



HAL
open science

Optimal position targeting via decoupling fields

Stefan Ankirchner, Alexander Fromm, Thomas Kruse, Alexandre Popier

► **To cite this version:**

Stefan Ankirchner, Alexander Fromm, Thomas Kruse, Alexandre Popier. Optimal position targeting via decoupling fields. 2017. hal-01500311v1

HAL Id: hal-01500311

<https://hal.science/hal-01500311v1>

Preprint submitted on 3 Apr 2017 (v1), last revised 16 Apr 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Optimal position targeting via decoupling fields

Stefan Ankirchner^{*1}, Alexander Fromm^{†1}, Thomas Kruse^{‡2} and Alexandre Popier^{§3}

¹Institute for Mathematics, University of Jena, Ernst-Abbe-Platz 2, 07743 Jena, Germany

²Faculty of Mathematics, University of Duisburg-Essen, Thea-Leymann-Str. 9, 45127 Essen, Germany

³Laboratoire Manceau de Mathématiques, Université du Maine, Avenue O. Messiaen, 72085 Le Mans cedex 9, France

March 28, 2017

Abstract

We consider a variant of the basic problem of the calculus of variations, where the Lagrangian is convex and subject to randomness adapted to a Brownian filtration. We solve the problem by reducing it, via a limiting argument, to an unconstrained control problem that consists in finding an absolutely continuous process minimizing the expected sum of the Lagrangian and the deviation of the terminal state from a given target position. Using the Pontryagin maximum principle we characterize a solution of the unconstrained control problem in terms of a fully coupled forward-backward stochastic differential equation (FBSDE). We use the method of decoupling fields for proving that the FBSDE has a unique solution.

Introduction

The basic problem of the calculus of variations consists in minimizing an integral functional over a set of functions satisfying an initial and terminal condition. In this article we consider a version of the basic problem, where the Lagrangian is convex and subject to random influences supported by a Brownian motion W on a probability space (Ω, \mathcal{F}, P) . More precisely, let $T \in (0, \infty)$ and $f : \Omega \times [0, \infty) \times \mathbb{R}^2 \rightarrow \mathbb{R}$ be a function, convex in the last two variables, such that for all $(x, a) \in \mathbb{R}^2$ the mapping $(\omega, t) \mapsto f(\omega, t, x, a)$ is progressively measurable with respect to $(\mathcal{F}_t)_{t \in [0, \infty)}$, the augmented filtration generated by the Brownian motion W . We show, under some additional analytic assumptions, existence of a solution of the following problem:

$$\text{Minimize } \hat{J}(X) = \mathbb{E} \left[\int_0^T f(t, X_t, \dot{X}_t) ds \right] \text{ over all absolutely continuous} \quad (\text{P})$$

and progr. mb. processes X satisfying $X_0 = x_0 \in \mathbb{R}$ and $X_T = 0$.

^{*}s.ankirchner@uni-jena.de

[†]alexander.fromm@uni-jena.de

[‡]thomas.kruse@uni-due.de

[§]apopier@univ-lemans.fr

We interpret t as time, X_t as the state and \dot{X}_t as the velocity at time t .

Minimizing $\hat{J}(X)$ is a classical problem with many applications e.g. in physics, economics and engineering. We refer to the scripts of Gelfand and Fomin [11], Clarke [6] and Evans [8] for explicit applications and an overview on the *deterministic* version of the basic problem. Stochastic examples of problem (P) have been recently analyzed in the context of closing financial asset positions in illiquid markets (see e.g. the introduction in [13] for an overview). In these applications f includes transaction costs, depending on the liquidation rate \dot{X} ; moreover f can incorporate measures of the risk exposure, depending on the volume X_t of the remaining position.

In order to prove existence of a process X minimizing the functional $\hat{J}(X)$ we study also a related control problem without the constraint $X_T = 0$, but with an additional term in the cost functional penalizing any deviation of X_T from zero. Let $g : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ be a function, convex in the second variable, such that for all $x \in \mathbb{R}$ the mapping $\omega \mapsto g(\omega, x)$ is \mathcal{F}_T -measurable. We solve, under some nice analytic assumptions, the following unconstrained control problem:

$$\text{Minimize } J(X) = \mathbb{E} \left[\int_0^T f(t, X_t, \dot{X}_t) ds + g(X_T) \right] \text{ over all absolutely continuous} \quad (\text{UP})$$

and progr. mb. processes X satisfying $X_0 = x_0$.

We show that by setting the penalty function equal to $g(x) = Lx^2$ and letting $L \rightarrow \infty$ one can reduce the constrained problem (P) to the unconstrained one (UP). We think that problem (UP) is interesting in itself and therefore we study it for a general class of convex functions g .

By following a classic Bellman approach for solving (UP) (at least if f and g are deterministic functions), one obtains a non-linear Hamilton-Jacobi-Bellman (HJB) equation that is difficult to solve. By the Pontryagin maximum principle an optimal solution of (UP) can be characterized in terms of a forward-backward stochastic differential equation (FBSDE), where the forward component describes the optimal state dynamics and the backward component the dynamics of the so-called costate. The FBSDE for (UP) takes the form

$$\begin{aligned} X_t &= x - \int_0^t f_y^*(s, X_s, Y_s) ds, \\ Y_t &= g'(X_T) - \int_t^T Z_s dW_s + \int_t^T f_x(s, X_s, f_y^*(s, X_s, Y_s)) ds, \end{aligned} \quad (1)$$

where $f^*(t, x, \cdot)$ denotes the convex conjugate of the function $a \mapsto f(t, x, a)$, f_y^* its derivative w.r.t. y and f_x the derivative of f w.r.t. x . Notice that the FBSDE (1) is fully coupled, i.e. the forward dynamics depend on the backward component Y , and the backward dynamics on the forward part X . It is a longstanding challenge to find conditions guaranteeing that a fully coupled FBSDE possesses a solution. Sufficient conditions are provided e.g. in [15], [19], [18], [20], [7], [16] (see also references therein). The method of decoupling fields, developed in [9] (see also the precursor articles [17], [10] and [16]), is practically useful for determining whether a solution exists. A decoupling field describes the functional dependence of the backward part Y on the forward component X . If the coefficients of a fully coupled FBSDE satisfy a Lipschitz condition, then there exists a maximal non-vanishing interval possessing a solution triplet (X, Y, Z) and a decoupling field with nice regularity properties. The method of decoupling fields consists in analyzing the dynamics of the decoupling field's gradient in order to determine whether the FBSDE has a solution on the whole time interval $[0, T]$.

The main idea of the present article is to use the method of decoupling fields in order to prove existence and uniqueness of a solution to the particular FBSDE (1) (see Theorem 3.1,

Corollary 3.3 and Theorem 3.6). With a solution at hand we can then solve the unconstrained problem (UP) (see Proposition 1.6) and finally also problem (P) (see Theorem 4.5).

The method works since the state derivative of the decoupling field of the FBSDE (1) turns out to be bounded (under some nice analytic assumptions). In particular, the decoupling field can not explode, which further implies that there exists a solution of (1) on the whole interval $[0, T]$.

Solutions of problem (P) and (UP) have been obtained under some additional structural assumptions on the function f . One focus of the literature so far is set on cost functions f that are *additive*. In [2] it is assumed that f takes the form $f(t, x, a) = \gamma_t|x|^p + \eta_t|a|^p$, where $p > 1$ and (η, γ) is a pair of non-negative progressively measurable processes. The particular form allows to decouple the FBSDE (1), after a variable change. As the penalty of any deviation of X_T from 0 increases to infinity, the backward part of the decoupled FBSDE converges to a solution of a BSDE with singular terminal condition, a concept developed in [21], [22]. In the setting of the present article, the solution processes (X, Y, Z) of (1) also converge as the penalty tends to infinity. The limiting processes turn out to be a solution of a coupled pair of stochastic differential equations, where an initial and terminal condition is imposed on the first equation, but no condition on the second (see Theorem 4.7).

Solving a fully coupled FBSDE is also not necessary in the case where f is additive and linear-quadratic (see [3] and [4]), and even in some additive and polynomial cases with a Poisson measure as an additional source of randomness (see [12] and [14]).

The article is organized as follows. In Section 1 we specify some basic assumptions on the functions f and g under which we solve problem (UP) and (P) respectively. Moreover, we explain how to obtain a solution of (UP) from a solution (X, Y, Z) of the FBSDE (1). Section 2 provides an overview of the method of decoupling fields. In Section 3 we apply the method to solve problem (UP). Finally, in Section 4 we solve problem (P) by reducing it to (UP).

1 Problem formulation via FBSDEs

Let $T > 0$ be a deterministic finite time horizon. Let W be a Brownian motion on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and denote by $(\mathcal{F}_t)_{t \in [0, T]}$ the smallest filtration satisfying the usual conditions and containing the filtration generated by W .

Let $A \subseteq \mathbb{R}$ be a closed and connected set of possible control values satisfying $\inf A \leq 0 < \sup A$. Let

$$g : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$$

be measurable and

$$f : \Omega \times [0, T] \times \mathbb{R} \times A \rightarrow \mathbb{R}$$

be measurable such that for all $(x, a) \in \mathbb{R} \times A$ the mapping $(\omega, t) \mapsto f(\omega, t, x, a)$ is progressively measurable. We make the following additional assumptions on f and g :

- (C0) For every fixed pair $(\omega, t) \in \Omega \times [0, T]$ the mappings $(x, a) \mapsto f(t, x, a)$ and $x \mapsto g(x)$ are convex, with f being strictly convex in a . Note that we follow the usual convention and omit the function argument ω .
- (C1) The mappings $A \ni a \mapsto f(t, x, a)$ and $\mathbb{R} \ni x \mapsto f(t, x, 0)$ attain a minimum at zero (for all ω, t, x and all ω, t respectively). We also assume that $f(t, 0, 0) = 0$ for all $t \in [0, T]$. Observe that f is then non-negative.

(C2) f is coercive, i.e. there exist $p > 1$ and $b > 0$ such that

$$\forall a \in A : f(t, x, a) \geq b|a|^p$$

independently of ω , t and x .

(C3) $g(\cdot)$ restricted to $[0, \infty)$ is twice continuously differentiable, $f(t, \cdot, \cdot)$ restricted to $[0, \infty) \times (A \cap [0, \infty))$ is continuously differentiable, while $f_x(t, \cdot, \cdot)$ and $f_a(t, \cdot, \cdot)$ are continuously differentiable on $[0, \infty) \times A_+$, where $A_+ := A \cap (0, \infty)$. Furthermore, for all compacts $K_1 \subset [0, \infty)$, $K_2 \subseteq A \cap [0, \infty)$ and $K_3 \subset A_+$:

$$\sup_{(\omega, t, x, a) \in \Omega \times [0, T] \times K_1 \times K_2} ((|g_x| + |g_{xx}|)(\omega, x) + (|f_x| + |f_a|)(\omega, t, x, a)) < \infty,$$

$$\sup_{(\omega, t, x, a) \in \Omega \times [0, T] \times K_1 \times K_3} (|f_{xx}| + |f_{xa}| + |f_{ax}| + |f_{aa}|)(\omega, t, x, a) < \infty.$$

(C4) Finally, we assume that the mapping $x \mapsto g(x)$ attains its minimum at zero (for all ω).

We also assume that $g(0) = 0$. Observe that g is then non-negative.

Remark 1.1. Note that the assumptions that $f(t, 0, 0) = 0$ for all $t \in [0, T]$ and $g(0) = 0$ can be relaxed to the assumptions that $f(\cdot, 0, 0) \in L^1(\Omega \times [0, T])$ and $g(0) \in L^1(\Omega)$. Indeed, in this case one can consider the problems (UP) and (P) with $\tilde{f}(t, x, a) = f(t, x, a) - f(t, 0, 0)$ and $\tilde{g}(x) = g(x) - g(0)$ instead of f and g and add $\mathbb{E} \left[\int_0^T f(t, 0, 0) dt + g(0) \right]$ outside the minimization problem.

For $x \in \mathbb{R}$ we define $\mathcal{A}(x)$ as the set of all progressively measurable $\alpha : \Omega \times [0, T] \rightarrow A$ such that

$$X_s^{x, \alpha} := x - \int_0^s \alpha_r dr \tag{2}$$

is well-defined for all $s \in [0, T]$. We can reformulate Problem (UP) as follows:

$$\text{Minimize } J(x, \alpha) := \mathbb{E} \left[\int_0^T f(s, X_s^{x, \alpha}, \alpha_s) ds + g(X_T^{x, \alpha}) \right] \text{ over all } \alpha \in \mathcal{A}(x). \tag{3}$$

For simplicity we sometimes write X^α or just X instead of $X^{x, \alpha}$. So, for given x , the goal is to choose the control α from the set $\mathcal{A}(x)$ of admissible controls in such a way that J is minimized.

The next result shows that when starting with a non-negative initial position, then it can not be optimal to choose α such that the position process is increasing or negative at some time point. This result is coherent with the absence of transaction-triggered price manipulation (see [1]). Note that for proving this statement we only need (C0), (C1) and (C4), but not (C2) and (C3).

Proposition 1.2. *Let $x \geq 0$. If $\alpha \in \mathcal{A}(x)$ is optimal, then $X_s^\alpha = x - \int_0^s \alpha_r dr$, $s \in [0, T]$, is non-increasing and non-negative. Moreover, for any $\alpha \in \mathcal{A}(x)$ there exists $\beta \in \mathcal{A}(x)$ such that X^β is non-increasing and non-negative and $J(x, \beta) \leq J(x, \alpha)$.*

Proof. For $\beta \in \mathcal{A}(x)$ we write $\mathcal{J}(x, \beta) := \int_0^T f(s, X_s^{x, \beta}, \beta_s) ds + g(X_T^{x, \beta})$.

Define the measurable set $B := \{\inf_{s \in [0, T]} X_s^\alpha < 0\}$ and the stopping time

$$\sigma := \inf\{s \in [0, T] \mid X_s < 0\}.$$

Note that $B = \{\sigma < T\}$ and that σ is equal to ∞ on B^c . Furthermore, $X_\sigma^\alpha = 0$ on B .

Now, define β as the strategy which coincides with α on $[0, \sigma \wedge T]$, but is 0 on the interval $(\sigma \wedge T, T]$. Notice that $\alpha = \beta$ on B^c . Now let $\omega \in B$. Then $X_s^\beta(\omega) = 0$ for all $s \in [\sigma(\omega), T]$, but $X_s^\beta(\omega) = X_s^\alpha(\omega)$ for $s \in [0, \sigma(\omega)]$. Also, for $s \in [\sigma(\omega), T]$,

$$f(s, X_s^\alpha, \alpha_s)(\omega) - f(s, X_s^\beta, \beta_s)(\omega) = f(s, X_s^\alpha, \alpha_s)(\omega) - f(s, 0, 0)(\omega) \geq 0,$$

due to (C1). Note that for $s > \sigma(\omega)$ in a sufficiently small neighbourhood of $\sigma(\omega)$ the left hand side of the previous inequality is strictly larger than zero. Moreover, (C4) yields

$$g(X_T^\alpha)(\omega) \geq g(X_T^\beta)(\omega) = g(0) = 0.$$

Therefore, $\mathcal{J}(x, \alpha)(\omega) > \mathcal{J}(x, \beta)(\omega)$ for all $\omega \in B$, which due to the optimality of α can only mean $\mathbb{P}(B) = 0$. In other words X^α is non-negative a.s.

Next, define stopping times

$$\rho := \inf\left\{s \in [0, T] \mid X_s^\alpha > \inf_{r \in [0, s]} X_r^\alpha\right\}$$

and

$$\tau := \inf\left\{s \in [0, T] \mid s > \rho \text{ and } X_s^\alpha = \inf_{r \in [0, s]} X_r^\alpha\right\}.$$

Now let γ be the strategy which is zero on $(\rho \wedge T, \tau \wedge T]$, but is equal α otherwise. Note that $(\rho \wedge T, \tau \wedge T]$ is empty and $\alpha = \gamma$ on

$$\widehat{B} := \left\{ \sup_{s \in [0, T]} \left(X_s^\alpha - \inf_{r \in [0, s]} X_r^\alpha \right) \leq 0 \right\},$$

the event where X^α is non-increasing.

Now let $\omega \in \widehat{B}^c$. Then $\rho(\omega) < T$ and $(\rho(\omega), \tau(\omega) \wedge T]$ is non-empty with $\gamma(\omega)$ vanishing on this interval and $X^\gamma(\omega)$ being constant $X_{\rho(\omega)}^\alpha(\omega)$ on this interval. Furthermore, condition (C1) implies that for all $s \in (\rho(\omega), \tau(\omega) \wedge T]$:

$$\begin{aligned} f(s, X_s^\alpha, \alpha_s)(\omega) - f(s, X_s^\gamma, \gamma_s)(\omega) &= (f(s, X_s^\alpha, \alpha_s)(\omega) - f(s, X_s^\alpha, 0)(\omega)) \\ &\quad + (f(s, X_s^\alpha, 0)(\omega) - f(s, X_s^\gamma, 0)(\omega)) \\ &\geq 0. \end{aligned}$$

Note that for $s > \rho(\omega)$ in a sufficiently small neighbourhood of $\rho(\omega)$ we have a strict inequality. Therefore, $\int_0^T f(s, X_s^\alpha, \alpha_s)(\omega) ds > \int_0^T f(s, X_s^\gamma, \gamma_s)(\omega) ds$. Moreover, (C4) and the convexity of g imply that it is non-decreasing on the positive half-line, such that

$$g(X_T^\alpha)(\omega) \geq g(X_T^\gamma)(\omega),$$

and hence $\mathcal{J}(x, \alpha)(\omega) > \mathcal{J}(x, \gamma)(\omega)$. Due to optimality of α we have $\mathbb{P}(\widehat{B}^c) = 0$. Therefore, the position process of an optimal strategy is non-increasing.

The second statement of the proposition can be shown with similar arguments: if X^α is not non-increasing, then we can construct β such that X^β is non-increasing and $J(x, \beta) \leq J(x, \alpha)$. If in addition X^β is not non-negative, then one can transform β to γ as above so that X^γ is non-negative, non-increasing and $J(x, \gamma) \leq J(x, \beta)$. \square

Proposition 1.2 shows that if the initial position is positive, then it is sufficient to consider only non-negative controls and positions. By symmetry, when starting in a negative position, one can restrict the analysis to non-positive controls and positions, with straightforward adjustments in the hypothesis (C3) (differentiability condition for non positive values). In the following we consider only the positive case and always assume that any positions and controls are non-negative.

The so-called Hamiltonian of the control problem (3) is defined by

$$\mathcal{H}(t, x, a, y) := -ay + f(t, x, a),$$

for $t \in [0, T]$ and $(x, a, y) \in \mathbb{R} \times A \times \mathbb{R}$. Notice that

$$\min_{a \in A} \mathcal{H}(t, x, a, y) = -f^*(t, x, y), \quad (4)$$

where $f^*(t, x, \cdot)$ is the convex conjugate of $f(t, x, \cdot)$. Observe that condition (C2) guarantees that f^* assumes real values only. For the following observation we need both (C2) and (C3).

Remark 1.3. For all $x, y \geq 0$ the following consideration holds: The minimum in (4) is attained at $a = f_y^*(t, x, y)$, where f_y^* denotes the partial derivative of f^* with respect to y . This partial derivative exists due to differentiability of f w.r.t. the last parameter a and (C2). More precisely, using Fermat's theorem applied to the minimization problem $\min_{a \in A} \mathcal{H}(t, x, a, y) = \min_{a \in A_+} \mathcal{H}(t, x, a, y)$ one can deduce

$$a = f_y^*(t, x, y) = f_a^{-1}(t, x, y), \quad \text{if } y \in (f_a(t, x, 0), f_a(t, x, a_{\text{sup}})),$$

where $a_{\text{sup}} := \sup A$ and where $f_a^{-1}(t, x, \cdot)$ denotes the inverse of the function $f_a(t, x, \cdot)$, which is strictly increasing.

Note that for $y \in [0, f_a(t, x, 0)]$ we have $f_y^*(t, x, y) = 0$ and for $y \in [f_a(t, x, a_{\text{sup}}), \infty)$ we have $f_y^*(t, x, y) = a_{\text{sup}}$. This means that we can extend $f_a^{-1}(t, x, \cdot)$ to the whole of $[0, \infty)$ canonically, s.t. $f_y^*(t, x, y) = f_a^{-1}(t, x, y)$ holds everywhere and this extended $f_a^{-1}(t, x, \cdot)$ is still continuous and non-decreasing.

Finally notice the following straightforwardly verifiable properties of f_y^* :

- $f_y^*(t, x, y) \geq 0$ for all $x, y \geq 0$,
- $f_y^*(t, x, 0) = 0$ for all $x \geq 0$,
- $f_y^*(t, x, \cdot)$, which is defined on $[0, \infty)$, has $A \cap [0, \infty)$ as its range, regardless of ω, t and $x \geq 0$.

The following observation will be of importance later on:

Remark 1.4. Assume that $f_{aa}(t, x, a) > 0$ holds for all $(\omega, t, x, a) \in \Omega \times [0, T] \times [0, \infty) \times A_+$ and that f_y^* is weakly differentiable w.r.t. (x, y) on $[0, \infty)^2$. By applying the chain rule to the equation

$$f_a(t, x, f_y^*(t, x, y)) = f_a(t, x, 0) \vee y \wedge f_a(t, x, a_{\text{sup}})$$

we obtain

$$(f_{yx}^*(t, x, y), f_{yy}^*(t, x, y)) = \begin{cases} (0, 0), & \text{if } y \notin (f_a(t, x, 0), f_a(t, x, a_{\text{sup}})), \\ \left(-\frac{f_{ax}}{f_{aa}}(t, x, f_y^*(t, x, y)), \frac{1}{f_{aa}}(t, x, f_y^*(t, x, y)) \right), & \text{otherwise.} \end{cases} \quad (5)$$

We remark that in order to have $f_{aa}(t, x, a) > 0$ for all $(x, a) \in [0, \infty) \times A_+$ it is sufficient to require local Lipschitz continuity of $f_y^*(t, x, \cdot)$ for all ω, t and all $x \geq 0$. Indeed, for fixed ω, t, x the above formula for $f_{yy}^*(t, x, y)$ is satisfied on the set of all $y \geq 0$ satisfying $f_{aa}(t, x, f_y^*(t, x, y)) > 0$ and this set cannot be empty (due to strict convexity). Due to the continuity of $f_{aa}(t, x, f_y^*(t, x, \cdot))$ and the assumption of local boundedness of $f_{yy}^*(t, x, \cdot)$, the reciprocal value of $f_{aa}(t, x, f_y^*(t, x, \cdot))$ must remain bounded on compact subsets of $[0, \infty)$. In particular $f_{aa}(t, x, f_y^*(t, x, y)) > 0$ for all $y \geq 0$, which means $f_{aa}(t, x, a) > 0$ for all $a \in A_+$.

Next we consider the so-called adjoint forward-backward stochastic differential equation (FBSDE) for the control problem (3), given by

$$\begin{aligned} X_s^x &= x - \int_0^s f_y^*(r, X_r^x, Y_r^x) dr, \\ Y_s^x &= g'(X_T^x) - \int_s^T Z_r^x dW_r + \int_s^T f_x(r, X_r^x, f_y^*(r, X_r^x, Y_r^x)) dr \end{aligned} \quad (6)$$

for all $s \in [0, T]$. To simplify the notations, when there is no ambiguity, (X^x, Y^x, Z^x) will be denoted by (X, Y, Z) . A solution to (6) is a triplet $(X, Y, Z) = (X^x, Y^x, Z^x)$ of progressively measurable processes such that

- X and Y are continuous and non-negative processes,
- the processes X, Y and $s \mapsto f_y^*(s, X_s, Y_s)$ are bounded and, finally,
- the two equations (6) are satisfied a.s. for every fixed $s \in [t, T]$.

Note that under the above and due to (C3) the processes $s \mapsto f(s, X_s^x, f_y^*(s, X_s^x, Y_s^x))$, $s \mapsto f_x(s, X_s^x, f_y^*(s, X_s^x, Y_s^x))$, as well as the random variables $g(X_T^x)$ and $g'(X_T^x)$ are also bounded. This implies that the stochastic integral $\int_0^T Z_r^x dW_r$ is a BMO martingale (see e.g. Proposition 1.1 in [5]). In particular for any $p \geq 1$ it holds that

$$\mathbb{E} \left[\left(\int_0^T |Z_r^x|^2 dr \right)^{p/2} \right] < +\infty. \quad (7)$$

Remark 1.5. We require that any solution satisfies $X, Y \geq 0$ to make sure that the equation (6) is well defined. Notice that we only assume that g is differentiable on $[0, \infty)$ and that $f_y^*(t, x, y)$, $f_x(t, x, f_y^*(t, x, y))$ are defined for $x, y \geq 0$.

Constructing solutions to the above FBSDE is important for the following reason:

Proposition 1.6. *If there exists a solution (X, Y, Z) of (6), then an optimal control for problem (3) is given by*

$$\alpha_s = f_y^*(s, X_s, Y_s), \quad s \in [0, T].$$

Proof. We adapt the proof of Theorem 5.2 in [23] to our setting. For $x = X_0 \geq 0$ and $\bar{\alpha} \in \mathcal{A}(x)$ let $\bar{X} = X^{x, \bar{\alpha}}$ be the associated state process. For the purpose of showing that $J(x, \bar{\alpha}) \geq J(x, \alpha)$ we can assume by Proposition 1.2 that \bar{X} is non-increasing and non-negative, such that it is in fact bounded. Let us define $\delta X := \bar{X} - X$ and $\delta \alpha := \bar{\alpha} - \alpha$. Note that $\bar{X}_0 = X_0 = x$, such that $\delta X_0 = 0$. Also, observe that δX is a bounded process. Since g is convex we have a.s.

$$Y_T \delta X_T - Y_0 \delta X_0 = Y_T \delta X_T = g'(X_T) \delta X_T \leq g(\bar{X}_T) - g(X_T). \quad (8)$$

At the same time Itô's formula proves that

$$\begin{aligned} Y_T \delta X_T - Y_0 \delta X_0 &= \int_0^T \delta X_s dY_s + \int_0^T Y_s d(\delta X_s) \\ &= - \int_0^T (\delta X_s) f_x(s, X_s, \alpha_s) ds + \int_0^T (\delta X_s) Z_s dW_s - \int_0^T Y_s (\delta \alpha_s) ds, \end{aligned} \quad (9)$$

where all expressions are well-defined due to the boundedness of X , Y and $\alpha = f_y^*(\cdot, X, Y)$. Now note that

$$\mathcal{H}_a(s, X_s, \alpha_s, Y_s) = -Y_s + f_a(s, X_s, f_y^*(s, X_s, Y_s)) = -Y_s + Y_s = 0.$$

Together with the convexity of \mathcal{H} this implies

$$\begin{aligned} \mathcal{H}(s, \bar{X}_s, \bar{\alpha}_s, Y_s) - \mathcal{H}(s, X_s, \alpha_s, Y_s) &\geq \mathcal{H}_x(s, X_s, \alpha_s, Y_s)(\delta X_s) + \mathcal{H}_a(s, X_s, \alpha_s, Y_s)(\delta \alpha_s) \\ &= f_x(s, X_s, \alpha_s)(\delta X_s). \end{aligned}$$

Thus, due to the definition of \mathcal{H}

$$\begin{aligned} f(s, \bar{X}_s, \bar{\alpha}_s) - f(s, X_s, \alpha_s) &= \mathcal{H}(s, \bar{X}_s, \bar{\alpha}_s, Y_s) - \mathcal{H}(s, X_s, \alpha_s, Y_s) + Y_s (\delta \alpha_s) \\ &\geq f_x(s, X_s, \alpha_s)(\delta X_s) + Y_s (\delta \alpha_s). \end{aligned}$$

This implies together with (9):

$$Y_T \delta X_T - Y_0 \delta X_0 \geq \int_0^T (f(s, X_s, \alpha_s) - f(s, \bar{X}_s, \bar{\alpha}_s)) ds + \int_0^T (\delta X_s) Z_s dW_s, \quad (10)$$

where $f(s, X_s, \alpha_s) - f(s, \bar{X}_s, \bar{\alpha}_s)$, although possibly not integrable w.r.t. s , is in any case bounded from above and where $(\delta X)Z$ is a BMO process, such that both integrals are well defined. Combining (10) with (8) we obtain

$$g(\bar{X}_T) - g(X_T) \geq \int_0^T (f(s, X_s, \alpha_s) - f(s, \bar{X}_s, \bar{\alpha}_s)) ds + \int_0^T (\delta X_s) Z_s dW_s$$

a.s. or

$$g(\bar{X}_T) + \int_0^T f(s, \bar{X}_s, \bar{\alpha}_s) ds \geq g(X_T) + \int_0^T f(s, X_s, \alpha_s) ds + \int_0^T (\delta X_s) Z_s dW_s$$

a.s., which is true even if $\left(\int_0^T f(s, \bar{X}_s, \bar{\alpha}_s) ds\right)(\omega) = \infty$ for some ω as the right-hand-side of the inequality is a finite number a.s. Taking the expectation on both sides leads to

$$J(x, \bar{\alpha}) \geq J(x, \alpha),$$

which shows optimality of α . □

Remark 1.7. Proposition 1.6 can be extended directly to the dynamic version of the problem (3), namely

$$\text{Minimize } J(t, x, \alpha) := \mathbb{E} \left[\int_t^T f(s, X_s^{t,x,\alpha}, \alpha_s) ds + g(X_T^{t,x,\alpha}) \middle| \mathcal{F}_t \right] \text{ over all } \alpha \in \mathcal{A}(t, x),$$

where $(t, x) \in [0, T] \times \mathbb{R}$ and $\mathcal{A}(t, x)$ is the set of all progressively measurable $\alpha : \Omega \times [t, T] \rightarrow A$ such that

$$X_s^{t,x,\alpha} := x - \int_t^s \alpha_r dr$$

is well-defined for all $s \in [t, T]$ and such that $\mathbb{E} \left[\int_t^T f \left(s, X_s^{t,x,\alpha}, \alpha_s \right) ds + g \left(X_T^{t,x,\alpha} \right) \right] < \infty$. In this case the FBSDE (6) becomes:

$$\begin{aligned} X_s^{t,x} &= x - \int_t^s f_y^*(r, X_r^{t,x}, Y_r^{t,x}) dr, \\ Y_s^{t,x} &= g'(X_T^{t,x}) - \int_s^T Z_r^{t,x} dW_r + \int_s^T f_x(r, X_r^{t,x}, f_y^*(r, X_r^{t,x}, Y_r^{t,x})) dr, \end{aligned} \quad (11)$$

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

Problem (P)

For all $x \in \mathbb{R}$ let $\mathcal{A}^0(x) := \{\alpha \in \mathcal{A}(x) \mid X_T^{x,\alpha} = 0 \text{ a.s.}\}$. The problem (P) reads as follows:

$$\text{Minimize } \hat{J}(x, \alpha) := \mathbb{E} \left[\int_0^T f(s, X_s^{t,x,\alpha}, \alpha_s) ds \right] \text{ over all } \alpha \in \mathcal{A}^0(x). \quad (12)$$

In order to reduce problem (12) to (3), we choose for every $L > 0$ the penalty function $g^L(x) := L \cdot x^2$ and define

$$J^L(x, \alpha) := \mathbb{E} \left[\int_0^T f(s, X_s, \alpha_s) ds + g^L(X_T) \right] \quad (13)$$

accordingly. We show, under some additional analytic conditions, that the FBSDE associated to the problem of minimizing J^L has a solution (X^L, Y^L, Z^L) . In Section 4 we show that the optimal controls $\alpha^L = f_y^*(t, X_t^L, Y_t^L)$ converge, as $L \rightarrow \infty$, to a control α^∞ in $\mathcal{A}^0(x)$ solving problem (12).

We remark that problem (12) can not be straightforwardly reduced to an FBSDE, since the terminal condition would not be well defined in a conventional sense. We show, however, that (P) can be linked to a pair of coupled SDEs, where the first equation describes again the optimal position process. The position constraints translate into an initial and a terminal condition on the first equation. On the second equation no conditions are imposed (see Theorem 4.7).

2 The method of decoupling fields

As mentioned above, solving (6) is crucial in constructing optimal controls. As a key result of this paper we prove in Section 3.1 the solvability of (6) under the assumption that the functions f_y^* , g' and f_x are uniformly Lipschitz continuous in $(x, y) \in [0, \infty)^2$ and are zero if we plug in the special value $(x, y) = (0, 0)$.

Note that even under these Lipschitz assumptions, it is not trivial to show well-posedness of (6) due to its coupled nature. It is necessary to take more subtle structural properties into account to conduct the proof. Our argumentation will be based on the so-called method of decoupling fields which we will briefly sum up in this section. Owing to their general significance, we treat the theory of FBSDEs and their decoupling fields in a more general framework than might be needed for investigating well-posedness of (6).

For a fixed finite time horizon $T > 0$, we consider a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$, where \mathcal{F}_0 consists of all null sets, $(W_t)_{t \in [0, T]}$ is a d -dimensional Brownian motion and $\mathcal{F}_t := \sigma(\mathcal{F}_0, (W_s)_{s \in [0, t]})$ with $\mathcal{F} := \mathcal{F}_T$. The dynamics of an FBSDE is given by

$$\begin{aligned} X_s &= X_0 + \int_0^s \mu(r, X_r, Y_r, Z_r) dr + \int_0^s \sigma(r, X_r, Y_r, Z_r) dW_r, \\ Y_t &= \xi(X_T) - \int_t^T f(r, X_r, Y_r, Z_r) dr - \int_t^T Z_r dW_r, \end{aligned}$$

for $s, t \in [0, T]$ and $X_0 \in \mathbb{R}^n$, where $(\xi, (\mu, \sigma, f))$ are measurable functions such that

$$\begin{aligned} \xi &: \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^m, & \mu &: [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}^n, \\ \sigma &: [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}^{n \times d}, & f &: [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m \times d} \rightarrow \mathbb{R}^m, \end{aligned}$$

for $d, n, m \in \mathbb{N}$. Throughout the whole section μ, σ and f are assumed to be progressively measurable with respect to $(\mathcal{F}_t)_{t \in [0, T]}$.

A decoupling field comes with an even richer structure than just a classical solution (X, Y, Z) .

Definition 2.1. Let $t \in [0, T]$. A function $u: [t, T] \times \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $u(T, \cdot) = \xi$ a.e. is called *decoupling field* for $(\xi, (\mu, \sigma, f))$ on $[t, T]$ if for all $t_1, t_2 \in [t, T]$ with $t_1 \leq t_2$ and any \mathcal{F}_{t_1} -measurable $X_{t_1}: \Omega \rightarrow \mathbb{R}^n$ there exist progressively measurable processes (X, Y, Z) on $[t_1, t_2]$ such that

$$\begin{aligned} X_s &= X_{t_1} + \int_{t_1}^s \mu(r, X_r, Y_r, Z_r) dr + \int_{t_1}^s \sigma(r, X_r, Y_r, Z_r) dW_r, \\ Y_s &= Y_{t_2} - \int_s^{t_2} f(r, X_r, Y_r, Z_r) dr - \int_s^{t_2} Z_r dW_r, \\ Y_s &= u(s, X_s), \end{aligned} \tag{14}$$

for all $s \in [t_1, t_2]$. In particular, we want all integrals to be well-defined.

Some remarks about this definition are in place.

- The first equation in (14) is called the *forward equation*, the second the *backward equation* and the third will be referred to as the *decoupling condition*.
- Note that, if $t_2 = T$, we get $Y_T = \xi(X_T)$ a.s. as a consequence of the decoupling condition together with $u(T, \cdot) = \xi$. At the same time $Y_T = \xi(X_T)$ together with decoupling condition implies $u(T, \cdot) = \xi$ a.e.
- If $t_2 = T$ we can say that a triplet (X, Y, Z) solves the FBSDE, meaning that it satisfies the forward and the backward equation, together with $Y_T = \xi(X_T)$. This relationship $Y_T = \xi(X_T)$ is referred to as the *terminal condition*.

In contrast to classical solutions of FBSDEs, decoupling fields on different intervals can be pasted together:

Lemma 2.2 ([9], Lemma 2.1.2). *Let u be a decoupling field for $(\xi, (\mu, \sigma, f))$ on $[t, T]$ and \tilde{u} be a decoupling field for $(u(t, \cdot), (\mu, \sigma, f))$ on $[s, t]$, for $0 \leq s < t < T$. Then, the map \hat{u} given by $\hat{u} := \tilde{u}\mathbf{1}_{[s, t]} + u\mathbf{1}_{(t, T]}$ is a decoupling field for $(\xi, (\mu, \sigma, f))$ on $[s, T]$.*

We want to remark that, if u is a decoupling field and \tilde{u} is a modification of u , i.e. for each $s \in [t, T]$ the functions $u(s, \omega, \cdot)$ and $\tilde{u}(s, \omega, \cdot)$ coincide for almost all $\omega \in \Omega$, then \tilde{u} is also a decoupling field to the same problem. Hence, u could also be referred to as a class of modifications and a progressively measurable and in some sense right-continuous representative exists if the decoupling field is Lipschitz continuous in x (Lemma 2.1.3 in [9]).

For the following we need to fix briefly further notation.

Let $I \subseteq [0, T]$ be an interval and $u : I \times \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ a map such that $u(s, \cdot)$ is measurable for every $s \in I$. We define

$$L_{u,x} := \sup_{s \in I} \inf \{L \geq 0 \mid \text{for a.a. } \omega \in \Omega : |u(s, \omega, x) - u(s, \omega, x')| \leq L|x - x'| \text{ for all } x, x' \in \mathbb{R}^n\},$$

where $\inf \emptyset := \infty$. We also set $L_{u,x} := \infty$ if $u(s, \cdot)$ is not measurable for every $s \in I$. One can show that $L_{u,x} < \infty$ is equivalent to u having a modification which is truly Lipschitz continuous in $x \in \mathbb{R}^n$.

We denote by $L_{\sigma,z}$ the Lipschitz constant of σ w.r.t. the dependence on the last component z and w.r.t. the Frobenius norms on $\mathbb{R}^{m \times d}$ and $\mathbb{R}^{n \times d}$. We set $L_{\sigma,z} = \infty$ if σ is not Lipschitz continuous in z .

By $L_{\sigma,z}^{-1} = \frac{1}{L_{\sigma,z}}$ we mean $\frac{1}{L_{\sigma,z}}$ if $L_{\sigma,z} > 0$ and ∞ otherwise.

For an integrable real valued random variable F the expression $\mathbb{E}_t[F]$ refers to $\mathbb{E}[F|\mathcal{F}_t]$, while $\mathbb{E}_{t,\infty}[F]$ refers to $\text{ess sup } \mathbb{E}[F|\mathcal{F}_t]$, which might be ∞ , but is always well defined as the infimum of all constants $c \in [-\infty, \infty]$ such that $\mathbb{E}[F|\mathcal{F}_t] \leq c$ a.s. Additionally, we write $\|F\|_\infty$ for the essential supremum of $|F|$.

Finally for a matrix $A \in \mathbb{R}^{N \times n}$ and a vector $v \in S^{n-1}$ we define $|A|_v := |Av|$ as the norm of A in the direction v .

In practice it is important to have explicit knowledge about the regularity of (X, Y, Z) . For instance, it is important to know in which spaces the processes live, and how they react to changes in the initial value.

Definition 2.3. Let $u : [t, T] \times \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a decoupling field to $(\xi, (\mu, \sigma, f))$.

1. We say u to be *weakly regular* if $L_{u,x} < L_{\sigma,z}^{-1}$ and $\sup_{s \in [t, T]} \|u(s, \cdot, 0)\|_\infty < \infty$.
2. A weakly regular decoupling field u is called *strongly regular* if for all fixed $t_1, t_2 \in [t, T]$, $t_1 \leq t_2$, the processes (X, Y, Z) arising in (14) are a.e unique and satisfy

$$\sup_{s \in [t_1, t_2]} \mathbb{E}_{t_1, \infty}[|X_s|^2] + \sup_{s \in [t_1, t_2]} \mathbb{E}_{t_1, \infty}[|Y_s|^2] + \mathbb{E}_{t_1, \infty} \left[\int_{t_1}^{t_2} |Z_s|^2 ds \right] < \infty, \quad (15)$$

for each constant initial value $X_{t_1} = x \in \mathbb{R}^n$. In addition they are required to be measurable as functions of (x, s, ω) and even weakly differentiable w.r.t. $x \in \mathbb{R}^n$ such that for every $s \in [t_1, t_2]$ the mappings X_s and Y_s are measurable functions of (x, ω) and even weakly differentiable w.r.t. x such that

$$\begin{aligned} \text{ess sup}_{x \in \mathbb{R}^n} \sup_{v \in S^{n-1}} \sup_{s \in [t_1, t_2]} \mathbb{E}_{t_1, \infty} \left[\left| \frac{d}{dx} X_s \Big|_v \right|^2 \right] &< \infty, \\ \text{ess sup}_{x \in \mathbb{R}^n} \sup_{v \in S^{n-1}} \sup_{s \in [t_1, t_2]} \mathbb{E}_{t_1, \infty} \left[\left| \frac{d}{dx} Y_s \Big|_v \right|^2 \right] &< \infty, \\ \text{ess sup}_{x \in \mathbb{R}^n} \sup_{v \in S^{n-1}} \mathbb{E}_{t_1, \infty} \left[\int_{t_1}^{t_2} \left| \frac{d}{dx} Z_s \Big|_v \right|^2 ds \right] &< \infty. \end{aligned} \quad (16)$$

3. We say that a decoupling field on $[t, T]$ is *strongly regular* on a subinterval $[t_1, t_2] \subseteq [t, T]$ if u restricted to $[t_1, t_2]$ is a strongly regular decoupling field for $(u(t_2, \cdot), (\mu, \sigma, f))$.

Under suitable conditions a rich existence, uniqueness and regularity theory for decoupling fields can be developed. The basis of the theory is Theorem 2.4 below, which is proven in Chapter 2 of [9].

Assumption (SLC): $(\xi, (\mu, \sigma, f))$ satisfies *standard Lipschitz conditions* (SLC) if

1. (μ, σ, f) are Lipschitz continuous in (x, y, z) with Lipschitz constant L ,
2. $\|(|\mu| + |f| + |\sigma|)(\cdot, \cdot, 0, 0, 0)\|_\infty < \infty$,
3. $\xi: \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ is measurable such that $\|\xi(\cdot, 0)\|_\infty < \infty$ and $L_{\xi, x} < L_{\sigma, z}^{-1}$.

Theorem 2.4 ([9], Theorem 2.2.1). *Suppose $(\xi, (\mu, \sigma, f))$ satisfies (SLC). Then there exists a time $t \in [0, T)$ such that $(\xi, (\mu, \sigma, f))$ has a unique (up to modification) decoupling field u on $[t, T]$ with $L_{u, x} < L_{\sigma, z}^{-1}$ and $\sup_{s \in [t, T]} \|u(s, \cdot, 0)\|_\infty < \infty$.*

A brief discussion of existence and uniqueness of classical solutions on sufficiently small intervals can be found in Remark 2.2.4 in [9].

This local theory for decoupling fields can be systematically extended to global results based on fairly simple “small interval induction” arguments (Lemma 2.5.1 and 2.5.2 in [9]).

Theorem 2.5 ([9], Corollary 2.5.3, 2.5.4 and 2.5.5). *Suppose that $(\xi, (\mu, \sigma, f))$ satisfies (SLC).*

1. *Global uniqueness: If there are two weakly regular decoupling fields $u^{(1)}, u^{(2)}$ to the corresponding problem on some interval $[t, T]$, then we have $u^{(1)} = u^{(2)}$ up to modifications.*
2. *Global regularity: If there exists a weakly regular decoupling field u to this problem on some interval $[t, T]$, then u is strongly regular.*
3. *If there exists a weakly regular decoupling field u of the corresponding FBSDE on some interval $[t, T]$, then for any initial condition $X_t = x \in \mathbb{R}^n$ there is a unique solution (X, Y, Z) of the FBSDE on $[t, T]$ satisfying*

$$\sup_{s \in [t, T]} \mathbb{E}[|X_s|^2] + \sup_{s \in [t, T]} \mathbb{E}[|Y_s|^2] + \mathbb{E} \left[\int_t^T |Z_s|^2 ds \right] < \infty.$$

In order to have a notion of global existence we need the following definition:

Definition 2.6. We define the maximal interval $I_{\max} \subseteq [0, T]$ of the problem given by $(\xi, (\mu, \sigma, f))$ as the union of all intervals $[t, T] \subseteq [0, T]$, such that there exists a weakly regular decoupling field u on $[t, T]$.

Note that the maximal interval might be open to the left. Also, let us remark that we define a decoupling field on such an interval as a mapping which is a decoupling field on every compact subinterval containing T . Similarly we can define weakly and strongly regular decoupling fields as mappings which restricted to an arbitrary compact subinterval containing T are weakly (or strongly) regular decoupling fields in the sense of the definitions given above.

Finally, we have global existence and uniqueness on the maximal interval:

Theorem 2.7 (Global existence in weak form, [9], Theorem 5.1.11 and Lemma 5.1.12). *Let $(\xi, (\mu, \sigma, f))$ satisfy SLC. Then there exists a unique weakly regular decoupling field u on I_{\max} . This u is even strongly regular. Furthermore, either $I_{\max} = [0, T]$ or $I_{\max} = (t_{\min}, T]$, where $0 \leq t_{\min} < T$. In the latter case we have*

$$\lim_{t \downarrow t_{\min}} L_{u(t, \cdot), x} = L_{\sigma, z}^{-1}. \quad (17)$$

Note that in particular cases the last statement allows to show “strong global existence”, i.e. $I_{\max} = [0, T]$, via contradiction and, thereby, is the basis of the so-called method of decoupling fields.

3 Solving problem (UP)

In this section we provide sufficient conditions for the FBSDE (6) to possess a unique solution on $[0, T]$. With Proposition 1.6 we then immediately obtain a solution of problem (UP). We first assume, in Subsection 3.1, that all the coefficients of the FBSDE are Lipschitz continuous. In Subsection 3.2 we consider the additive case and show solvability by reducing the FBSDE to the previously considered Lipschitz case via a cut-off argument.

3.1 The Lipschitz case

In addition to conditions (C0) to (C4) we assume throughout this subsection that

(D1) $f_y^*(t, x, y)$, $g'(x)$ and $f_x(t, x, f_y^*(t, x, y))$ are uniformly Lipschitz continuous in $(x, y) \in [0, \infty) \times [0, \infty)$,

(D2) $g'(0) = f_x(t, 0, 0) = 0$ for all ω, t ,

(D3) $\sup_{x \geq 0, a \in A_+} |f_{xx}(t, x, a)|$ is bounded uniformly in (ω, t) .

Notice that in general the functions $f_y^*(t, \cdot, \cdot)$, g' , $f_x(t, \cdot, f_y^*(t, \cdot, \cdot))$ are defined only on $[0, \infty) \times [0, \infty)$, i.e. for non-negative x, y . We extend them to the whole of \mathbb{R}^2 by projecting (x, y) to $(x \vee 0, y \vee 0)$ and plugging the projected value in the respective function. Note that condition (D1) is still maintained after this extension. By a slight abuse of notation we denote the new functions again by $f_y^*(t, \cdot, \cdot)$, g' , $f_x(t, \cdot, f_y^*(t, \cdot, \cdot))$. Note that (D1) implies that these three functions are weakly differentiable on the whole of \mathbb{R}^2 . Furthermore the weak derivatives w.r.t. x vanish whenever $x \leq 0$ and the weak derivatives w.r.t. y vanish whenever $y \leq 0$.

Now the parameter functions of the problem (6) satisfy (SLC), such that the theory of decoupling fields can be applied. We will see in the following result that the problem is well-posed and can be solved on the whole of $[0, T]$. Note, however, that in order to construct, for given $t \in [0, T)$ and $x = X_t \in [0, \infty)$, an optimal control via $\alpha_s := f_y^*(s, X_s, Y_s)$, $s \in [t, T]$, we also need to prove $X_s \in [0, \infty)$ and $Y_s \in [0, \infty)$ everywhere such that the initial domain of f_y^* is never left. This follows, for the most part, from the non-negativity of f_y^* , g' and f_x in our domains of interest. Only the property $X_s \geq 0$ is not clear a priori. This, however, follows from the dynamics of X , the decoupling condition $u(s, X_s) = Y_s$, the statement $f_y^*(t, 0, 0) = 0$ from Remark 1.3 and the property $u(t, x) = 0$, for all $x \leq 0$, which is proven in Theorem 3.1.

Theorem 3.1. *The maximal interval associated with the above problem satisfies $I_{\max} = [0, T]$. Furthermore, the unique weakly regular decoupling field u on $[0, T]$ satisfies $u(t, x) = 0$ for all $x \leq 0$ and $t \in [0, T]$.*

Proof. Suppose $I_{\max} = [0, T]$ does not hold. Then $I_{\max} = (t_{\min}, T]$, where $t_{\min} \in [0, T)$. We show that this leads to a contradiction. Let u be the unique weakly regular decoupling field on $(t_{\min}, T]$. We next show that there exists a uniform bound on u_x . In the following let $t_0 \in (t_{\min}, T]$ and $x \in \mathbb{R}$ be arbitrary. Let (X, Y, Z) be the solution of the corresponding FBSDE on $[t_0, T]$ with initial condition $X_{t_0} = x$. We define $D_t := \frac{dX_t}{dx}$, $C_t = \frac{dY_t}{dx}$ as well as $\Psi_t = u_x(t, X_t)$. These objects are well defined due to strong regularity of u .

Firstly, let us consider the case $x \leq 0$. It is straightforward to verify that X, Y are both constant: $X = x$ and $Y = 0$. This is due to uniqueness and the fact that these constant processes together with $Z = 0$ solve the FBSDE as g' vanishes for non-positive values. This means that $u(t_0, x) = 0$ for $x \leq 0$, which implies $u_x(t_0, x) = 0$ for $x \leq 0$.

Secondly, let us assume $x > 0$. We claim that $X_t, Y_t \geq 0$ holds for all $t \in [t_0, T]$: Note that X is non-increasing due to non-negativity of f_y^* . Define the stopping time $\tau := \inf\{s \geq t_0 \mid X_s = 0\} \wedge T$ and a process \tilde{X} in such a way that it vanishes on $(\tau, T]$, but is equal to X on $[t_0, \tau]$. Similarly we define $\tilde{Y}_s := Y_s \mathbf{1}_{\{s \leq \tau\}}$. Now note that if $\tau < T$, then $X_\tau = 0$ and $X_s \leq 0$ on $\{s > \tau\}$ for all $s \in [t_0, T]$ and, therefore, $Y_s \mathbf{1}_{\{s > \tau\}} = u(s, X_s) \mathbf{1}_{\{s > \tau\}} = 0$ such that $Y_\tau \mathbf{1}_{\{\tau < T\}} = 0$, due to continuity of Y . This implies that \tilde{Y} is continuous and an Itô process as is \tilde{X} . It is now straightforward to check that \tilde{X}, \tilde{Y} solve the same FBSDE, which is solved by X, Y implying that \tilde{X}, \tilde{Y} coincide with X, Y according to Theorem 2.5. Furthermore, we have shown that for all $s \in [t_0, T]$: $\{X_s = 0\} \subseteq \{Y_s = 0\}$ up to a null set.

Now notice that D satisfies

$$dD_t = - (f_{yx}^*(t, X_t, Y_t)D_t + f_{yy}^*(t, X_t, Y_t)C_t) dt, \quad (18)$$

and $D_{t_0} = 1$. We now show that D remains positive on the whole interval $[t_0, T]$. To this end let $\tau \in [t_0, T]$ be a stopping time such that D is positive on $[t_0, \tau]$. Then we have $\Psi_t = C_t D_t^{-1}$ on $[t_0, \tau]$. Note that

$$\begin{aligned} dD_t^{-1} &= \frac{f_{yx}^*(t, X_t, Y_t)D_t + f_{yy}^*(t, X_t, Y_t)C_t}{D_t^2} dt \\ &= D_t^{-1} (f_{yx}^*(t, X_t, Y_t) + f_{yy}^*(t, X_t, Y_t)\Psi_t) dt, \end{aligned}$$

on the stochastic interval $[t_0, \tau]$. Consequently,

$$D_t^{-1} = \exp \left(\int_{t_0}^t f_{yx}^*(s, X_s, Y_s) + f_{yy}^*(s, X_s, Y_s)\Psi_s ds \right).$$

Since the decoupling field is weakly regular, u_x and hence Ψ_t are bounded on $[t_0, T]$. Moreover, f_{yx}^*, f_{yy}^* are uniformly bounded. Therefore, D^{-1} is bounded on $[t_0, \tau]$ by a constant that does not depend on τ . This implies that D can never reach 0 and, therefore, we can choose $\tau = T$.

Moreover, C satisfies the BSDE

$$\begin{aligned} -dC_t &= -\frac{dZ_t}{dx} dW_t + f_{xx}(t, X_t, \alpha_t)D_t dt \\ &\quad + f_{xa}(t, X_t, \alpha_t) [f_{yx}^*(t, X_t, Y_t)D_t + f_{yy}^*(t, X_t, Y_t)C_t] dt, \end{aligned} \quad (19)$$

with terminal condition $C_T = g''(X_T)D_T$, where $\alpha_t := f_y^*(t, X_t, Y_t)$.

The dynamics of Ψ are now deduced from those of C and D using the product rule:

$$\begin{aligned}
d\Psi_t &= D_t^{-1}dC_t + C_t dD_t^{-1} \\
&= D_t^{-1} \frac{dZ_t}{dx} dW_t + f_{yy}^*(t, X_t, Y_t) \Psi_t^2 dt \\
&\quad + [f_{yx}^*(t, X_t, Y_t) - f_{xa}(t, X_t, \alpha_t) f_{yy}^*(t, X_t, Y_t)] \Psi_t dt \\
&\quad - [f_{xx}(t, X_t, \alpha_t) + f_{xa}(t, X_t, \alpha_t) f_{yx}^*(t, X_t, Y_t)] dt \\
&= D_t^{-1} \frac{dZ_t}{dx} dW_t - f_{xx}(t, X_t, \alpha_t) dt \\
&\quad + [f_{yy}^*(t, X_t, Y_t) \Psi_t + f_{yx}^*(t, X_t, Y_t)] [\Psi_t - f_{xa}(t, X_t, \alpha_t)] dt. \tag{20}
\end{aligned}$$

Considering (5) in Remark 1.4 we have:

$$\begin{aligned}
f_{yy}^*(t, x, y) \mathbf{1}_{(0, \infty)}(x) &= \frac{1}{f_{aa}}(t, x, f_y^*(t, x, y)) \mathbf{1}_{(f_a(t, x, 0+), f_a(t, x, a_{\text{sup}}))}(y) \mathbf{1}_{(0, \infty)}(x), \\
f_{yx}^*(t, x, y) \mathbf{1}_{(0, \infty)}(x) &= -\frac{f_{ax}}{f_{aa}}(t, x, f_y^*(t, x, y)) \mathbf{1}_{(f_a(t, x, 0+), f_a(t, x, a_{\text{sup}}))}(y) \mathbf{1}_{(0, \infty)}(x).
\end{aligned}$$

Together with (20) we obtain:

$$d\Psi_t = D_t^{-1} \frac{dZ_t}{dx} dW_t - f_{xx}(t, X_t, \alpha_t) dt + f_{yy}^*(t, X_t, Y_t) (\Psi_t - f_{xa}(t, X_t, \alpha_t))^2 dt. \tag{21}$$

To sum up, Ψ_t solves the BSDE with $\Psi_T = g''(X_T)$ and driver

$$h(t, \psi) = -\frac{(\psi - f_{xa}(t, X_t, \alpha_t))^2}{f_{aa}(t, X_t, \alpha_t)} \mathbf{1}_{(f_a(t, X_t, 0), f_a(t, X_t, a_{\text{sup}}))}(Y_t) + f_{xx}(t, X_t, \alpha_t).$$

Lemma 3.2. *We have $\Psi \geq 0$ and $\Psi \leq g''(X_T) + \int_0^T f_{xx}(t, X_t, f_y^*(t, X_t, Y_t)) dt$. In particular, if $f_{xx}(t, X_t, f_y^*(t, X_t, Y_t))$ and $g''(X_T)$ are bounded, say by $K \in \mathbb{R}_+$, then Ψ is bounded from above by $K(1 + T)$.*

Proof. We first show $\Psi \geq 0$ using the comparison theorem. To this end define

$$\hat{h}(t, \psi) := 0 \wedge \left(\left(-\frac{(\psi - f_{xa}(t, X_t, \alpha_t))^2}{f_{aa}(t, X_t, \alpha_t)} + f_{xx}(t, X_t, \alpha_t) \right) \mathbf{1}_{(f_a(t, X_t, 0), f_a(t, X_t, a_{\text{sup}}))}(Y_t) \right).$$

Clearly, $\hat{h}(t, \psi) \leq h(t, \psi)$. We claim that the zero process solves the BSDE given by the driver \hat{h} and the terminal condition 0: Indeed, plugging $\psi = 0$ into the definition of \hat{h} , we notice that

$$\left(-\frac{f_{xa}^2}{f_{aa}} + f_{xx} \right) (t, X_t, \alpha_t) = \left(\frac{1}{f_{aa}} \det(D^2 f) \right) (t, X_t, \alpha_t),$$

where $D^2 f$ is the Hessian matrix of f with respect to the two variables x and a . $D^2 f$ is symmetrical and positive definite everywhere, since f is strictly convex. Therefore, the eigenvalues of $D^2 f$ are always positive. Moreover, we can calculate $\det(D^2 f)$ as the product of those eigenvalues. Therefore, $\hat{h}(t, 0) = 0$, showing that 0 is a solution of the BSDE with parameters $(\hat{h}, 0)$. Finally, the comparison theorem implies that $\Psi \geq 0$.

In order to show the upper estimate notice that $h(t, \psi) \leq f_{xx}(t, X_t, f_y^*(t, X_t, Y_t))$. With the comparison theorem we get $\Psi \leq g''(X_T) + \int_0^T f_{xx}(t, X_t, f_y^*(t, X_t, Y_t)) dt$. \square

According to our assumptions f_{xx} and g'' are bounded. Therefore, Ψ_t is bounded from above, say by $C \in \mathbb{R}_+$, and also non-negative. Notice that we can choose C such that it does not depend on x and t_0 . We have $|u_x(t_0, x)| = |u_x(t_0, X_{t_0})| = |\Psi_{t_0}| \leq C$, which is a contradiction to (17) in Theorem 2.7.

Therefore, there exists a unique weakly regular decoupling field u on $[0, T]$ to our FBSDE which implies the existence of classical solutions according to Theorem 2.5 as well.

The property $u(t, x) = 0$, $x \leq 0$, is a simple consequence of the fact that $(X, Y, Z) = (0, 0, 0)$ solve the FBSDE with the initial condition $X_t = x \leq 0$. \square

Corollary 3.3. *For $x = X_0 > 0$ let (X, Y, Z) satisfy the FBSDE together with the decoupling condition $u(s, X_s) = Y_s$. Then X, Y are both bounded and hence (X, Y, Z) is a solution to (6) in the sense of Section 1.*

Proof. We have seen $X, Y \geq 0$ in the proof of Theorem 3.1. Furthermore, X is decreasing and, therefore, $X \leq x$. Boundedness of X together with the decoupling condition and Lipschitz continuity of u in x as well as $u(\cdot, 0) = 0$ imply that Y must also be bounded. Finally, Lipschitz continuity of f_y^* together with $f_y^*(\cdot, 0, 0) = 0$ implies boundedness of $s \mapsto f_y^*(s, X_s, Y_s)$. \square

As another consequence of the Lipschitz continuity of u in the last component we obtain

Lemma 3.4. *Let (X, Y, Z) be the solution of FBSDE (11) on $[t, T]$ with initial condition $X_t = x > 0$. Then $X_s > 0$ for all $s \in [t, T]$.*

Proof. The forward equation together with the decoupling condition imply

$$X_s = x - \int_t^s f_y^*(r, X_r, u(r, X_r)) dr,$$

for all $s \in [t, T]$. We have $f_y^*(r, 0, 0) = u(r, 0) = 0$ and, therefore,

$$|f_y^*(r, X_r, u(r, X_r))| \leq L_{f_y^*}(|X_r| + L_u \cdot |X_r|) \leq L_{f_y^*}(1 + L_u)|X_r|,$$

where $L_{f_y^*}$ is a Lipschitz constant of f_y^* w.r.t. the last two components (x, y) and L_u is a Lipschitz constant of u w.r.t. x . Let $\theta : [t, T] \rightarrow \mathbb{R}$ be the unique solution to the ODE

$$\theta_s = x - \int_t^s L_{f_y^*}(1 + L_u)|\theta_r| dr, \quad s \in [t, T].$$

It is straightforward to verify that $\theta_s = x \cdot \exp\left(-L_{f_y^*}(1 + L_u)(s - t)\right)$, $s \in [t, T]$, is the solution. A comparison principle for ODEs implies that $X \geq \theta > 0$. \square

3.2 The case of an additive generator

In this subsection we assume that $f_{xa} = 0$. This is actually equivalent to assuming that f has the form $f(t, x, a) = f(t, x, 0) + f(t, 0, a) - f(t, 0, 0)$. In this case, in order to have solvability of (6), it is sufficient to require boundedness of $\frac{1}{f_{aa}}$ on $\Omega \times [0, T] \times [0, x_0] \times (A \cap (0, a_0))$ for all $a_0 > 0$ and $x_0 \geq 0$, together with $g'(0) = f_x(t, 0, 0) = 0$, for all $t \in [0, T]$.

For the following result we define the set $A_{\geq 0} := A \cap [0, \infty)$ and assume, as usual, that f, g satisfy conditions (C0) to (C4). Note that under $f_{xa} = 0$ the derivative f_x does not depend on a and f_a does not depend on x , such that we sometimes write $f_x(t, x)$ instead of $f_x(t, x, a)$ or $f_a(t, a)$ instead of $f_a(t, x, a)$ and use similar notations for the second derivatives as well. Also, we may write $f_y^*(t, y)$ instead of $f_y^*(t, x, y)$ for the same reason.

Lemma 3.5. *Assume that f_{xa} , $g'(0)$ and $f_x(\cdot, 0, 0)$ all vanish, let $x_0 > 0$ and $a_0 \in A_+$ be such that $\frac{1}{f_{aa}}$ is bounded on $\Omega \times [0, T] \times (0, a_0]$. Then there exists a progressively measurable $\tilde{f} : \Omega \times [0, T] \times \mathbb{R} \times A_{\geq 0} \rightarrow \mathbb{R}$, as well as a measurable $\tilde{g} : \Omega \times \mathbb{R} \rightarrow \mathbb{R}$ with the following properties:*

1. \tilde{f} and \tilde{g} satisfy (C0) - (C4) with $A_{\geq 0}$ as the set of admissible controls (instead of A).
2. (\tilde{f}, \tilde{f}_x) coincides with (f, f_x) on $\Omega \times [0, T] \times [0, x_0] \times [0, a_0]$.
3. (\tilde{g}, \tilde{g}') coincides with (g, g') on $\Omega \times [0, x_0]$.
4. \tilde{f}_y^* coincides with f_y^* on $\{(\omega, t, y) \in \Omega \times [0, T] \times [0, \infty) \mid y \leq f_a(t, a_0)\}$.
5. \tilde{f}_x, \tilde{g}' are uniformly Lipschitz continuous in $x \in [0, \infty)$.
6. \tilde{f}_y^* is uniformly Lipschitz continuous in $y \in [0, \infty)$.

Proof. For an arbitrary $r \in (0, \infty)$ we define an operator

$$G_r : C([0, r], \mathbb{R}) \cap C^2((0, r], \mathbb{R}) \rightarrow C(\mathbb{R}, \mathbb{R}) \cap C^2((0, \infty), \mathbb{R})$$

as follows: for any $\varphi \in C([0, r], \mathbb{R}) \cap C^2((0, r], \mathbb{R})$ let

$$G_r(\varphi)(x) := \varphi(r) + \varphi'(r)(x - r) + \frac{1}{2}\varphi''(r)(x - r)^2$$

for $x \geq r$, while defining $G_r(\varphi)(x) := \varphi(x)$ for $x \in [0, r)$ and setting $G_r(\varphi)(x) := G_r(\varphi)(-x)$ for $x \in (-\infty, 0)$. It is straightforward to verify that

- $G_r(\varphi)$ is well defined in the sense that it is continuous on \mathbb{R} and twice differentiable on $(0, \infty)$ with continuous derivatives,
- $G_r(\varphi)$ is strictly convex if φ is strictly convex with $\varphi''(r) > 0$,
- $G_r(\varphi)$ and φ coincide on $[0, r]$, while
- $G_r(\varphi)'(x) = \varphi'(r) + \varphi''(r)(x - r)$ and $G_r(\varphi)''(x) = \varphi''(r)$ for $x \geq r$.

Now define for all $(\omega, t, x, a) \in \Omega \times [0, T] \times \mathbb{R} \times A_{\geq 0}$

$$\tilde{f}(t, x, a) := G_{x_0}(f(t, \cdot, 0))(x) + G_{a_0}(f(t, 0, \cdot))(a) - f(t, 0, 0)$$

and for all $(\omega, x) \in \Omega \times \mathbb{R}$

$$\tilde{g}(x) := G_{x_0}(g)(x).$$

Note that \tilde{f} is additive similarly to f . We now check that \tilde{f}, \tilde{g} satisfy properties 1. to 6.:

1. \tilde{f} and \tilde{g} satisfy (C0) to (C4) with $A_{\geq 0}$ as the set of admissible controls:
 - \tilde{f} inherits convexity in (x, a) from f as the Hessian matrix is always positive-semidefinite. Similarly, \tilde{g} is convex in x . Furthermore, \tilde{f} is even strictly convex in a , since \tilde{f}_{aa} is bounded away from zero.
 - (C4) holds as \tilde{g} coincides with g for $x \in [0, x_0]$ and is increasing in $x \in [0, \infty)$ due to $g'(0) = 0$.

- The second statement in (C1), i.e. the claim that $x \mapsto \tilde{f}(t, x, 0)$ attains its minimum in 0, is verified analogously to (C4).
- The remaining part of (C1), i.e. the claim that $a \mapsto G_{a_0}(f(t, 0, \cdot))(a)$ attains its minimum in 0, follows from the analogous property of $a \mapsto f(t, 0, a)$.
Notice at this point that (C1) implies that $\tilde{f}_a(t, 0) \geq 0$ and $\tilde{f}(t, x, 0) \geq \tilde{f}(t, 0, 0) = 0$ for all $x \geq 0$.
- (C3) follows from basic properties of G_r combined with the property (C3) satisfied by f, g .
- It remains to verify (C2): Due to boundedness of $\frac{1}{f_{aa}}$ on $\Omega \times [0, T] \times [0, x_0] \times (0, a_0]$, the second derivative \tilde{f}_{aa} is uniformly bounded away from 0 by some constant $\kappa > 0$. This implies for $x \geq 0$ and $a \in A_+$:

$$\begin{aligned}
\tilde{f}(t, x, a) &= \tilde{f}(t, x, 0) + \int_0^a \left(\tilde{f}_a(t, x, 0) + \int_0^b \tilde{f}_{aa}(t, x, c) dc \right) db \\
&\geq \tilde{f}(t, x, 0) + \int_0^a \left(\tilde{f}_a(t, x, 0) + \int_0^b \kappa dc \right) db \\
&\geq \tilde{f}(t, x, 0) + \tilde{f}_a(t, x, 0) \cdot a + \frac{1}{2} a^2 \kappa \geq \frac{1}{2} a^2 \kappa.
\end{aligned}$$

This proves (C2) for \tilde{f} .

2. to 4.: Note that $g = \tilde{g}$ on $\Omega \times [0, x_0]$ and that $f = \tilde{f}$ on $\Omega \times [0, T] \times [0, x_0] \times [0, a_0]$. In particular, f_a and \tilde{f}_a coincide on $\Omega \times [0, T] \times (0, a_0]$. The property $f_a(t, a) = \tilde{f}_a(t, a)$ for $a \leq a_0$ implies $\tilde{f}_y^*(t, y) = f_y^*(t, y)$ for $y \in [0, f_a(t, a_0)]$ according to Remark 1.3.

5. \tilde{f}_x and \tilde{g}' are uniformly Lipschitz continuous in $x \in [0, \infty)$ according to our construction and the fact that the second derivatives of f, g are uniformly bounded on compacts.

6. Finally, we show that \tilde{f}_y^* is uniformly Lipschitz continuous in $(x, y) \in [0, \infty) \times [0, \infty)$:

Firstly, note that $\frac{1}{f_{aa}}$ is bounded on $\Omega \times [0, T] \times [0, \infty) \times A_+$ due to uniform boundedness of $\frac{1}{f_{aa}}$ on $\Omega \times [0, T] \times [0, x_0] \times (0, a_0]$. According to Remark 1.4 we have

$$\left(\tilde{f}_{yx}^*(t, x, y), \tilde{f}_{yy}^*(t, x, y) \right) = \begin{cases} (0, 0), & \text{if } y \notin \left(\tilde{f}_a(t, x, 0), \tilde{f}_a(t, x, a_{\text{sup}}) \right) \\ \left(0, \frac{1}{f_{aa}}(t, x, \tilde{f}_y^*(t, x, y)) \right), & \text{otherwise.} \end{cases}$$

This already implies that the weak derivative of \tilde{f}_y^* w.r.t. (x, y) is uniformly bounded. Consequently, \tilde{f}_y^* is uniformly Lipschitz continuous in y . \square

Theorem 3.6. *Assume that $f_{xa}, g'(0)$ and $f_x(\cdot, 0, 0)$ all vanish and assume that for every $a_0 > 0$ the function $\frac{1}{f_{aa}}$ is bounded on $\Omega \times [0, T] \times (0, a_0]$. Then there exists a solution (X, Y, Z) for the FBSDE (6) with $X_0 = x_0 > 0$.*

Proof. Let $c := \|g'(x_0)\|_{L^\infty(\Omega)} + T\|f_x(\cdot, x_0)\|_{L^\infty([0, T] \times \Omega)}$ and choose an $a_0 > 0$ such that $ba_0^{p-1} \geq c$ is satisfied, where the constants $b > 0$ and $p > 1$ stem from assumption (C2). Now let \tilde{f} and \tilde{g} be the functions constructed in Lemma 3.5. These functions satisfy the conditions required for Theorem 3.1: (D1), (D3) are satisfied due to 5. and 6., while (D2) follows from

2. and 3. Hence, there exists a process (X, Y, Z) satisfying (6) if we replace f, g by \tilde{f}, \tilde{g} . More precisely it holds, for every $s \in [0, T]$, a.s. that

$$\begin{aligned} X_s &= x_0 - \int_0^s \tilde{f}_y^*(r, Y_r) dr, \\ Y_s &= \tilde{g}'(X_T) - \int_s^T Z_r dW_r + \int_s^T \tilde{f}_x(r, X_r) dr, \end{aligned} \quad (22)$$

where we use that \tilde{f}_y^* does not depend on x and \tilde{f}_x does not depend on a .

Since $\tilde{f}_y^* \geq 0$ it follows that $X_s \leq x_0$ for all $s \in [0, T]$ and also that $\tilde{f}_x(s, X_s) = f_x(s, X_s)$ by Lemma 3.5. Moreover, convexity of \tilde{g} implies that $\tilde{g}'(X_T) \leq \tilde{g}'(x_0)$. Similarly, convexity of \tilde{f} in x implies that $\tilde{f}_x(s, X_s) \leq \tilde{f}_x(s, x_0)$. Thus, it holds that $Y_t \leq c$ for all $t \in [0, T]$. It follows from convexity of f in a and assumption (C2) that $f_a(t, 0, a_0) \geq \frac{f(t, 0, a_0)}{a_0} \geq b a_0^{p-1} \geq c$. By Lemma 3.5 it holds that $\tilde{f}_y^*(t, y) = f_y^*(t, y)$ for all $t \in [0, T]$ and $y \leq f_a(t, a_0)$. In particular, it follows that $\tilde{f}_y^*(t, Y_s) = f_y^*(t, Y_s)$ for all $s \in [0, T]$. Therefore, we can replace $\tilde{f}, \tilde{g}, \tilde{f}_y^*$ by f, g, f_y^* in (22). \square

Example 3.7. Let $p \in (1, 2]$, let $q, r \in [2, \infty)$, let $\eta, \gamma: [0, T] \times \Omega \rightarrow [0, \infty)$ be progressively measurable stochastic processes that are bounded away from zero and bounded from above uniformly in (t, ω) and let $\xi: \Omega \rightarrow [0, \infty)$ be a random variable that is bounded away from zero and bounded from above. Assume that for all $t \in [0, T]$, $x \in \mathbb{R}$ and $a \in A$ it holds that $f(t, x, a) = \eta_t |a|^p + \gamma_t |x|^q$ and $g(x) = \xi |x|^r$. Then (C0), (C1), (C2), (C3), (C4) and all assumptions of Theorem 3.6 are satisfied. In particular, there exists a solution (X, Y, Z) for the FBSDE (6).

4 Solving problem (P) in the Lipschitz case

In this section we turn to problem (P). We solve it by making similar assumptions on f as in subsection 3.1. More precisely, we assume conditions (C0) - (C4) as well as (D1) - (D3). However, instead of (D3) we make the even stronger assumption that

(D3') the whole Hessian matrix $D^2 f(s, x, a)$ of f w.r.t. $(x, a) \in [0, \infty) \times A_+$ is uniformly bounded independently of (ω, s, x, a) .

Moreover, we assume that

(D4) g has the form $g(x) = g^L(x) = Lx^2$, $x \in \mathbb{R}$, with some fixed $L > 0$.

Finally, we require that

(D5) $\sup A = \infty$ and $f_a(t, x, 0) = 0$ for all $(\omega, t, x) \in \Omega \times [0, T] \times [0, \infty)$.

For every initial value $x_0 > 0$ and every penalty function $g^L(x) = Lx^2$, $L > 0$, we have, according to Theorem 3.1 combined with Theorem 2.5, a unique solution (X^L, Y^L, Z^L) to the FBSDE (6), as well as a unique weakly regular decoupling field u^L associated with (6). According to Proposition 1.6 the strategy $\alpha_t^L := f_y^*(t, X_t^L, Y_t^L)$ minimizes J^L defined in (13).

The main objective of this section is to show that, under the above assumptions, α^L converges for $L \rightarrow \infty$ to an admissible strategy $\alpha^\infty \in \mathcal{A}^0(x_0)$, which minimizes $\hat{J}(x_0, \cdot)$, i.e. provides an optimal strategy for problem (12). We do so by first proving convergence of u^L to some limit u^∞ and then showing convergence of X^L to a limit X^∞ . This will finally lead us to the limit $\alpha^\infty \in \mathcal{A}^0(x_0)$.

Note that assumption (D3') is equivalent to assuming that the components f_{xx} , f_{aa} and $f_{xa} = f_{ax}$ of the Hessian D^2f are bounded. We denote by $\|f_{xx}\|_\infty$, $\|f_{aa}\|_\infty$ and $\|f_{xa}\|_\infty$ the respective uniform bounds of these mappings.

Due to Remark 1.4 assumption (D5) implies that $f_{yy}^*(t, x, y) = \frac{1}{f_{aa}}(t, x, f_y^*(t, x, y))$ for $y > 0$, while $f_{yy}^*(t, x, y) = 0$ if $y \leq 0$. Together with boundedness of f_{aa} this implies that for positive y the value $f_{yy}^*(t, x, y)$ can be uniformly bounded away from 0, while assumption (D1) implies that f_{yy}^* is bounded from above by some value $\|f_{yy}^*\|_\infty \in [0, \infty)$ as well.

Lemma 4.1. *There exist constants $C_1, C_2 \in (0, \infty)$, which depend on T , $\|f_{yy}^*\|_\infty$, $\|f_{xx}\|_\infty$, $\|f_{aa}\|_\infty$ and $\|f_{xa}\|_\infty$ only and are monotonically increasing in these values, such that for all $L > 0$ we have the following estimates for u_x^L :*

$$\frac{1}{C_1 \left(\frac{1}{2L} + (T-t) \right)} =: \kappa_t^L \leq u_x^L(t, x) \leq \gamma_t := C_2 \left(1 + \frac{1}{T-t} \right), \quad (23)$$

for all $t \in [0, T)$ and a.a. $x > 0$. As a consequence

$$x \cdot \kappa_t^L \leq u^L(t, x) \leq x \cdot \gamma_t, \quad (24)$$

for all $t \in [0, T)$ and $x > 0$.

Proof. Let $t \in [0, T)$ and $x > 0$ be fixed. Consider the processes $(X, Y, Z) = (X^L, Y^L, Z^L)$ on $[t, T]$ solving FBSDE (11) with initial condition $X_t = x$, terminal condition $g = g^L$ and satisfying the decoupling condition via $u = u^L$. Define $\Psi_s := u_x(s, X_s)$. According to the proof of Theorem 3.1 the non-negative process Ψ satisfies

$$d\Psi_s = \tilde{Z}_s dW_s - \left(f_{xx}(s, X_s, \alpha_s) - f_{yy}^*(s, X_s, Y_s) (\Psi_s - f_{xa}(s, X_s, \alpha_s))^2 \right) ds,$$

$s \in [t, T]$, (see (21)) and $\Psi_T = g''(X_T) = 2L$, where $\alpha_s := f_y^*(s, X_s, Y_s)$. Note that $\det(D^2f) = f_{xx}f_{aa} - f_{xa}^2$ and $f_{aa}(s, X_s, \alpha_s) = \frac{1}{f_{yy}^*}(s, X_s, Y_s)$, which leads to

$$\begin{aligned} d\Psi_s &= \tilde{Z}_s dW_s - \left(f_{xx}(s, X_s, \alpha_s) \mathbf{1}_{\{Y_s \leq 0\}} \right. \\ &\quad \left. + f_{yy}^*(s, X_s, Y_s) (\det(D^2f)(s, X_s, \alpha_s) + 2f_{xa}(s, X_s, \alpha_s)\Psi_s - \Psi_s^2) \right) ds. \end{aligned} \quad (25)$$

We now define a generator h via

$$h(s, \psi) := f_{xx}(s, X_s, \alpha_s) \mathbf{1}_{\{Y_s \leq 0\}} + f_{yy}^*(s, X_s, Y_s) (\det(D^2f) + 2\psi f_{xa} - \psi^2)(s, X_s, \alpha_s),$$

such that $d\Psi_s = \tilde{Z}_s dW_s - h(s, \Psi_s) ds$. Furthermore, we define a generator \hat{h} via

$$\hat{h}(s, \psi) := \|f_{yy}^*\|_\infty (-2\|f_{xa}\|_\infty \psi - \psi^2),$$

such that $\hat{h} \leq h$ for non-negative ψ , due to $f_{xx}, \det(D^2f) \geq 0$, which holds because f is convex in (x, a) . For $K_1 := 2\|f_{xa}\|_\infty \|f_{yy}^*\|_\infty$ define the deterministic and bounded non-negative process $\hat{\psi}$ via

$$\hat{\psi}_s := \left(\frac{e^{K_1(T-s)}}{2L} + \|f_{yy}^*\|_\infty (T-s) \frac{e^{K_1(T-s)} - 1}{K_1(T-s)} \right)^{-1},$$

where an expression of the form $\frac{e^x-1}{x}$ is to be replaced by $1 = \lim_{x \rightarrow 0} \frac{e^x-1}{x}$ in case x is zero. It is straightforward to verify that $\hat{\psi}$ solves the ODE given by $d\hat{\psi}_s = -\hat{h}(s, \hat{\psi}_s)ds$, $\hat{\psi}_T = 2L$. The comparison principle implies $\Psi \geq \hat{\psi}$, which in turn implies $\kappa_s^L \leq \Psi_s$ for an appropriately chosen C_1 depending on $T \cdot K_1$ and $\|f_{yy}^*\|_\infty$.

This shows the lower bound in (23). Combined with Lemma 3.4 and $u(t, 0) = 0$ this implies that $Y_s = u(s, X_s) > 0$ for all $s \in [t, T)$, which simplifies the generator h as the first summand vanishes a.e.

We now prove the upper bound on $\Psi_t = u_x(t, x)$. To this end define another locally Lipschitz generator via

$$\check{h}(s, \psi) := K_2 - \frac{1}{2\|f_{aa}\|_\infty} \psi^2,$$

where $K_2 := \|f_{xx}\|_\infty + 3\|f_{xa}\|_\infty^2 \|f_{yy}^*\|_\infty \geq 0$.

Note that $h \leq \check{h}$, because $\det(D^2 f) = f_{xx}f_{aa} - f_{xa}^2$,

$$2f_{xa}\psi = 2\left(\sqrt{2}f_{xa}\right)\left(\frac{1}{\sqrt{2}}\psi\right) \leq 2\|f_{xa}\|_\infty^2 + \frac{1}{2}\psi^2$$

and $f_{yy}^*(s, x, y) \geq \frac{1}{\|f_{aa}\|_\infty}$ for positive y .

We define $K_3 := \frac{1}{2\|f_{aa}\|_\infty}$. Note that $K_3 \in (0, \infty)$, but $K_2 = 0$ is possible.

Now define the deterministic and bounded process $\check{\psi}$ via

$$\check{\psi}_s = \sqrt{\frac{K_2}{K_3}} \left(1 + \frac{2}{\left(1 + 2\left(2L\sqrt{\frac{K_3}{K_2}} - 1\right)^{-1}\right) \exp(2\sqrt{K_2 K_3}(T-s)) - 1} \right),$$

for $K_2 > 0$. If $K_2 = 0$ we obtain the right definition for $\check{\psi}$ by passing to the limit $K_2 \downarrow 0$ in the above expression:

$$\check{\psi}_s = \sqrt{\frac{1}{K_3}} \cdot \frac{1}{(2L\sqrt{K_3})^{-1} + \sqrt{K_3}(T-s)}.$$

In any case it is straightforward to verify that $\check{\psi}$ solves the ODE given by $d\check{\psi}_s = -\check{h}(s, \check{\psi}_s)ds$, $\check{\psi}_T = 2L$. The comparison principle implies $\Psi \leq \check{\psi}$. Now observe that $\check{\psi}$ is monotonically increasing in L and if $K_2 > 0$ then it converges to

$$\sqrt{\frac{K_2}{K_3}} \left(1 + \frac{2}{\exp(2\sqrt{K_2 K_3}(T-s)) - 1} \right).$$

If $K_2 = 0$ then $\check{\psi}_s$ converges to the upper bound

$$\frac{1}{K_3(T-s)}$$

as $L \uparrow \infty$. In any case we can control $\check{\psi}$ by

$$\sqrt{\frac{K_2}{K_3}} + \frac{1}{K_3(T-s)},$$

which is controlled from above by γ_s with $C_2 := \sqrt{\frac{K_2}{K_3}} \vee \frac{1}{K_3}$. Thus we have proven $\Psi \leq \gamma$ and, therefore, the upper bound in (23).

The estimate (24) follows from (23) since $u(t, 0) = 0$. \square

In particular for every pair $(s, x) \in [0, T] \times (0, \infty)$ the value $|u^L(s, x)|$ can be uniformly bounded independently of L and ω . Also, $u^L(s, x) = 0$ for $x \leq 0$ according to Theorem 3.1. In order to have convergence of u^L for $L \rightarrow \infty$ it is now sufficient to show

Lemma 4.2. *The mapping $(t, \omega, L, x) \mapsto u^L(t, \omega, x)$ is progressively measurable while being continuous and non-decreasing in L .*

Proof. Consider the FBSDE given by the backward equation as in (6) and the forward equation, which is two-dimensional with the first component being identical to the forward equation in (6) and the second having the form $L_t = L_0$, where $L_0 \in \mathbb{R}$ is the initial value for this second component of the two-dimensional forward equation. Furthermore, instead of $g'(X_T)$, we choose the terminal condition to be $\hat{g}(x, L) := (0 \vee (2x) \wedge C) \cdot (0 \vee L \wedge C)$, where $C > 0$ is an arbitrary but fixed positive constant. Clearly, such an FBSDE satisfies (SLC). We claim that its maximal interval is $[0, T]$. To this end choose arbitrary $t_0 \in I_{\max}$ and $x = X_{t_0}, L = L_{t_0} \in \mathbb{R}$ and consider the corresponding X, Y, Z on $[t_0, T]$.

We write $\hat{C}_t = \frac{dY_t}{dL}$, $\hat{D}_t = \frac{dX_t}{dL}$ and $\gamma_t = u_L(t, X_t, L_t)$, where $t \in [t_0, T]$ and where u is the unique weakly regular decoupling field on $[t_0, T]$. Using the decoupling condition $Y_t = u(t, X_t, L_t)$ we obtain $\gamma_t = \hat{C}_t - \Psi_t \hat{D}_t$, where $\Psi_t = u_x(t, X_t, L_t)$. Observe that \hat{D}, \hat{C} and Ψ satisfy similarly to (18) the dynamics

$$d\hat{D}_t = - \left(f_{yx}^*(t, X_t, Y_t) \hat{D}_t + f_{yy}^*(t, X_t, Y_t) \hat{C}_t \right) dt,$$

similarly to (19) the dynamics

$$\begin{aligned} -d\hat{C}_t &= - \frac{dZ_t}{dx} dW_t + f_{xx}(t, X_t, \alpha_t) \hat{D}_t dt \\ &\quad + f_{xa}(t, X_t, \alpha_t) \left[f_{yx}^*(t, X_t, Y_t) \hat{D}_t + f_{yy}^*(t, X_t, Y_t) \hat{C}_t \right] dt, \end{aligned}$$

and similarly to (21) the dynamics

$$\begin{aligned} d\Psi_t &= D_t^{-1} \frac{dZ_t}{dx} dW_t - f_{xx}(t, X_t, \alpha_t) dt \\ &\quad + f_{yy}^*(t, X_t, Y_t) (\Psi_t - f_{xa}(t, X_t, \alpha_t))^2 dt, \end{aligned}$$

where $\alpha_t = f_y^*(t, X_t, Y_t)$. This implies

$$\begin{aligned} d\gamma_t &= d\hat{C}_t - \Psi_t d\hat{D}_t - \hat{D}_t d\Psi_t \\ &= \frac{dZ_t}{dx} dW_t - f_{xx}(t, X_t, \alpha_t) \hat{D}_t dt \\ &\quad - f_{xa}(t, X_t, \alpha_t) \left[f_{yx}^*(t, X_t, Y_t) \hat{D}_t + f_{yy}^*(t, X_t, Y_t) \hat{C}_t \right] dt \\ &\quad + \Psi_t \left(f_{yx}^*(t, X_t, Y_t) \hat{D}_t + f_{yy}^*(t, X_t, Y_t) \hat{C}_t \right) dt \\ &\quad - \hat{D}_t D_t^{-1} \frac{dZ_t}{dx} dW_t + \hat{D}_t f_{xx}(t, X_t, \alpha_t) dt \\ &\quad - \hat{D}_t f_{yy}^*(t, X_t, Y_t) (\Psi_t - f_{xa}(t, X_t, \alpha_t))^2 dt \\ &= f_{yy}^*(t, X_t, Y_t) (\Psi_t - f_{xa}(t, X_t, \alpha_t)) \gamma_t dt + \Gamma_t dW_t, \end{aligned}$$

where $\Gamma_t = \frac{dZ_t}{dL} - \widehat{D}_t D_t^{-1} \frac{dZ_t}{dx}$. Hence, the bounded process γ satisfies a linear BSDE with a bounded trend coefficient as Ψ is uniformly bounded according to the proof of Lemma 3.2. Note that

$$u(T, x, L) = (0 \vee (2x) \wedge C) \cdot (0 \vee L \wedge C),$$

such that $u_x(T, X_T, L_T)$ and $u_L(T, X_T, L_T)$ are both non-negative and uniformly bounded (with the respective bounds depending on C). It follows that $u_L(t, x, L)$ is non-negative whenever X_T is non-negative and also that γ is uniformly bounded, i.e. independently of t_0 .

Therefore, for every $C > 0$, we have a unique weakly regular decoupling field u on $[0, T]$ such that $u_L \geq 0$. Finally, if we choose the initial values x, L between 0 and C and consider the corresponding FBSDE, we observe that X_t, L_t both stay in $[0, C]$, such that the terminal condition $Y_T = 2X_T L$ is satisfied, which means that X, Y solve FBSDE (6) with terminal condition g^L . Uniqueness of solutions implies $u^L(t_0, x) = u(t_0, x, L)$ and the monotonicity of u in L is inherited by u^L as $C > 0$ can be chosen arbitrarily large. \square

Now we can define $u^\infty : \Omega \times [0, T] \times \mathbb{R}$ via $u^\infty(s, x) := \lim_{L \rightarrow \infty} u^L(s, x)$. Note that u^∞ inherits progressive measurability from u^L . Also note that for all $s < T$ the mapping $u^L(s, \cdot)$ is Lipschitz continuous w.r.t. x with Lipschitz constant γ_s , which does not depend on L . Therefore, $u^\infty(s, \cdot)$ is also Lipschitz continuous with the same Lipschitz constant. Finally, note that for all $s < T$ and all $x > 0$

$$u^\infty(s, x) \geq x \cdot \lim_{L \rightarrow \infty} \kappa_s^L = \frac{x}{C_1(T-s)}. \quad (26)$$

Now let $x_0 > 0$. Since u^∞ restricted to $\Omega \times [0, T - \varepsilon] \times \mathbb{R}$ is progressively measurable and uniformly Lipschitz continuous in the last component for every $\varepsilon \in (0, T)$, we can define X^∞ as the unique solution to the ODE

$$X_t^\infty := x_0 - \int_0^t f_y^*(s, X_s^\infty, u^\infty(s, X_s^\infty)) ds, \quad t \in [0, T],$$

which is motivated by passing to the limit $L \rightarrow \infty$ in

$$X_t^L = x_0 - \int_0^t f_y^*(s, X_s^L, u^L(s, X_s^L)) ds.$$

Note that X^∞ is defined on $[0, T)$ only. We now prove

Lemma 4.3. *It holds that $\lim_{L \rightarrow \infty} X^L = X^\infty$ and $\lim_{L \rightarrow \infty} u^L(\cdot, X^L) \rightarrow u^\infty(\cdot, X^\infty)$ almost everywhere on $\Omega \times [0, T)$.*

Proof. Note that the generators $f_y^*(s, \cdot, u^L(s, \cdot))$ are monotonically non-decreasing in L according to Lemma 4.2. A comparison principle for ODEs entails that X^L is non-increasing in L . As X^L is non-negative for all $L > 0$ we can define the progressively measurable process $\hat{X} := \lim_{L \rightarrow \infty} X^L$ as an a.e. limit. Next observe that $u^L(\cdot, X^L) \rightarrow u^\infty(\cdot, \hat{X})$ a.e. as $L \rightarrow \infty$. Indeed,

$$\begin{aligned} |u^L(s, X_s^L) - u^\infty(s, X_s^\infty)| &\leq \left| u^L(s, X_s^L) - u^L(s, \hat{X}_s) \right| + \left| u^L(s, \hat{X}_s) - u^\infty(s, \hat{X}_s) \right| \\ &\leq \gamma_s \left| X_s^L - \hat{X}_s \right| + \left| u^L(s, \hat{X}_s) - u^\infty(s, \hat{X}_s) \right|, \end{aligned}$$

$s \in [0, T]$, with both summands on the right-hand-side converging to 0 as $L \rightarrow \infty$. Now dominated convergence implies

$$\lim_{L \rightarrow \infty} \int_0^t f_y^*(s, X_s^L, u^L(s, X_s^L)) ds = \int_0^t f_y^*(s, \hat{X}_s, u^\infty(s, \hat{X}_s)) ds,$$

which yields

$$\hat{X}_t = x_0 - \int_0^t f_y^*(s, \hat{X}_s, u^\infty(s, \hat{X}_s)) ds,$$

for all $t < T$. This already implies $\hat{X} = X^\infty$ a.e. \square

Note that X^L and therefore X^∞ is non-negative. Furthermore, X^∞ is non-increasing, since f_y^* is non-negative. Therefore, $\lim_{t \rightarrow T} X_t^\infty$ exists and we can continuously extend the process X^∞ to the whole of $[0, T]$ via $X_T^\infty := \lim_{t \rightarrow T} X_t^\infty$. Remark that $X_T^\infty \geq 0$ a.s.

Lemma 4.4. *It holds that $X_T^\infty = 0$.*

Proof. Firstly note that for every fixed (ω, s, x) the function $f_y^*(s, x, \cdot) : [0, \infty) \rightarrow [0, \infty)$ starts in zero and is strictly increasing with a derivative which is larger or equal $\frac{1}{\|f_{aa}\|_\infty}$ almost everywhere. Using (26) and since $t \mapsto X_t^\infty(\omega)$ is non increasing, we obtain for any $0 \leq t < T$

$$\begin{aligned} X_t^\infty(\omega) &= x_0 - \int_0^t f_y^*(s, X_s^\infty, u^\infty(s, X_s^\infty))(\omega) ds \leq x_0 - \int_0^t \frac{1}{\|f_{aa}\|_\infty} \cdot \frac{X_s^\infty(\omega)}{C_1(T-s)} ds \\ &\leq x_0 - \frac{X_t^\infty(\omega)}{\|f_{aa}\|_\infty C_1} \int_0^t \frac{1}{(T-s)} ds. \end{aligned}$$

Hence

$$X_t^\infty(\omega) \left[1 + \frac{1}{\|f_{aa}\|_\infty C_1} \int_0^t \frac{1}{(T-s)} ds \right] \leq x_0.$$

When t tends to T , we deduce that $X_T^\infty = 0$ a.s. \square

We now show the two main results of this chapter. We define $\alpha_s^\infty := f_y^*(s, X_s^\infty, u^\infty(s, X_s^\infty))$, for $s \in [0, T]$, while setting $\alpha_T^\infty := 0$. Note that $\alpha^\infty \in \mathcal{A}^0(x_0)$ according to Lemma 4.4.

Theorem 4.5. *The strategy α^∞ minimizes $\hat{J}(x_0, \cdot)$.*

Proof. Let $\alpha \in \mathcal{A}^0(x_0)$ be any admissible control such that $X_T^{0, x_0, \alpha} = 0$. Since α^L minimizes J^L we have:

$$\mathbb{E} \left[\int_0^T f(s, X_s^{0, x_0, \alpha}, \alpha_s) ds + g^L(X_T^{0, x_0, \alpha}) \right] \geq \mathbb{E} \left[\int_0^T f(s, X_s^{0, x_0, \alpha^L}, \alpha_s^L) ds + g^L(X_T^{0, x_0, \alpha^L}) \right],$$

which implies

$$\mathbb{E} \left[\int_0^T f(s, X_s^{0, x_0, \alpha}, \alpha_s) ds \right] \geq \mathbb{E} \left[\int_0^T f(s, X_s^{0, x_0, \alpha^L}, \alpha_s^L) ds \right].$$

Note that $X^{0, x_0, \alpha^L} = X^L$, which converges to $X^\infty = X^{0, x_0, \alpha^\infty}$ a.s. for $L \rightarrow \infty$ by Lemma 4.3. We now aim at showing that $\alpha^L = f_y^*(\cdot, X^L, u^L(\cdot, X^L))$ converges for $L \rightarrow \infty$ to α^∞

a.e. as well. Since $f_y^*(s, \cdot, \cdot)$ is continuous, it is sufficient to show that $u^L(s, X_s^L)$ converges to $u^\infty(s, X_s^\infty)$ a.s. The latter follows from

$$\begin{aligned} |u^L(s, X_s^L) - u^\infty(s, X_s^\infty)| &\leq |u^L(s, X_s^L) - u^L(s, X_s^\infty)| + |u^L(s, X_s^\infty) - u^\infty(s, X_s^\infty)| \\ &\leq \gamma_s |X_s^L - X_s^\infty| + |u^L(s, X_s^\infty) - u^\infty(s, X_s^\infty)|, \end{aligned}$$

for all $s < T$. Note that both summands on the right-hand-side converge to 0 as $L \rightarrow \infty$.

Due to continuity of $f(s, \cdot, \cdot)$ the process $f(\cdot, X_s^{0, x_0, \alpha^L}, \alpha_s^L)$ converges to $f(\cdot, X_s^{0, x_0, \alpha^\infty}, \alpha_s^\infty)$ for $L \rightarrow \infty$ almost everywhere. Note also that at a time $s \in [0, T)$ the value $f(s, X_s^L, \alpha_s^L)$ is bounded by $f(s, x_0, f_y^*(s, x_0, x_0 \gamma_s))$. For every fixed $\varepsilon \in (0, T)$ the latter value is uniformly bounded independently of $\omega \in \Omega$ and $s \in [0, T - \varepsilon]$, since $f(s, 0, f_y^*(s, 0, 0)) = 0$, (C3) and (D1). Therefore, by dominated convergence, we have

$$\lim_{L \rightarrow \infty} \mathbb{E} \left[\int_0^{T-\varepsilon} f(s, X_s^{0, x_0, \alpha^L}, \alpha_s^L) ds \right] = \mathbb{E} \left[\int_0^{T-\varepsilon} f(s, X_s^{0, x_0, \alpha^\infty}, \alpha_s^\infty) ds \right]$$

and hence

$$\mathbb{E} \left[\int_0^T f(s, X_s^{0, x_0, \alpha}, \alpha_s) ds \right] \geq \mathbb{E} \left[\int_0^{T-\varepsilon} f(s, X_s^{0, x_0, \alpha^\infty}, \alpha_s^\infty) ds \right].$$

Finally, by monotone convergence we obtain

$$\mathbb{E} \left[\int_0^T f(s, X_s^{0, x_0, \alpha}, \alpha_s) ds \right] \geq \mathbb{E} \left[\int_0^T f(s, X_s^{0, x_0, \alpha^\infty}, \alpha_s^\infty) ds \right].$$

□

Example 4.6. Let $C \in (1, \infty)$ and let $\eta, \gamma: [0, T] \times \Omega \rightarrow [0, \infty)$ be progressively measurable stochastic processes such that for all $t \in [0, T]$ it holds a.s. that $\eta_t \geq \frac{1}{C}$ and $\max(\eta_t, \gamma_t) \leq C$. Assume that for all $t \in [0, T]$, $x \in \mathbb{R}$ and $a \in \mathbb{R}$ it holds that $f(t, x, a) = \eta_t \frac{|a|^3 + 2|a|^2}{|a| + 1} + \gamma_t |x|^2$. Then conditions (C0), (C1), (C2), (C3) and (C4) are satisfied. Moreover, it holds for all $t \in [0, T]$, $x \in \mathbb{R}$ and $a \in \mathbb{R}$ a.s. that $\frac{2}{C} \leq 2\eta_t \leq f_{aa}(t, x, a) \leq 4\eta_t \leq 4C$. This together with (5) implies that $[0, \infty)^2 \ni (x, y) \mapsto f_y^*(t, x, y)$ is uniformly Lipschitz continuous. It follows that (D1) is satisfied. Moreover, it follows that (D3') is satisfied. Observe that $f_a(t, x, a) = 0$ if and only if $a = 0$. This implies (D2). Condition (D5) is also satisfied. Therefore, it follows from Theorem 4.5 that α^∞ is an optimal control in Problem (P).

Theorem 4.7. Let $Y^\infty: [0, T] \times \Omega \rightarrow \mathbb{R}$ satisfy for all $t \in [0, T)$ a.s. that $Y_t^\infty = u^\infty(t, X_t^\infty)$. Then the sequence Z^L converges in $L^2((0, t) \times \Omega)$ for any $t < T$ to Z^∞ and the process $(X^\infty, Y^\infty, Z^\infty)$ satisfies for all $0 \leq t \leq r < T$ a.s. that

$$\begin{aligned} X_t^\infty &= x_0 - \int_0^t f_y^*(s, X_s^\infty, Y_s^\infty) ds, & X_T^\infty &= 0 \\ Y_t^\infty &= Y_r^\infty + \int_t^r f_x(s, X_s^\infty, f_y^*(s, X_s^\infty, Y_s^\infty)) ds - \int_t^r Z_s^\infty dW_s. \end{aligned} \quad (27)$$

Proof. Let $0 \leq t \leq r < T$. The fact that X^∞ satisfies $X_t^\infty = x_0 - \int_0^t f_y^*(s, X_s^\infty, Y_s^\infty) ds$ follows directly from the definition of X^∞ and Y^∞ . By Lemma 4.4 it holds that $X_T^\infty = 0$. It follows from Theorem 3.1 that for all $L > 0$ it holds

$$Y_t^L = Y_r^L + \int_t^r f_x(s, X_s^L, f_y^*(s, X_s^L, Y_s^L)) ds - \int_t^r Z_s^L dW_s. \quad (28)$$

Applying Itô's formula for $L \leq L'$ we have for any $0 \leq t \leq r < T$

$$\begin{aligned} |Y_t^L - Y_t^{L'}|^2 &= |Y_r^L - Y_r^{L'}|^2 - \int_t^r |Z_s^L - Z_s^{L'}|^2 ds - \int_t^r (Z_s^L - Z_s^{L'}) dW_s \\ &\quad + 2 \int_t^r (Y_s^L - Y_s^{L'}) (f_x(s, X_s^L, f_y^*(s, X_s^L, Y_s^L)) - f_x(s, X_s^{L'}, f_y^*(s, X_s^{L'}, Y_s^{L'}))) ds. \end{aligned}$$

By assumption (D1) there exists $C > 0$ such that for all $s \in [t, r]$ and $0 < L \leq L'$

$$|f_x(s, X_s^L, f_y^*(s, X_s^L, Y_s^L)) - f_x(s, X_s^{L'}, f_y^*(s, X_s^{L'}, Y_s^{L'}))| \leq C(|X_s^L - X_s^{L'}| + |Y_s^L - Y_s^{L'}|).$$

Thereby with Young's inequality and taking the expectation, we obtain

$$\begin{aligned} \mathbb{E} \int_t^r |Z_s^L - Z_s^{L'}|^2 ds &\leq \mathbb{E} |Y_r^L - Y_r^{L'}|^2 \\ &\quad + 2C \mathbb{E} \int_t^r |Y_s^L - Y_s^{L'}| (|X_s^L - X_s^{L'}| + |Y_s^L - Y_s^{L'}|) ds \\ &\leq \mathbb{E} |Y_r^L - Y_r^{L'}|^2 \\ &\quad + (2C + C^2) \mathbb{E} \int_t^r |Y_s^L - Y_s^{L'}|^2 ds + \mathbb{E} \int_t^r |X_s^L - X_s^{L'}|^2 ds. \end{aligned} \tag{29}$$

Lemma 4.3 implies that $X_t^L \rightarrow X_t^\infty$ and $Y_t^L \rightarrow Y_t^\infty$ for $L \rightarrow \infty$ a.s. Moreover, it holds for all $s \in [t, r]$ and $L > 0$ that $\max(X_s^L, X_s^\infty) \leq x_0$ and by (24) that $\max(|Y_s^L|, |Y_s^\infty|) \leq \gamma_r x_0$. By the dominated convergence theorem, the two sequences X^L and Y^L also converge in the space $L^2((0, r) \times \Omega)$. Combining this with (29), we deduce that Z^L is a Cauchy sequence in $L^2((0, r) \times \Omega)$ and thus converges to Z^∞ . It is straightforward that the limit Z^∞ does not depend on the particular choice of $r < T$.

Again by assumption (D1)

$$|f_x(s, X_s^L, f_y^*(s, X_s^L, Y_s^L)) - f_x(s, X_s^\infty, f_y^*(s, X_s^\infty, Y_s^\infty))| \leq C(|X_s^L - X_s^\infty| + |Y_s^L - Y_s^\infty|),$$

and we can take the limit $L \rightarrow \infty$ in (28) to obtain

$$Y_t^\infty = Y_r^\infty + \int_t^r f_x(s, X_s^\infty, f_y^*(s, X_s^\infty, Y_s^\infty)) ds - \int_t^r Z_s^\infty dW_s.$$

This completes the proof. \square

Remark 4.8. From the estimate (26), we obtain

$$\liminf_{t \rightarrow T} \left(\frac{Y_t^\infty}{X_t^\infty} \right) = +\infty.$$

This behaviour is similar to the weak terminal condition (1.3) in [2]. Moreover from (27), the process

$$Y^\infty + \int_0^\cdot f_x(s, X_s^\infty, f_y^*(s, X_s^\infty, Y_s^\infty)) ds$$

is a non negative local martingale on $[0, T)$. Hence its limit at time T exists in $[0, \infty)$ a.s. By the monotone convergence theorem, the integral has also a limit. Since both terms are non negative, we deduce that Y^∞ has a limit at time T a.s. and Y_T^∞ is finite a.s. Note that here Y_T^∞ is not a given terminal condition, but part of the solution.

References

- [1] A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *SIAM J. Financial Math.*, 3(1):511–533, 2012.
- [2] S. Ankirchner, M. Jeanblanc, and T. Kruse. BSDEs with Singular Terminal Condition and a Control Problem with Constraints. *SIAM J. Control Optim.*, 52(2):893–913, 2014.
- [3] S. Ankirchner and T. Kruse. Optimal position targeting with stochastic linear-quadratic costs. In *Advances in mathematics of finance*, volume 104 of *Banach Center Publ.*, pages 9–24. Polish Acad. Sci. Inst. Math., Warsaw, 2015.
- [4] P. Bank and M. Voß. Linear quadratic stochastic control problems with singular stochastic terminal constraint. *ArXiv e-prints*, Nov. 2016.
- [5] P. Barrieu, N. Cazanave, and N. El Karoui. Closedness results for BMO semi-martingales and application to quadratic BSDEs. *C. R. Math. Acad. Sci. Paris*, 346(15-16):881–886, 2008.
- [6] F. Clarke. *Functional analysis, calculus of variations and optimal control*, volume 264 of *Graduate Texts in Mathematics*. Springer, London, 2013.
- [7] F. Delarue. On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stochastic Process. Appl.*, 99(2):209–286, 2002.
- [8] L. C. Evans. An introduction to mathematical optimal control theory. *Lecture Notes, University of California, Department of Mathematics, Berkeley*, 2005.
- [9] A. Fromm. *Theory and applications of decoupling fields for forward-backward stochastic differential equations*. PhD thesis, Humboldt-Universität zu Berlin, 2015.
- [10] A. Fromm and P. Imkeller. Existence, Uniqueness and Regularity of Decoupling Fields to Multi-dimensional Fully Coupled FBSDEs. *Preprint arXiv:1310.0499*, 2013.
- [11] I. M. Gelfand and S. V. Fomin. *Calculus of variations*. Revised English edition translated and edited by Richard A. Silverman. Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.
- [12] P. Graewe, U. Horst, and J. Qiu. A non-Markovian liquidation problem and backward SPDEs with singular terminal conditions. *SIAM J. Control Optim.*, 53(2):690–711, 2015.
- [13] T. Kruse. *Inverse Optimal Stopping and Optimal Closure of Illiquid Positions*. PhD thesis, Universitäts- und Landesbibliothek Bonn, 2014.
- [14] T. Kruse and A. Popier. Minimal supersolutions for BSDEs with singular terminal condition and application to optimal position targeting. *Stochastic Process. Appl.*, 126(9):2554–2592, 2016.
- [15] J. Ma, P. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitly – a four step scheme. *Probab. Theory Relat. Fields*, 98(3):339–359, 1994.
- [16] J. Ma, Z. Wu, D. Zhang, and J. Zhang. On well-posedness of forward-backward SDEs—a unified approach. *Ann. Appl. Probab.*, 25(4):2168–2214, 2015.
- [17] J. Ma, H. Yin, and J. Zhang. On non-Markovian forward-backward SDEs and backward stochastic PDEs. *Stochastic Process. Appl.*, 122(12):3980–4004, 2012.
- [18] J. Ma and J. Yong. *Forward-backward stochastic differential equations and their applications*, volume 1702 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1999.
- [19] E. Pardoux and S. Tang. Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probab. Theory Relat. Fields*, 114(2):123–150, 1999.
- [20] S. Peng and Z. Wu. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM J. Control Optim.*, 37(3):825–843, 1999.

- [21] A. Popier. Backward stochastic differential equations with singular terminal condition. *Stochastic Process. Appl.*, 116(12):2014–2056, 2006.
- [22] A. Popier. Backward stochastic differential equations with random stopping time and singular final condition. *Ann. Probab.*, 35(3):1071–1117, 2007.
- [23] J. Yong and X. Y. Zhou. *Stochastic controls*, volume 43 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1999. Hamiltonian systems and HJB equations.