^n_u - Z^Q_u) dW_u \right] \right| \underset{n\to\infty}{\longrightarrow} 0\,.$$

Furthermore, (CON) yields that F is convex in (Z, Δ, Γ).

Step 4: Minimality of $(Z^Q, \Delta^Q, \Gamma^Q)$. We claim that $F(Z^Q, \Delta^Q, \Gamma^Q) = \inf_{(\Delta,\Gamma)\in\mathcal{L}^2\times\mathcal{L}^2} F(Z, \Delta, \Gamma)$ which would then in turn finally imply (4.12). To this end, it suffices to prove that $F(Z^Q, \Delta^Q, \Gamma^Q) \leq \inf_{(\Delta,\Gamma)\in\mathcal{L}^2\times\mathcal{L}^2} F(Z, \Delta, \Gamma)$, since the reverse inequality is naturally satisfied. Observe now that

$$\begin{aligned}\inf_{(\Delta,\Gamma)\in\mathcal{L}^2\times\mathcal{L}^2} F(Z, \Delta, \Gamma) &= \lim_n F\left(Z^n, \Delta^n, \Gamma^n\right) = \lim_n \sum_{k=n}^{M(n)} \lambda_k^{(n)} F\left(Z^k, \Delta^k, \Gamma^k\right) \\ &\geq \lim_n F\left(\sum_{k=n}^{M(n)} \lambda_k^{(n)} Z^k, \sum_{k=n}^{M(n)} \lambda_k^{(n)} \Delta^k, \sum_{k=n}^{M(n)} \lambda_k^{(n)} \Gamma^k\right) \\ &= \lim_n F\left(\tilde{Z}^n, \tilde{\Delta}^n, \tilde{\Gamma}^n\right) \geq F(Z^Q, \Delta^Q, \Gamma^Q)\end{aligned}$$

where we denoted by $\lambda_k^{(n)}$, $n \leq k \leq M(n)$, $\sum_k \lambda_k^{(n)} = 1$ the convex weights of the sequence $((\tilde{Z}^n \tilde{\Delta}^n, \tilde{\Gamma}^n))$ and made use of the convexity and lower semicontinuity of F.

Step 5: Uniqueness of $(Z^Q, \Delta^Q, \Gamma^Q)$. As to the uniqueness, assume that there are (Δ^1, Γ^1) and (Δ^2, Γ^2) with corresponding Z^1 and Z^2, respectively, both attaining the supremum such that $P \otimes dt[(\Delta^1, \Gamma^1) \neq (\Delta^2, \Gamma^2)] > 0$. Setting $(\bar{Z}, \bar{\Delta}, \bar{\Gamma}) := \frac{1}{2}[(Z^1, \Delta^1, \Gamma^1) + (Z^2, \Delta^2, \Gamma^2)]$ together with $Q \sim P$ and the strict convexity of F inherited by g yields that $F(\bar{Z}, \bar{\Delta}, \bar{\Gamma}) < F(Z^1, \Delta^1, \Gamma^1)$, a contradiction to the optimality of $(Z^1, \Delta^1, \Gamma^1)$. □

Remark 4.6. Since for a given $Q \in \mathcal{Q}$ with $\frac{dQ}{dP} \in L_b^\infty$ and a strictly convex generator the maximizer $(Z^Q, \Delta^Q, \Gamma^Q)$ is unique by the preceding proposition, it has to be (conditionally) optimal at all times $t \in [0, T]$. Indeed, assume to the contrary the existence of (Δ, Γ) such that $z + \int_0^t \Delta_u du + \int_0^t \Gamma_u dW_u = Z_t = Z^Q_t$ and $E_Q[\int_t^T -g_u(Z_u, \Delta_u, \Gamma_u)du + \int_t^T Z_u dW_u|\mathcal{F}_t] > E_Q[\int_t^T -g_u(Z^Q_u, \Delta^Q_u, \Gamma^Q_u)du + \int_t^T Z^Q_u dW_u|\mathcal{F}_t]$ holds true for some $t \in [0, T]$. Then, however, the concatenated processes $(\bar{Z}, \bar{\Delta}, \bar{\Gamma}) := (Z^Q, \Delta^Q, \Gamma^Q)1_{[0,t]} + (Z, \Delta, \Gamma)1_{]t,T]}$ satisfy $E_Q[\int_0^T -g_u(\bar{Z}_u, \bar{\Delta}_u, \bar{\Gamma}_u)du + \int_0^T \bar{Z}_u dW_u] > E_Q[\int_0^T -g_u(Z^Q_u, \Delta^Q_u, \Gamma^Q_u)du + \int_0^T Z^Q_u dW_u]$, which is a contradiction to the opimality of $(Z^Q, \Delta^Q, \Gamma^Q)$ at time zero. ♦

Notice that, for $d = 1$ and the case of a quadratic generator which is in addition independent of z, that is $g_u(\delta, \gamma) = |\delta|^2 + |\gamma|^2$, the processes (Δ^Q, Γ^Q) attaining $\mathcal{E}_0^*(Q)$ can be explicitly computed and (Δ^Q_t, Γ^Q_t) depends on the whole path of q up to time t, as illustrated in the following proposition. It thus constitutes a useful tool for the characterization of the dual optimizers and its proof is closely related to the Euler-Lagrange equation arising in classical calculus of variation.

Proposition 4.7. *Assume that $d = 1$ and that g is defined by $g_u(\delta, \gamma) = |\delta|^2 + |\gamma|^2$. For $Q \in \mathcal{Q}$ with $\frac{dQ}{dP} \in L_b^\infty$, let (Δ^Q, Γ^Q) be the optimizer attaining $\mathcal{E}_0^*(Q)$. Then there exist $c_1, c_2 \in \mathbb{R}$ such that*

$$\Delta_t^Q = -\frac{1}{2} \int_0^t q_s ds + c_1 \tag{4.17}$$

$$\Gamma_t^Q = -\frac{1}{2} q_t \left(\int_0^t q_s ds + c_2 \right) , \tag{4.18}$$

for all $t \in [0, T]$.

Proof. For the purpose of this proof we assume without loss of generality that $z = 0$, since the initial value does not affect the optimization with respect to (Δ, Γ). Hence, the generator only depending on δ and γ in combination with (4.7) gives $\mathcal{E}_0^*(Q) = -\inf_{(\Delta,\Gamma)}\{F_1(\Delta) + F_2(\Gamma)\}$ where $F_1(\Delta) := E_Q[\int_0^T (|\Delta_u|^2 - q_u \int_0^u \Delta_s ds) du]$ and $F_2(\Gamma) := E_Q[\int_0^T (|\Gamma_u|^2 - q_u \int_0^u q_s \Gamma_s ds) du]$. We will proceed along an ω-wise criterion of optimality, since any pair (Δ^Q, Γ^Q) that is optimal for almost all $\omega \in \Omega$ then naturally also optimizes the expectation under Q. The uniqueness obtained in Proposition 4.5 then assures that the path-wise optimizer is the only one. We define

$$J_1(\Delta) = \int_0^T \left(|\Delta_u|^2 - q_u \int_0^u \Delta_s ds \right) du \quad \text{and} \quad J_2(\Gamma) = \int_0^T \left(|\Gamma_u|^2 - q_u \int_0^u q_s \Gamma_s ds \right) du$$

and observe that it is sufficient to elaborate how to obtain conditions for a minimizer of J_1, as the functional J_2 is of a similar structure. Introducing $X(u) := \int_0^u \Delta_s ds$ we obtain $X'(u) := \frac{d}{du} X(u) = \Delta_u$ and

$$J_1(\Delta) = \tilde{J}_1(X) = \int_0^T L(u, X(u), X'(u)) du$$

where $L(u, a, b) = |b|^2 - q_u a$. If X is a local minimum of $\tilde{J}_1$, then $\tilde{J}_1(X) \leq \tilde{J}_1(X + \varepsilon\eta)$ for sufficiently small $\varepsilon > 0$ and all differentiable $\eta \in C([0, T], \mathbb{R})$ the derivatives of which are square integrable and which satisfy $\eta(0) = \eta(T) = 0$. In particular, with $\phi(\varepsilon) := \tilde{J}_1(X + \varepsilon\eta)$, it has to hold that $\frac{d}{d\varepsilon}\phi(\varepsilon)_{|_{\varepsilon=0}} = 0$. Using the specific form of L we get

$$\begin{aligned} \frac{d}{d\varepsilon}\phi(\varepsilon)_{|_{\varepsilon=0}} &= \lim_{h \to 0} \frac{1}{h} \int_0^T \big[L\big(u, X(u) + h\eta(u), X'(u) + h\eta'(u)\big) - L\big(u, X(u), X'(u)\big) \big] du \\ &= \lim_{h \to 0} \int_0^T \big[-q_u \eta(u) + 2X'(u)\eta'(u) + h(\eta'(u))^2 \big] du \,. \end{aligned}$$

Having assumed η' to be square integrable allows us to exchange limit and integration, yielding

$$0 = \frac{d}{d\varepsilon}\phi(\varepsilon)_{|_{\varepsilon=0}} = \int_0^T \left[-q_u\eta(u) + 2\Delta_u\eta'(u)\right] du\,. \tag{4.19}$$

Using integration by parts we obtain

$$-\int_0^T q_u\eta(u)du = \left(\int_0^u -q_s ds\right)\eta(u)\Big|_0^T - \int_0^T\left(\int_0^u -q_s ds\right)\eta'(u)du\,.$$

The first term on the right-hand side above vanishes and so, by plugging this back into (4.19) we end up with

$$\int_0^T \left(2\Delta_u + \int_0^u q_s ds\right)\eta'(u)du = 0\,. \tag{4.20}$$

Let us next introduce the constant $c := \frac{1}{T}\int_0^T (2\Delta_u + \int_0^u q_s ds)du$ and observe that, using $\int_0^T \eta'(u)du = 0$, Equation (4.20) may be rewritten as

$$\int_0^T \left(2\Delta_u + \int_0^u q_s ds - c\right)\eta'(u)du = 0\,. \tag{4.21}$$

Moreover, the function

$$\bar\eta(t) := \int_0^t \left(2\Delta_u + \int_0^u q_s ds - c\right) du$$

satisfies $\bar\eta(0) = \bar\eta(T) = 0$ by construction as well as $\bar\eta'(u) = 2\Delta_u + \int_0^u q_s ds - c$, which is square integrable for almost all $\omega \in \Omega$, since Δ and q are square integrable[6]. Hence, (4.21) applied to our particular function $\bar\eta$ yields

$$\int_0^T \left(2\Delta_u + \int_0^u q_s ds - c\right)^2 du = 0$$

and we deduce that

$$2\Delta_u + \int_0^u q_s ds = \frac{1}{T}\int_0^T \left(2\Delta_r + \int_0^r q_s ds\right) dr \qquad \text{for almost all } u \in [0,T]\,. \tag{4.22}$$

[6]More precisely, it holds $P(\int_0^T |q_u|^2 du < \infty) = P(\int_0^T |\Delta_u|^2 du < \infty) = 1$.

Specifically, (4.22) shows that, for almost all $\omega \in \Omega$, there exists a set $\mathcal{I}(\omega) \subseteq [0,T]$ with Lebesgue measure T such that, for all $u \in \mathcal{I}(\omega)$,

$$2\Delta_u + \int_0^u q_s ds = M\,, \tag{4.23}$$

where M of course depends on $\omega \in \Omega$ and is thus a random variable. This in turn implies, for dt-almost all $u \in [0,T]$, the existence of $\Omega_u \subseteq \Omega$ with $P(\Omega_u) = 1$ such that (4.23) holds for all $\omega \in \Omega_u$. In particular, on Ω_u the above M equals an $\mathcal{F}_u$-measurable random variable. We choose a sequence $(u_n) \subset [0,T]$ with $\lim_n u_n = 0$ and hence obtain that on $\bar{\Omega} := \bigcap_n \Omega_{u_n}$, where $P(\bar{\Omega}) = Q(\bar{\Omega}) = 1$, M equals an $\bigcap_n \mathcal{F}_{u_n}$-measurable random variable and is thus $\mathcal{F}_0$-measurable by the right-continuity of our filtration, that means it is a constant on $\bar{\Omega}$ due to Blumenthal's zero-one law. Since modifying our optimizer on a Q-nullset does not alter the value of the functional to be optimized, we have shown (4.17) by putting $c_1 := \frac{M}{2}$.

As to the case of our optimal Γ^Q, assume first that $q_u \neq 0$ for all $u \in [0,T]$ and observe that, with $Y(u) := \int_0^u q_s \Gamma_s ds$, we obtain $J_2(\Gamma) = \tilde{J}_2(Y) = \int_0^T K(u, Y(u), Y'(u))du$ where we set $K(u,a,b) = (1/q_u)^2 b^2 - q_u a$. Thus, an argumentation identical to that above would yield (4.18) in that case. Furthermore, it holds that $\lim_{q_u \to 0} \Gamma_u^Q = 0$, a value that is consistent with the "pointwise" minimization consideration that $\arg\min_{\Gamma_u}\{|\Gamma_u|^2 - q_u \int_{[0,u)} q_s \Gamma_s ds\}|_{q_u=0} = 0$ and hence justifies expression (4.18). □

For $\xi \in L^1$ such that $\mathcal{A}(\xi, g, z) \neq \emptyset$, the ensuing theorem states that, given the existence of an equivalent probability measure $\hat{Q} \in \mathcal{Q}$ such that the sup in (4.2) is attained, the BSDE with generator g and terminal condition ξ admits a solution under constraints.

Theorem 4.8. *Assume that, for $\xi \in L^1$ with $\mathcal{A}(\xi,g,z) \neq \emptyset$, there exists a $\hat{Q} \in \mathcal{Q}$ with $\frac{dQ}{dP} \in L_b^\infty$ such that $\mathcal{E}_0^g(\xi) = E_{\hat{Q}}[\xi] - \mathcal{E}_0^*(\hat{Q})$. Then there exists a solution $(Y,Z) \in \mathcal{A}(\xi,g,z)$ of the BSDE with parameters (ξ, g).*

Proof. Starting with (4.2) in combination with Proposition 4.5, it holds

$$\mathcal{E}_0^g(\xi) = E_{\hat{Q}}[\xi] - \mathcal{E}_0^*(\hat{Q}) = E_{\hat{Q}}\left[\xi + \int_0^T g_u(Z_u^{\hat{Q}}, \Delta_u^{\hat{Q}}, \Gamma_u^{\hat{Q}})du - \int_0^T Z_u^{\hat{Q}} dW_u\right]. \tag{4.24}$$

We recall that $\mathcal{A}(\xi,g,z) \neq \emptyset$ and $\xi \in L^1$. Hence, by Theorem 3.7 there exists (Δ, Γ, Z) such that

$$E_{\hat{Q}}\left[\xi + \int_0^T g_u(Z_u^{\hat{Q}}, \Delta_u^{\hat{Q}}, \Gamma_u^{\hat{Q}})du - \int_0^T Z_u^{\hat{Q}} dW_u\right] - \int_0^T g_u(Z_u, \Delta_u, \Gamma_u)du + \int_0^T Z_u dW_u \geq \xi$$

holds true. Taking expectation under $\hat{Q}$ on both sides of the inequality above yields

$$E_{\hat{Q}}\left[-\int_0^T g_u(Z_u^{\hat{Q}},\Delta_u^{\hat{Q}},\Gamma_u^{\hat{Q}})du+\int_0^T Z_u^{\hat{Q}}dW_u\right] \\ \leq E_{\hat{Q}}\left[-\int_0^T g_u(Z_u,\Delta_u,\Gamma_u)du+\int_0^T Z_u dW_u\right].$$

However, the expression on the left-hand side is maximal for $(Z^{\hat{Q}},\Delta^{\hat{Q}},\Gamma^{\hat{Q}})$ by means of Proposition 4.5 and thus equality has to hold. Hence, it follows that $\mathcal{E}_0^*(\hat{Q}) = E_{\hat{Q}}[-\int_0^T g_u(Z_u,\Delta_u,\Gamma_u)du+\int_0^T Z_u dW_u]$. By plugging this back into (4.24) we obtain

$$\mathcal{E}_0^g(\xi) = E_{\hat{Q}}[\xi]+E_{\hat{Q}}[\int_0^T g_u(Z_u,\Delta_u,\Gamma_u)du-\int_0^T Z_u dW_u]$$

which is equivalent to

$$E_{\hat{Q}}\left[\mathcal{E}_0^g(\xi)-\int_0^T g_u(Z_u,\Delta_u,\Gamma_u)du+\int_0^T Z_u dW_u-\xi\right]=0\,.$$

Since the expression within the expectation is P- and thereby also $\hat{Q}$-almost surely positive, we finally conclude that

$$\mathcal{E}_0^g(\xi)-\int_0^T g_u(Z_u,\Delta_u,\Gamma_u)du+\int_0^T Z_u dW_u=\xi\,,$$

and thus $(\mathcal{E}_0^g(\xi)-\int_0^\cdot g(Z,\Delta,\Gamma)du+\int_0^\cdot ZdW,Z)$ constitutes a solution of the BSDE with parameters (ξ,g). □

References

[1] P. Briand and Y. Hu. BSDE with quadratic growth and unbounded terminal value. *Probability Theory and Related Fields*, 136(4):604–618, 2006.

[2] M. Broadie, J. Cvitanic, and H. M. Soner. Optimal Replication of Contingent Claims under Portfolio Constraints. *The Review of Financial Studies*, 11:59–79, 1998.

[3] P. Cheridito, H. M. Soner, and N. Touzi. The multi-dimensional super-replication problem under gamma constraints. *Annales de l'Institut Henri Poincaré (C) Analyse Non Linéaire*, 22(5):633–666, 2005.

[4] P. Cheridito, H. M. Soner, N. Touzi, and N. Victoir. Second Order Backward Stochastic Differential Equations and Fully Non-Linear Parabolic PDEs. *Communications in Pure and Applied Mathematics*, 60(7):1081–1110, 2007.

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

[6] F. Delbaen and W. Schachermayer. A General Version of the Fundamental Theorem of Asset Pricing. *Mathe. Annalen*, 300:463–520, 1994.

[7] F. Delbaen and W. Schachermayer. A Compactness Principle for Bounded Sequences of Martingales with Applications. *Proceedings of the Seminar of Stochastic Analysis, Random Fields and Applications, Progress in Probability, 133–173*, Birkhäuser, 1996.

[8] F. Delbaen, Y. Hu, and X. Bao. Backward SDEs with Superquadratic Growth. *Probability Theory and Related Fields*, 150(1-2):145–192, 2011.

[9] S. Drapeau, G. Heyne, and M. Kupper. Minimal Supersolutions of Convex BSDEs. Forthcoming in *Annals of Probability*, 2013.

[10] S. Drapeau, M. Kupper, E. R. Gianin, and L. Tangpi. Dual Representation of Minimal Supersolutions of Convex BSDEs. *arXiv e-prints*, 2013.

[11] N. El Karoui, S. Peng, and M. C. Quenez. Backward Stochastic Differential Equations in Finance. *Mathematical Finance*, 7(1):1–71, 1997.

[12] G. Heyne. *Essays on Minimal Supersolutions of BSDEs and on Cross Hedging in Incomplete Markets*. PhD thesis, Humboldt-Universität zu Berlin, 2012.

[13] G. Heyne, M. Kupper, and C. Mainberger. Minimal Supersolutions of BSDEs with Lower Semicontinuous Generators. Forthcoming in *Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques*, 2012.

[14] E. Jouini and H. Kallal. Arbitrage in securities markets with short-sales constraints. *Mathematical Finance*, 5(3):197–232, 1995.

[15] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus (Graduate Texts in Mathematics)*. Springer, August 2004.

[16] N. Kazamaki. *Continuous Exponential Martingales and BMO*, volume 1579 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1994. ISBN 3-540-58042-5.

[17] M. Kobylanski. Backward Stochastic Differential Equations and Partial Differential Equations with Quadratic Growth. *Annals of Probability*, 28(2):558–602, 2000.

[18] E. Pardoux and S. Peng. Adapted Solution of a Backward Stochastic Differential Equation. *System & Control Letters*, 14(1):55–61, 1990.

[19] S. Peng. Monotonic Limit Theorem of BSDE and Nonlinear Decomposition Theorem of Doob–Meyer's Type. *Probability Theory and Related Fields*, 113(4):473–499, 1999.

[20] P. E. Protter. *Stochastic Integration and Differential Equations*. Springer, 2nd edition, 2005.

[21] H. M. Soner and N. Touzi. Superreplication under Gamma constraints. *SIAM Journal on Control and Optimization*, 39(1):73–96, 2000.

[22] H. M. Soner, N. Touzi, and J. Zhang. Dual Formulation of Second Order Target Problems. *Annals of Applied Probability*, 23(1):308–347, 2013.