HAL
open science

Solving fully coupled FBSDEs by minimizing a directly calculable error functional

Stefan Ankirchner, Alexander Fromm

To cite this version:

Stefan Ankirchner, Alexander Fromm. Solving fully coupled FBSDEs by minimizing a directly calculable error functional. 2018. hal-01915772

HAL Id: hal-01915772

https://hal.science/hal-01915772

Preprint submitted on 8 Nov 2018

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.

Solving fully coupled FBSDEs by minimizing a directly calculable error functional

Stefan Ankirchner*[1] and Alexander Fromm †[1]

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

November 7, 2018

Abstract

We present a new scheme for approximating solutions of forward-backward stochastic differential equations (FBSDEs). The scheme is particularly suitable for fully coupled FBSDEs. The approximation relies on a piecewise in time approximation by minimizing an error functional that measures how well a process triplet satifies the FBSDE. The error functional is minimized in a finite-dimensional linear space based on iterated integrals. We provide sufficient conditions for the approximations to converge at the rate $N^{-1/2}$, where N is the time discretization parameter.

2010 Mathematics Subject Classification. primary: 60H35, 65C30; secondary: 65G20, 93E20.

Keywords. Forward-backward stochastic differential equation, decoupling field, error functional, stochastic polynomial, convergence of approximations.

Introduction

In this article we present and analyze a new scheme for approximating the solution of a forward-backward stochastic differential equation (FBSDE) of the form

$$X_t = X_0 + \int_0^t \mu(r, X_r, Y_r, Z_r)dr + \int_0^t \sigma(r, X_r, Y_r, Z_r)dW_r, \tag{1}$$

$$Y_t = \xi(X_T) - \int_t^T f(r, X_r, Y_r, Z_r)dr - \int_t^T Z_r dW_r, \tag{2}$$

where $T > 0$ is a fixed time horizon, $t \in [0, T]$, W is a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$, and μ, σ, f and ξ are Lipschitz continuous parameters. Notice that the FBSDE is fully coupled, i.e. the dynamics of the backward equation (2) depends on the solution of

*s.ankirchner@uni-jena.de

†alexander.fromm@uni-jena.de

Financial support from the *German Research Foundation* through the project AN 1024/4-1 is gratefully acknowledged.

the forward equation (1) and, vice versa, the forward equation depends on the backward part. Recall that even under Lipschitz assumptions on the parameters a fully coupled FBSDE does not necessarily possess a solution on the whole interval $[0,T]$ (see e.g. in [MY99]). There does, however, exist a non-vanishing interval on which a solution exists (see e.g. [Ant93], [PT99], [Del02], [Zha06], [MWZZ15], [Fro15]). One can identify a *maximal* existence interval and on this interval there does also exist a so-called decoupling field, a function u that establishes the dependence of the backward process on the forward process via $Y_t = u(t, X_t)$ (see e.g. [Fro15]). The scheme that we present in the following provides an approximation of the decoupling field and can be used to determine the maximal existence interval in the first place.

We now describe our scheme for approximating the solution of a fully coupled FBSDE and its decoupling field u. For simplicity, we restrict ourselves to the one-dimensional case. Suppose first that there exists a decoupling field $u : [0,T] \times \Omega \times \mathbb{R} \to \mathbb{R}$ on the *whole* interval $[0,T]$. This implies that for all $(t,x) \in [0,T] \times \mathbb{R}$ a unique solution $(X^{t,x}, Y^{t,x}, Z^{t,x})$ of (1) and (2) on the subinterval $[t,T]$ with initial condition $X_t^{t,x} = x$ exists and satifies $u(s, X_s^{t,x}) = Y_s^{t,x}$ for all $s \in [t,x]$.

We approximate the decoupling field along an equidistant time grid $0 = t_0 < t_1 < \cdots < t_N = T$. Our algorithm approximates $u(t_j, \cdot)$ via a backward recursion along j. Note that $u(t_N, \cdot) = \xi$. So u is already known at the terminal time t_N.

Suppose that we have approximated $u(t_{j+1}, \cdot)$ by some function $\tilde{u}(t_{j+1}, \cdot)$ already. In order to approximate $u(t_j, \cdot)$ we first choose a space discretization $x_1 < x_2 < \cdots < x_S$. For any given $k \in \{1, \ldots, S\}$ we approximate the solution components X and Y of the FBSDE on $[t_j, t_{j+1}]$ with the initial condition $X_{t_j} = x_k$ and the terminal condition $\tilde{u}(t_{j+1}, \cdot)$ with processes X^α and Y^β of the form

$$X_s^\alpha = \alpha_1 + \alpha_2(s - t_j) + \alpha_3(W_s - W_{t_j}) + \alpha_4 \int_{t_1}^{s} \left(W_r - W_{t_j}\right) \mathrm{d}W_r, \tag{3}$$

$$Y_s^\beta = \beta_1 + \beta_2(s - t_j) + \beta_3(W_s - W_{t_j}) + \beta_4 \int_{t_j}^{s} \left(W_r - W_{t_j}\right) \mathrm{d}W_r, \tag{4}$$

for $s \in [t_j, t_{j+1}]$, where $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4) \in \mathbb{R}^4$ and $\beta = (\beta_1, \beta_2, \beta_3, \beta_4) \in \mathbb{R}^4$. We do so by minimizing the error functional

$$\begin{aligned} &G(\alpha, \beta) := \\ &\mathbb{E}\Bigg[|\alpha_1 - x_k|^2 + \frac{T}{N} \int_{t_j}^{t_{j+1}} |\alpha_2 - \mu(r, \Theta_r)|^2 \,\mathrm{d}r + \int_{t_j}^{t_{j+1}} |\alpha_3 + \alpha_4(W_r - W_{t_j}) - \sigma(r, \Theta_r)|^2 \,\mathrm{d}r \\ &+ \frac{T}{N} \int_{t_j}^{t_{j+1}} |\beta_2 - f(r, \Theta_r)|^2 \,\mathrm{d}r + |Y_{t_{j+1}}^\beta - \tilde{u}(t_{j+1}, X_{t_{j+1}}^\alpha)|^2\Bigg], \end{aligned} \tag{5}$$

where $\Theta_r = (X_r^\alpha, Y_r^\beta, \beta_3 + \beta_4(W_r - W_{t_j}))$ for $r \in [t_j, t_{j+1}]$.

Notice that the problem of minimizing (5) is a minimization problem in $\mathbb{R}^8$ and one can use, e.g., the Newton method to obtain a minimizer. We comment on the details regarding the application of the Newton method at a later point (see Section: 4). For now we simply assume that a global minimizer (α^*, β^*) of the functional $(\alpha, \beta) \mapsto G_x(X^\alpha, Y^\beta)$ can be calculated for every x_k, $k \in \{1, \ldots, S\}$.

One can interpret β_1^* as an approximation of $u(t_j, x_k)$. We set $w(t_j, x_k) := \beta_1^*$. We next interpolate the function $w(t_j, \cdot)$. A plain piecewise linear interpolation along the grid

points x_k, $k \in \{1, \dots, S\}$, does not necessarily provide a smooth approximation of $u(t_j, \cdot)$. By applying some standard smoothing techniques, e.g. a Gaussian kernel smoothing, one can obtain a smooth approximation $\tilde{u}(t_j, \cdot)$ of the function $u(t_j, \cdot)$.

By going backwards in time we obtain thus an approximation of the decoupling field on all times $t_0, t_1, \dots, t_N$. We can summarize the approximation method as follows:

The algorithm in a nutshell

1. Select a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$.
2. Select a set of supporting points $\{x_1, x_2, \dots, x_S\} \subset \mathbb{R}$.
3. Set $j = N - 1$.
4. For every supporting point x_k find the parameters $\alpha_1^*, \dots, \alpha_4^*$ and $\beta_1^*, \dots, \beta_4^*$ that minimize $(\alpha, \beta) \mapsto G(x_k, X^\alpha, Y^\beta)$. Set $w(t_j, x_k) = \beta_1^*$.
5. Choose a smooth function $\tilde{u}(t_j, \cdot)$ such that $\tilde{u}(t_j, x_k)$ is close to $w(t_j, x_k)$ for all $k \in \{1, \dots, S\}$.
6. If $j \neq 0$, then set $j = j - 1$ and go to 4.

If the maximal existence interval is not $[0, T]$ but of the form $(t_{\min}, T]$ with $t_{\min} \in [0, T]$, then the algorithm can be used for approximating the decoupling field on any compact subinterval of $(t_{\min}, T]$. Convergence of the algorithm on a specific subinterval $[a, b] \subset [0, T]$ indicates that $a > t_{\min}$. In this way, the maximal existence interval can be identified.

We provide a bunch of conditions on the FBSDE and the smoothing procedure in Step 5 of the algorithm guaranteeing that the approximations from the algorithm converge to the true solution of the FBSDE as the time step size converges to zero. One of the conditions is that the regularity from the terminal condition of the FBSDE is transferred, in a quantifiable way, to the decoupling field. Another condition assumes that some suitable semi-norms of X and Y are controllable on small intervals. Moreover we assume a sufficiently strong stability property with respect to the terminal condition. We show that if these conditions are satisfied, then the approximations converge to the true solution at the rate $N^{-1/2}$ (see Section 3).

We now explain why the scheme provides arbitrarily precise approximations. The error functional $G(\alpha, \beta)$ in (11) measures how well the processes X^α, Y^β and $Z^\beta = \beta_3 + \beta_4(W_\cdot - W_{t_j})$ satisfy the FBSDE (1), (2) on the interval $[t_j, t_{j+1}]$ with terminal condition $\tilde{u}(t_{j+1}, \cdot)$. Indeed, notice that if the error is zero, then $(X^\alpha, Y^\beta, Z^\beta)$ is an exact solution. The distance of a pair of processes (X^α, Y^β) to the true solution can be estimated in terms of the error functional G and, vice versa, the error functional can be estimated against the distance (see Section 2). Moreover, the true solution processes X and Y can be approximated with stochastic polynomials of the type X^α and Y^β arbitrarily well, respectively, if the time interval length is sufficiently small. Indeed, the approximation error is of the order $N^{-3/2}$. One can show that the distance of the true solution processes X and Y to the processes $(X^{\alpha^*}, Y^{\beta^*})$ is also of the order $N^{-3/2}$. By aggregating the errors over all time intervals we obtain a global approximation error of the order $N^{-1/2}$.

We close the introduction by comparying our new scheme to existing approximation methods. One of first methods described in the literature is the four step scheme of [MPY94]. The

four step scheme approximates an FBSDE by solving a related PDE numerically. The approach necessitates numerical methods and existence of a solution for the related PDE. In constrast, our approach is probabilistic by drawing on the dynamics of the FBSDE in the error functional.

The literature comprises already several schemes that directly approximate the FBSDEs, see e.g. the first part of [BS12] for a survey. In most existing approaches the forward dynamics are approximated explicitly along a discrete time grid, e.g. by using the Euler scheme. Moreover, the schemes rely on an explicit approximation formula for the backward process in terms of a conditional expectation. [BT04] consider a scheme for decoupled FBSDEs and prove a convergence rate of 1/2 of the approximation error due to the time discretization. [CM14] propose a scheme, in the decoupled case, with a more accurate representation of the control process allowing to obtain a convergence rate of 2. [BZ08] show a time discretization error of 1/2 for a scheme that covers also coupled FBSDEs. In contrast to the above, in the scheme proposed here the approximating dynamics are determined implicitly by minimizing the error functional $G(\alpha, \beta)$. The error functional incorporates the forward and backward dynamics simultaneously, and thus inherently reflects any coupling of the forward and backward part. Compared to iterative schemes that consider decoupled systems in each Picard sweep (e.g. as in [BZ08]), our approach directly captures the coupling within the FBSDE.

We remark that the scheme described can be modified to obtain approximations that converge at an arbitrarily fast rate. Indeed, by using stochastic polynomials of a higher order than that of X^α and Y^β one can increase the local and hence also the global convergence rate. To keep the convergence analysis simpler, we analyze the algorithm only for process approximations of order 3/2.

We stress that the practical implementation of the algorithm does not necessarily require Monte-Carlo estimations. Indeed, one can modify the error functional $G(\alpha, \beta)$ by replacing μ, σ, f and ξ with approximating polynomials. One thus obtains an approximation of G in closed form, the minimizer of which can be determined with analytic methods (see Section 4 for a more detailed explanation).

Finally, observe that obtaining approximations on small intervals by minimizing the error functional has the by-product of each approximation being endowed with an estimate for the distance to the real solution. Contrary to existing approaches this allows to assess the quality of the output obtained for a chosen discretization of $[0, T]$, and hence to choose the discretization dynamically to satisfy a required approximation quality. Such refinements of the scheme are also reserved for future research.

1 Decoupling fields under SLC

In this section we summarize some known results on decoupling fields for coupled FBSDEs.

Throughout the article let $T > 0$ be a fixed finite time horizon. Besides let W be a d-dimensional 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.

The dynamics of an FBSDE are given by

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

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

$$\xi\colon \Omega \times \mathbb{R}^n \to \mathbb{R}^m, \qquad \mu\colon [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^n,$$
$$\sigma\colon [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^{n\times d}, \qquad f\colon [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \to \mathbb{R}^m,$$

for $d, n, m \in \mathbb{N}$. Throughout the whole article μ, σ and f are assumed to be progressively measurable with respect to $(\mathcal{F}_t)_{t\in[0,T]}$, i.e. $\mu\mathbf{1}_{[0,t]}, \sigma\mathbf{1}_{[0,t]}, f\mathbf{1}_{[0,t]}$ must be $\mathcal{B}([0,T]) \otimes \mathcal{F}_t \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{R}^m) \otimes \mathcal{B}(\mathbb{R}^{m\times d})$ - measurable for all $t \in [0,T]$. Moreover, we assume that ξ is $\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R}^n)$-measurable.

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

Definition 1.1. Let $t \in [0, T]$. A function $u\colon [t, T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ 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 $X_{t_1}\colon \Omega \to \mathbb{R}^n$ there exist progressively measurable processes (X, Y, Z) on $[t_1, t_2]$ such that

$$\begin{aligned} X_s &= X_{t_1} + \int_{t_1}^s \mu(r, X_r, Y_r, Z_r)dr + \int_{t_1}^s \sigma(r, X_r, Y_r, Z_r)dW_r, \\ Y_s &= Y_{t_2} - \int_s^{t_2} f(r, X_r, Y_r, Z_r)dr - \int_s^{t_2} Z_r dW_r, \\ Y_s &= u(s, X_s), \end{aligned} \tag{6}$$

a.s. for all $s \in [t_1, t_2]$. In particular, the definition assumes that all integrals are well-defined.

Some remarks about this definition are in place.

- The first equation in (6) is called the *forward equation*, the second the *backward equation* and the third is 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 the decoupling condition implies $u(T, \cdot) = \xi$, a.e.
- If $t_2 = T$ we 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 adjacent intervals can be pasted together (see e.g. Lemma 2.1.2 of [Fro15]).

We 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 [Fro15]).

For the following we need to fix further notation.

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

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

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 Lipschitz continuous in $x \in \mathbb{R}^n$ for every (s,ω).

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 $\operatorname{ess\,sup} \mathbb{E}[F|\mathcal{F}_t]$, 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, where S^{n-1} is the $(n-1)$ - dimensional sphere.

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 1.2. Let $u\colon [t,T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ 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 (6) are a.e. unique and satisfy

$$\sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}[|X_s|^2] + \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}[|Y_s|^2] + \mathbb{E}_{t_1,\infty}\left[\int_{t_1}^{t_2} |Z_s|^2 ds\right] < \infty, \tag{7}$$

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

$$\begin{aligned}
&\operatorname{ess\,sup}_{x\in\mathbb{R}^n} \sup_{v\in S^{n-1}} \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}\left[\left|\frac{\mathrm{d}}{\mathrm{d}x} X_s\right|_v^2\right] < \infty,\\
&\operatorname{ess\,sup}_{x\in\mathbb{R}^n} \sup_{v\in S^{n-1}} \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}\left[\left|\frac{\mathrm{d}}{\mathrm{d}x} Y_s\right|_v^2\right] < \infty,\\
&\operatorname{ess\,sup}_{x\in\mathbb{R}^n} \sup_{v\in S^{n-1}} \mathbb{E}_{t_1,\infty}\left[\int_{t_1}^{t_2} \left|\frac{\mathrm{d}}{\mathrm{d}x} Z_s\right|_v^2 \mathrm{d}s\right] < \infty.
\end{aligned} \tag{8}$$

Under suitable conditions a rich existence, uniqueness and regularity theory for decoupling fields can be developed. The basis of the theory is Theorem 1.3 below, which is proven in Chapter 2 of [Fro15].

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. $\xi \colon \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ is measurable such that $\|\xi(\cdot, 0)\|_\infty < \infty$ and $L_{\xi,x} < L_{\sigma,z}^{-1}$.

Theorem 1.3 ([Fro15], 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 [Fro15].

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 [Fro15]). In order to have a notion of global existence we need the following definition:

Definition 1.4. 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 1.5 (Global existence in weak form, [Fro15], 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* $0 \le t_{\min} < T$*. In the latter case we have* $\lim_{t\downarrow t_{\min}} L_{u(t,\cdot),x} = L_{\sigma,z}^{-1}$.

2 Solving FBSDE by minimizing a continuous error functional

The purpose of this section is to introduce and study an error functional on a set of Itô processes. Our aim is to show that minimizing the error functional is in some sense equivalent to minimizing the distance to the actual solution of the FBSDE. While the distance between an approximation and the actual solution is usually not known, because the actual solution is not explicitly known, the error functional can still be evaluated for all admissible approximations. This allows to construct approximations close to the actual solution without explicitly knowing it.

Before introducing the error functional we fix some notations and show some basic estimates. Throughout we suppose that $(\xi, (\mu, \sigma, f))$ satisfy (SLC).

Let $I = [a,b] \subseteq [0,T]$. Define $\mathcal{L}(I)$ as the Banach space of all progressively measurable $\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d}$ - valued processes (X,Y,Z) on $I \times \Omega$ such that

$$\|(X,Y,Z)\|_\infty := \max\Bigg(\sup_{s\in I} \sqrt{\mathbb{E}[|X_s|^2]},\, (1+L_{\sigma,z}) \sup_{s\in I} \sqrt{\mathbb{E}[|Y_s|^2]},$$
$$(1+L_{\sigma,z})\sqrt{\mathbb{E}\left[\int_a^b |Z_s|^2 \,\mathrm{d}s\right]}\Bigg) < \infty. \tag{9}$$

Next, we define $\mathcal{M}(I)$ as the Banach space of all progressively measurable $\mathbb{R}^n \times \mathbb{R}^m$ - valued processes (X,Y) on $I \times \Omega$ such that X, Y can be written in the form

$$X_s = X_a + \int_a^s DX_r \,\mathrm{d}r + \int_a^s \mathfrak{D}X_r \,\mathrm{d}W_r,$$
$$Y_s = Y_a + \int_a^s DY_r \,\mathrm{d}r + \int_a^s \mathfrak{D}Y_r \,\mathrm{d}W_r,$$

a.s. for all $s \in I$ and such that

$$\|(X,Y)\|_1 := \max\Bigg\{ \left(\mathbb{E}\left[|X_a|^2\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[|I| \int_a^b |DX_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\int_a^b |\mathfrak{D}X_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}},$$
$$(1+L_{\sigma,z})\left(\left(\mathbb{E}\left[|I| \int_a^b |DY_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[|Y_b|^2\right]\right)^{\frac{1}{2}}\right)\Bigg\} < \infty,$$

where $|I|$ denotes the length of the interval I, i.e. $|I| = b-a$.

Lemma 2.1. *For all $(X,Y) \in \mathcal{M}(I)$ we have $(X,Y,\mathfrak{D}Y) \in \mathcal{L}(I)$ and $\|(X,Y,\mathfrak{D}Y)\|_\infty \leq \|(X,Y)\|_1$.*

Proof. Firstly note that using the Minkowski and the Cauchy-Schwarz inequalities we get

$$\sqrt{\mathbb{E}[|X_s|^2]} \leq \left(\mathbb{E}\left[|X_a|^2\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[(s-a)\int_a^s |DX_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\int_a^s |\mathfrak{D}X_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}},$$

for all $s \in I$. This value does not exceed $\|(X,Y)\|_1$. Secondly, note that for all $s \in I$

$$Y_s + \int_s^b \mathfrak{D}Y_r \,\mathrm{d}W_r = Y_b - \int_s^b DY_r \,\mathrm{d}r.$$

The orthogonality of Y_s and $\int_s^b \mathfrak{D}Y_r \,\mathrm{d}W_r$, together with the Minkowski and the Cauchy-Schwarz inequalities as well as the Itô isometry, imply

$$\left(\mathbb{E}[|Y_s|^2] + \mathbb{E}\left[\int_s^b |\mathfrak{D}Y_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} \leq \left(\mathbb{E}\left[|Y_b|^2\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[(b-s)\int_s^b |DY_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}}.$$

This shows

$$(1+L_{\sigma,z}) \max\left(\sup_{s\in I} \sqrt{\mathbb{E}[|Y_s|^2]},\, \sqrt{\mathbb{E}\left[\int_a^b |\mathfrak{D}Y_s|^2 \,\mathrm{d}s\right]}\right) \leq \|(X,Y)\|_1.$$

□

For given $(X,Y) \in \mathcal{M}(I)$ the process $Z := \mathfrak{D}Y$ is well-defined and is such that $(X,Y,Z) \in \mathcal{L}(I)$. Thus, $\mathcal{M}(I)$ can be canonically interpreted as a subspace of $\mathcal{L}(I)$.

Now let $x \in \mathbb{R}^n$. Moreover, let $\zeta : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ be measurable w.r.t. $\mathcal{F}_b \otimes \mathcal{B}(\mathbb{R}^n)$ and Lipschitz continuous w.r.t. x such that $L_{\zeta,x} < L_{\sigma,z}^{-1}$, where $L_{\zeta,x}$ denotes the Lipschitz constant of ζ w.r.t. x. For any $(X^0,Y^0,Z^0) \in \mathcal{L}(I)$ there is a unique $(X^1,Y^1,Z^1) = F_{x,I,\zeta}(X^0,Y^0,Z^0) \in \mathcal{L}(I)$ s.t.

$$X_s^1 := x + \int_a^s \mu(r, X_r^0, Y_r^0, Z_r^0)\,\mathrm{d}r + \int_a^s \sigma(r, X_r^0, Y_r^0, Z_r^0)\,\mathrm{d}W_r,$$

$$Y_s^1 := \zeta(X_b^1) - \int_s^b f(r, X_r^1, Y_r^0, Z_r^0)\,\mathrm{d}r - \int_s^b (Z_r^1)\,\mathrm{d}W_r,$$

a.s. for all $s \in I$. This is a well-known consequence of the martingale representation theorem. This defines a mapping $F_{x,I,\zeta} : \mathcal{L}(I) \to \mathcal{L}(I)$. We simply write F instead of $F_{x,I,\zeta}$ in case there is no ambiguity. It can be shown that F is a contraction w.r.t. $\|\cdot\|_\infty$, if $|I| := b - a$ is sufficiently small. In this case, the fixed point (X^*,Y^*,Z^*) of this mapping is the solution of the FBSDE on I for the terminal condition ζ and the initial value x.

It is straightforward to show that for $(X,Y) \in \mathcal{M}(I)$ the triplet $(\hat{X},\hat{Y},\hat{Z}) := F(X,Y,Z)$ is such that $(\hat{X},\hat{Y}) \in \mathcal{M}(I)$ while $\hat{Z} = \mathfrak{D}\hat{Y}$. This defines a mapping on $\mathcal{M}(I)$ to itself, which can be interpreted as the restriction of F to the space $\mathcal{M}(I)$. By a slight abuse of notation we denote the restricted mapping again by $F = F_{x,I,\zeta} : \mathcal{M}(I) \to \mathcal{M}(I)$.

Finally, we define the *error functional* $G_{x,I,\zeta} : \mathcal{M}(I) \to [0,\infty)$ via

$$G_{x,I,\zeta}(X,Y) := \mathbb{E}\Bigg[|X_a - x|^2 + |I| \cdot \int_a^b |DX_r - \mu(r,\Theta_r)|^2\,\mathrm{d}r + \int_a^b |\mathfrak{D}X_r - \sigma(r,\Theta_r)|^2\,\mathrm{d}r +$$

$$+ |I| \cdot \int_a^b |DY_r - f(r,\Theta_r)|^2\,\mathrm{d}r + |Y_b - \zeta(X_b)|^2\Bigg],$$

where Θ_r stands for $(X_r, Y_r, \mathfrak{D}Y_r)$. In cases of no ambiguity we simply write $G(X,Y)$ instead of $G_{x,I,\zeta}(X,Y)$. Notice that, applied to the processes (3) and (4), the error functional $G_{x,I,\zeta}$ coincides with the function G defined in (5).

Similar to the definition of $L_{\sigma,z}$ we define the constants $L_{f,x}$, $L_{f,y}$, $L_{f,z}$ as the minimal Lipschitz constants w.r.t. the applicable Euclidian/Frobenius norms. We further define $L_f := L_{f,x} \vee L_{f,y} \vee L_{f,z}$. Similarly, we define L_μ and L_σ.

We next estimate a norm of the difference $(X,Y) - F(X,Y)$ against $G(X,Y)$ and vice versa. To this end we define two constants

$$C_1 := \max\left\{3, (1 + L_{\sigma,z})\,(2 + 3\,(L_f|I| + L_{\zeta,x}))\right\} \text{ and } C_2 := \sqrt{2\left(2 + (L_f|I|)^2 + L_{\zeta,x}^2\right)}.$$

Proposition 2.2. *On the one hand we have*

$$\|(X,Y) - F(X,Y)\|_1 \le C_1\sqrt{G(X,Y)}$$

and on the other hand

$$\sqrt{G(X,Y)} \le C_2\|(X,Y) - F(X,Y)\|_1$$

for all $(X,Y) \in \mathcal{M}(I)$.

Proof. Using the definition of F we have

$$\|(X,Y)-F(X,Y)\|_1 =$$

$$\max\Bigg\{ \left(\mathbb{E}\left[|X_a-x|^2\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[|I|\int_a^b |DX_r-\mu(r,\Theta_r)|^2\,\mathrm{d}r\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\int_a^b |\mathfrak{D}X_r-\sigma(r,\Theta_r)|^2\,\mathrm{d}r\right]\right)^{\frac{1}{2}},$$

$$(1+L_{\sigma,z})\cdot\left(\left(\mathbb{E}\left[|I|\int_a^b |DY_r-f(r,\check{X}_r,Y_r,Z_r)|^2\,\mathrm{d}r\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[|Y_b-\zeta(\check{X}_b)|^2\right]\right)^{\frac{1}{2}}\right)\Bigg\},$$

where $\check{X}$ is the first component of $F(X,Y)$. Clearly, the three summands in the first argument of max can be bounded by $\sqrt{G(X,Y)}$ each. The second argument in max could be bounded by $2(1+L_{\sigma,z})\sqrt{G(X,Y)}$ if we were to replace $\check{X}$ by X in it. The error caused by doing so would not exceed

$$(1+L_{\sigma,z})\left(L_f|I|+L_{\zeta,x}\right)\sup_{s\in[a,b]}\sqrt{\mathbb{E}[|X_s-\check{X}_s|^2]},$$

where L_f is a Lipschitz constant for f. Now note that $\sqrt{\mathbb{E}[|X_s-\check{X}_s|^2]}$ is controlled by

$$\sqrt{\mathbb{E}\left[|X_a-x|^2\right]} + \left(\mathbb{E}\left[|I|\int_a^b |DX_r-\mu(r,\Theta_r)|^2\,\mathrm{d}r\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\int_a^b |\mathfrak{D}X_r-\sigma(r,\Theta_r)|^2\,\mathrm{d}r\right]\right)^{\frac{1}{2}},$$

via the Minkowski inequality. This expression, however, is bounded by $3\sqrt{G(X,Y)}$ as already mentioned. Therefore, we have

$$\|(X,Y)-F(X,Y)\|_1 \le \max\left\{3, 2(1+L_{\sigma,z})+(1+L_{\sigma,z})\left(L_f|I|+L_{\zeta,x}\right)\cdot 3\right\}\sqrt{G(X,Y)}.$$

At the same time $G(X,Y)$ would be bounded by $2\|(X,Y)-F(X,Y)\|_1^2$ if we were to replace X by $\check{X}$ in $f(r,\Theta_r)$ and in $\zeta(X_b)$ in the definition of G. However, we still have

$$G(X,Y)\le 4\|(X,Y)-F(X,Y)\|_1^2 + 2((L_f|I|)^2+L_{\zeta,x}^2)\sup_{s\in[a,b]}\mathbb{E}[|X_s-\check{X}_s|^2],$$

which, due to the above considerations, is controlled by

$$2\left(2+(L_f|I|)^2+L_{\zeta,x}^2\right)\|(X,Y)-F(X,Y)\|_1^2.$$

□

We next estimate $\|(X,Y)-F(X,Y)\|_1$ against $\|(X,Y)-(X^*,Y^*)\|_1$, where (X^*,Y^*) denotes the solution of the FBSDE. For this purpose we define

$$C_3 := \frac{1}{1-\kappa},\quad C_4 := 1+\kappa,\quad \kappa=\kappa(|I|,L_{\zeta,x}) := \sqrt{|I|}\cdot C_5 + \frac{L_{\sigma,z}}{1+L_{\sigma,z}} \vee \left(L_{\zeta,x}L_{\sigma,z}\right), \tag{10}$$

where

$$C_5 := \max\Big\{\left(2\sqrt{|I|}+1\right)L_\mu + 2L_\sigma,$$

$$\left(2\sqrt{|I|}+1\right)L_f + (1+L_{\sigma,z})\left(L_f|I|+L_{\zeta,x}\right)\left(\left(2\sqrt{|I|}+1\right)L_\mu+2L_\sigma\right) + L_{\sigma,z}L_f\sqrt{I}\Big\}.$$

Note that C_5 depends only on $|I|$, $L_{\zeta,x}$ and the Lipschitz constants of μ, σ, f and is monotonically increasing in these values.

In the following we say that $|I|$ is sufficiently small if $|I|$ satisfies $\kappa(|I|,L_{\zeta,x})<1$.

Proposition 2.3. *F is a contraction on $\mathcal{M}(I)$ w.r.t. $\|\cdot\|_1$, if $|I|$ is sufficiently small. In addition, for sufficiently small $|I|$ we have $\|(X^*,Y^*)-(X,Y)\|_1 \le C_3\|(X,Y)-F(X,Y)\|_1$ and $\|(X,Y)-F(X,Y)\|_1 \le C_4\|(X^*,Y^*)-(X,Y)\|_1$ for all $(X,Y)\in\mathcal{M}(I)$.*

Proof. Let (X,Y), $(X',Y')\in\mathcal{M}(I)$, $Z=\mathfrak{D}Y$ and $Z'=\mathfrak{D}Y'$. Then we have

$$\|F(X,Y)-F(X',Y')\|_1 =$$

$$= \max\Bigg\{ \left(\mathbb{E}\left[|I|\int_a^b |\mu(r,\Theta_r)-\mu(r,\Theta'_r)|^2\,\mathrm{d}r\right]\right)^{\frac12} + \left(\mathbb{E}\left[\int_a^b |\sigma(r,\Theta_r)-\sigma(r,\Theta'_r)|^2\,\mathrm{d}r\right]\right)^{\frac12},$$

$$(1+L_{\sigma,z})\left(\left(\mathbb{E}\left[|I|\int_a^b |f(r,\check\Theta_r)-f(r,\check\Theta'_r)|^2\,\mathrm{d}r\right]\right)^{\frac12} + \left(\mathbb{E}\left[|\zeta(\check X_b)-\zeta(\check X'_b)|^2\right]\right)^{\frac12}\right)\Bigg\},$$

where $\check\Theta$ refers to $(\check X,Y,Z)$ and $\check\Theta'$ refers to $(\check X',Y',Z')$ respectively. Let L_μ, L_σ, L_f be Lipschitz constants for μ, σ and f. Remembering the definition of $\|\cdot\|_\infty$ we have:

$$\|F(X,Y)-F(X',Y')\|_1 \le \max\Bigg\{ \left(2|I|+\sqrt{|I|}\right)\cdot L_\mu\|(X,Y,Z)-(X',Y',Z')\|_\infty +$$

$$+2\sqrt{|I|}\cdot L_\sigma\|(X,Y,Z)-(X',Y',Z')\|_\infty + \frac{L_{\sigma,z}}{1+L_{\sigma,z}}\|(X,Y,Z)-(X',Y',Z')\|_\infty,$$

$$(2|I|+\sqrt{|I|})\cdot L_f\|(X,Y,Z)-(X',Y',Z')\|_\infty + (1+L_{\sigma,z})(L_f|I|+L_{\zeta,x})\sqrt{\mathbb{E}[|\check X_s-\check X'_s|^2]}\Bigg\}.$$

Note that, using the Minkowski inequality, the value $\sqrt{\mathbb{E}[|\check X_s-\check X'_s|^2]}$ is bounded by

$$\left(\mathbb{E}\left[|I|\int_a^b |\mu(r,\Theta_r)-\mu(r,\Theta'_r)|^2\,\mathrm{d}r\right]\right)^{\frac12} + \left(\mathbb{E}\left[\int_a^b |\sigma(r,\Theta_r)-\sigma(r,\Theta'_r)|^2\,\mathrm{d}r\right]\right)^{\frac12} \le$$

$$\left(\left(2|I|+\sqrt{|I|}\right)\cdot L_\mu + 2\sqrt{|I|}\cdot L_\sigma + \frac{L_{\sigma,z}}{1+L_{\sigma,z}}\right)\|(X,Y,Z)-(X',Y',Z')\|_\infty.$$

Note that the above constant in front of $\|(X,Y,Z)-(X',Y',Z')\|_\infty$ can be controlled by 1 for $|I|$ sufficiently small. Putting everything together allows us to conclude that for the constant C_5 and sufficiently small $|I|$

$$\|F(X,Y)-F(X',Y')\|_1 \le \left(\sqrt{|I|}\cdot C_5 + \frac{L_{\sigma,z}}{1+L_{\sigma,z}} \vee (L_{\zeta,x}L_{\sigma,z})\right)\|(X,Y,Z)-(X',Y',Z')\|_\infty \le$$

$$\le \left(\sqrt{|I|}\cdot C_5 + \frac{L_{\sigma,z}}{1+L_{\sigma,z}} \vee (L_{\zeta,x}L_{\sigma,z})\right)\|(X,Y)-(X',Y')\|_1.$$

Therefore, F is a contraction under $\|\cdot\|_1$ if $|I|$ is sufficiently small enough. Furthermore,

$$\|(X,Y)-F(X,Y)\|_1 \le \|(X,Y)-(X^*,Y^*)\|_1 + \|F(X^*,Y^*)-F(X,Y)\|_1 \le$$

$$\le (1+\kappa)\|(X,Y)-(X^*,Y^*)\|_1.$$

At the same time

$$\|(X,Y)-(X^*,Y^*)\|_1 \le \|(X,Y)-F(X,Y)\|_1 + \|F(X,Y)-(X^*,Y^*)\|_1 \le$$
$$\le \|(X,Y)-F(X,Y)\|_1 + \kappa\|(X,Y)-(X^*,Y^*)\|_1.$$

Therefore, we have shown

$$(1-\kappa)\|(X,Y)-(X^*,Y^*)\|_1 \le \|(X,Y)-F(X,Y)\|_1 \le (1+\kappa)\|(X,Y)-(X^*,Y^*)\|_1,$$

which yields the claim. □

By combining Propositions 2.3 and 2.2 we obtain the main result of this section.

Theorem 2.4. *For sufficiently small I we have*

$$\|(X^*,Y^*)-(X,Y)\|_1 \le C_1C_3\sqrt{G(X,Y)},$$

and

$$\sqrt{G(X,Y)} \le C_2C_4\|(X^*,Y^*)-(X,Y)\|_1.$$

Theorem 2.4 means that seeking to minimize G is in a sense equivalent to minimizing the distance to the actual solution of the FBSDE. Put otherwise, if our objective is to choose the best approximation for (X^*,Y^*) from a given set of admissible approximations we may simply choose the approximation for which the functional G assumes the smallest value. Note that while (X^*,Y^*) is usually unknown (this is what we want to approximate) the functional G can be evaluated for any $(X,Y)\in\mathcal{M}(I)$ based on the input parameters x,μ,σ,f,ζ alone.

Let v be a weakly regular decoupling field on the interval I for the terminal condition ζ. As we are primarily focused on approximating $v(a,x)$ for a given $x\in\mathbb{R}^n$, the following statement is of importance:

Corollary 2.5. *Let $|I|$ be sufficiently small and let $(X,Y)\in\mathcal{M}(I)$. Then*

$$\sqrt{\mathbb{E}[|v(a,x)-Y_a|^2]} \le \frac{\max\left\{3, 2+3\left(L_f|I|+L_{\zeta,x}\right)\right\}}{1-\kappa(|I|,L_{\zeta,x})}\sqrt{G}(X,Y).$$

In particular, if $v(a,x)$ and Y_a are deterministic we obtain a control for $|v(a,x)-Y_a| = \sqrt{\mathbb{E}[|v(a,x)-Y_a|^2]}$.

Proof. According to the proofs of Propositions 2.3 and 2.2 we have

$$\|(X,Y)-(X^*,Y^*)\|_1 \le \frac{1}{1-\kappa}\|(X,Y)-F(X,Y)\|_1$$
$$\le \frac{1}{1-\kappa}\max\left\{3, (1+L_{\sigma,z})[2+3\left(L_f|I|+L_{\zeta,x}\right)]\right\}\sqrt{G(X,Y)}.$$

Also, note that

$$\sqrt{\mathbb{E}[|v(a,x)-Y_a|^2]} \le \frac{1}{1+L_{\sigma,z}}\|(X,Y,Z)-(X^*,Y^*,Z^*)\|_\infty \le \frac{1}{1+L_{\sigma,z}}\|(X,Y)-(X^*,Y^*)\|_1.$$

Combining the estimates yields the claim. □

Remark 2.6. Note that we can assume without any loss of generality that $L_{\sigma,z}$ and $L_{\zeta,x}$ both do not exceed 1:

This is because we can modify a given FBSDE by multplying the forward equation by a parameter $\lambda > 0$, thus obtaining the process $\tilde{X} := \lambda X$ and then replace all occurrences of X in the FBSDE by $\frac{1}{\lambda}\tilde{X}$. In particular we obtain a different terminal condition $\tilde{\zeta}$ defined via $\tilde{\zeta}(x') := \zeta(\lambda^{-1}x')$. We also obtain different $\tilde{\mu}, \tilde{\sigma}, \tilde{f}$ and, therefore, a different target functional G^λ. Note that $L_{\tilde{\zeta},x} = \frac{1}{\lambda}L_{\zeta,x}$ and $L_{\tilde{\sigma},z} = \lambda L_{\sigma,z}$. However, $L_{\tilde{\zeta},x}L_{\tilde{\sigma},z} = L_{\zeta,x}L_{\sigma,z} < 1$ regardless of λ.

Now if $L_{\zeta,x} > 1$ we can set $\lambda := L_{\zeta,x}$, thus making sure that $L_{\tilde{\zeta},x} = 1$, while $L_{\tilde{\sigma},z} = L_{\zeta,x}L_{\sigma,z} < 1$. If, however, $L_{\sigma,z} > 1$ we can set $\lambda := \frac{1}{L_{\sigma,z}}$, thereby making sure that $L_{\tilde{\sigma},z} = 1$, while $L_{\tilde{\zeta},x} = L_{\zeta,x}L_{\sigma,z} < 1$.

Note that choosing an appropriate λ can considerably sharpen the estimate of Corollary 2.5 by preventing $L_{\zeta,x}$ or $L_{\sigma,z}$ from becoming large. The choice of λ effects how the backward and the forward equation are weighted inside the error functional G^λ. One should aim at avoiding error functionals in which too much emphasis is put on fulfilling the forward or the backward equation while the other is somewhat neglected.

3 Approximation errors for stochastic polynomials

In this section we derive error estimates for the minimizers of the error functional $G_{x,I,\zeta}$ in a set of stochastic polynomials. To this end we need to introduce finite dimensional spaces of stochastic polynomials.

Let $L^2(I)$ be the linear space of all real-valued square-integrable progressive processes X defined on $I := [a,b] \subseteq [0,T]$ equipped with the norm $\|\cdot\|_{L^2}$ given by

$$\|X\|_{L^2(I)} := \left(\int_a^b \mathbb{E}[|X_s|^2]\,\mathrm{d}s\right)^{\frac{1}{2}}.$$

$\mathcal{P}(I)$ is defined as the smallest subset of $L^2(I)$ satisfying

1. $\mathcal{P}(I)$ is an $\mathbb{R}$ - vector space, which contains all constants;
2. if $X \in \mathcal{P}(I)$, then $s \mapsto \int_a^s X_r\,\mathrm{d}r$ is also in $\mathcal{P}(I)$;
3. if $X \in \mathcal{P}(I)$, then $s \mapsto \int_a^s X_r\,\mathrm{d}W_r^i$ is also in $\mathcal{P}(I)$ for every $i = 1, \ldots, d$.

In other words, $\mathcal{P}(I)$ is the smallest space which is closed under integration w.r.t. time or the Brownian motion W. In case of no ambiguity we simply write $\mathcal{P}$ instead of $\mathcal{P}(I)$.

Moreover, let $L^{2,\infty}(I)$ be the linear space of all $X \in L^2(I)$ such that

$$\|X\|_{L^{2,\infty}(I)} := \left(\sup_{s\in[a,b]} \mathbb{E}[|X_s|^2]\right)^{\frac{1}{2}} < \infty.$$

For all real-valued Itô processes X, which can be written in the form

$$X_s = X_a + \int_a^s DX_r\,\mathrm{d}r + \int_a^s \mathfrak{D}X_r\,\mathrm{d}W_r,$$

where X_a is a constant, we define

$$\|X\|_0 := |X_a| + \left(\mathbb{E}\left[|I| \int_a^b |DX_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\int_a^b |\mathfrak{D}X_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}}.$$

We refer to $\mathbb{H}^0(I)$ as the set of all such Itô processes X satisfying $\|X\|_0 < \infty$. It is straightforward to show, using the Minkowski and the Cauchy-Schwarz inequalities together with the Itô isometry, that $\mathbb{H}^0 \subseteq L^{2,\infty}(I)$ and $\|X\|_{L^{2,\infty}(I)} \leq \|X\|_0$.

For every $z \in \mathcal{Z} := \left\{\frac{n}{2} \,|\, n \in \mathbb{N}_1\right\}$ we define the spaces $\mathbb{H}^z(I) \subseteq \mathbb{H}^0(I)$ recursively as follows:

- $\mathbb{H}^{\frac{1}{2}}$ is the space of all $X \in \mathbb{H}^0$ such that $DX \in L^2$ and $\mathfrak{D}X^i \in L^{2,\infty}$ for all $i = 1, \ldots, d$.
- $\mathbb{H}^1$ is the space of all $X \in \mathbb{H}^0$ such that $DX \in L^{2,\infty}$ and $\mathfrak{D}X^i \in \mathbb{H}^{\frac{1}{2}}$ for all $i = 1, \ldots, d$.
- For $z \geq \frac{3}{2}$ we define $\mathbb{H}^z$ as the space of all $X \in \mathbb{H}^0$ such that $DX \in \mathbb{H}^{z-1}$ and $\mathfrak{D}X^i \in \mathbb{H}^{z-\frac{1}{2}}$ for all $i = 1, \ldots, d$.

Furthermore, we equip the spaces $\mathbb{H}^z$ with the following semi-norms:

- For $X \in \mathbb{H}^{\frac{1}{2}}$ we set $\|X\|_{\frac{1}{2}} := \|DX\|_{L^2} \vee \max\left\{\|\mathfrak{D}X^i\|_{L^{2,\infty}} \,\middle|\, i = 1, \ldots, d\right\}$.
- For $X \in \mathbb{H}^1$ we set $\|X\|_1 := \|DX\|_{L^{2,\infty}} \vee \max\left\{\|\mathfrak{D}X^i\|_{\frac{1}{2}} \,\middle|\, i = 1, \ldots, d\right\}$.
- For $X \in \mathbb{H}^z$, $z \geq \frac{3}{2}$ we set $\|X\|_z := \|DX\|_{z-1} \vee \max\left\{\|\mathfrak{D}X^i\|_{z-\frac{1}{2}} \,\middle|\, i = 1, \ldots, d\right\}$.

Finally, we define $\mathcal{P}^z$, where $z \in \mathcal{Z}$, as the set of all $X \in \mathbb{H}^z$ such that $\|X\|_z = 0$.

Lemma 3.1. *$\mathcal{P}^z \subseteq \mathcal{P}$ for all $z \in \mathcal{Z}$.*

Proof. This is straightforward to show using induction over z:

- For $z = \frac{1}{2}$ the equation $\|X\|_z = 0$ implies that X is constant.
- For $z = 1$ the equation $\|X\|_z = 0$ implies that DX vanishes and $\mathfrak{D}X^i \in \mathcal{P}^{\frac{1}{2}}$. The latter means that $\mathfrak{D}X$ is a constant process. Therefore, X is a stochastic polynomial.
- Let $z \in \mathcal{Z}$ satisfy $z \geq \frac{3}{2}$ and assume that for all smaller z the statement is already proven. Then $\|X\|_z = 0$ implies that $DX \in \mathcal{P}^{z-1}$ and $\mathfrak{D}X^i \in \mathcal{P}^{z-\frac{1}{2}}$. So, DX and $\mathfrak{D}X^i$ are stochastic polynomials for all i and, therefore, X is a stochastic polynomial as well.

□

One can also show using the definition of $\mathcal{P}$ that $\mathcal{P} \subseteq \bigcup_{z \in \mathcal{Z}} \mathcal{P}^z$ as the set $\bigcup_{z \in \mathcal{Z}} \mathcal{P}^z$ satisfies the three properties which characterize $\mathcal{P}$. Furthermore, one can show $\mathcal{P}^z \subseteq \mathcal{P}^{z+\frac{1}{2}}$ for all $z \in \mathcal{Z}$, using induction over z.

We call the elements of $\mathcal{P}^z \backslash \mathcal{P}^{z-\frac{1}{2}}$ *stochastic polynomials of order z*, where we set $\mathcal{P}^0 := \{0\}$. Clearly, every non-vanishing stochastic polynomial has a uniquely defined order.

Proposition 3.2. *For every $z \in \mathcal{Z}$ there exists a constant $C_{z,d}$, which depends on z and d only, such that for every $X \in \mathbb{H}^z(I)$ there exists an $A \in \mathcal{P}^z(I)$ such that*

$$\sqrt{\mathbb{E}[|X_s - A_s|^2]} \leq C_{z,d} \cdot (s-a)^z \cdot \|X\|_z,$$

for all $s \in [a,b]$.
Furthermore, one can choose $C_{z,d}$ such that $C_{z,d} \leq \frac{2\left(1+\sqrt{d}\right)^{2z-1}}{\sqrt{(2z)!}}$ for all z.

Proof. Let $z = \frac{1}{2}$. Then we can set $A = X_a$ and use $C_{\frac{1}{2},d} = 2$. Indeed, in this case

$$\sqrt{\mathbb{E}[|X_s - X_a|^2]} \leq \left(\mathbb{E}\left[(s-a)\int_a^s |DX_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\int_a^s |\mathfrak{D}X_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} \leq$$

$$\leq (s-a)^{\frac{1}{2}}\|X\|_z + (s-a)^{\frac{1}{2}}\|X\|_z.$$

Now let $z = 1$. Then we define A via $A_a := X_a$, $DA = 0$ and $\mathfrak{D}A := \mathfrak{D}X_a$. Then

$$\sqrt{\mathbb{E}[|X_s - A_s|^2]} \leq \left(\mathbb{E}\left[(s-a)\int_a^s |DX_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} + \left(\mathbb{E}\left[\int_a^s |\mathfrak{D}X_r - \mathfrak{D}X_a|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} \leq$$

$$\leq (s-a)^{\frac{1}{2}}(s-a)^{\frac{1}{2}}\|X\|_z + \left(\int_a^s 4d(r-a)\|X\|_z^2 \,\mathrm{d}r\right)^{\frac{1}{2}} = (s-a)\|X\|_z + \sqrt{2d}(s-a)\|X\|_z,$$

so we can set $C_{1,d} := 1 + \sqrt{2d}$.

Finally, let $z \geq \frac{3}{2}$ and assume that the statement is true for all lower z. Let $X \in \mathbb{H}^z$ and define A by setting $A_a := X_a$ and choosing DA, $\mathfrak{D}A$ such that

$$\sqrt{\mathbb{E}[|DX_s - DA_s|^2]} \leq C_{z-1,d} \cdot (s-a)^{z-1} \cdot \|DX\|_{z-1},$$

$$\sqrt{\mathbb{E}[|\mathfrak{D}X_s^i - \mathfrak{D}A_s^i|^2]} \leq C_{z-\frac{1}{2},d} \cdot (s-a)^{z-\frac{1}{2}} \cdot \|\mathfrak{D}X^i\|_{z-\frac{1}{2}}.$$

The Minkowski inequality together with the Itô isometry imply

$$\begin{aligned}
\sqrt{\mathbb{E}[|X_s - A_s|^2]} &\leq \int_a^s \left(\mathbb{E}\left[|DX_r - DA_r|^2\right]\right)^{\frac{1}{2}} \,\mathrm{d}r + \left(\mathbb{E}\left[\int_a^s |\mathfrak{D}X_r - \mathfrak{D}A_r|^2 \,\mathrm{d}r\right]\right)^{\frac{1}{2}} \\
&\leq C_{z-1,d}\frac{1}{z}(s-a)^z\|DX\|_{z-1} + \left(\int_a^s C_{z-\frac{1}{2},d}^2 (r-a)^{2z-1} \cdot d \cdot \|X\|_z^2 \,\mathrm{d}r\right)^{\frac{1}{2}} \\
&= C_{z-1,d}\frac{1}{z}(s-a)^z\|X\|_z + C_{z-\frac{1}{2},d}\frac{\sqrt{d}}{\sqrt{2z}}(s-a)^z\|X\|_z,
\end{aligned}$$

so we can set $C_{z,d} := C_{z-1,d}\frac{1}{z} + C_{z-\frac{1}{2},d}\frac{\sqrt{d}}{\sqrt{2z}}$.
Note, that if we define $C_{z,d}$ recursively as above, one can show using induction that

$$C_{z,d} \leq \frac{2(1+\sqrt{2d})^{2z-1}}{\sqrt{(2z)!}}.$$

□

The above approximation result is straightforward to generalize to multi-dimensional processes X by applying it component-wise. Note that the definition of the semi-norms $\|\cdot\|_z$ is canonically extended to multi-dimensional processes using the standard definitions for $\|X\|_{L^2(I)}$ and $\|X\|_{L^{2,\infty}(I)}$ for vector-valued X. We further define $\mathcal{P}_n^z(I) := (\mathcal{P}^z(I))^n$. This finite dimensional linear space is used for the admissible approximations of $\mathbb{R}^n$ - valued X.

For the next result we use the notations of the previous section.

Theorem 3.3. *Let $x \in \mathbb{R}^n$ and let $\zeta : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ be measurable w.r.t. $\mathcal{F}_b \otimes \mathcal{B}(\mathbb{R}^n)$ such that $L_{\zeta,x} < L_{\sigma,z}^{-1}$. Suppose that I is sufficiently small and let $(\hat{X}, \hat{Y})$ be a minimizer of $G_{x,I,\zeta}$ in $\mathcal{P}_n^{\frac{3}{2}}(I) \times \mathcal{P}_m^{\frac{3}{2}}(I)$. Then*

$$\sqrt{G_{x,I,\zeta}(\hat{X}, \hat{Y})} \leq C_2 C_4 K(\|X^*\|_{\frac{3}{2}}, \|Y^*\|_{\frac{3}{2}})\, |I|^{\frac{3}{2}},$$

where

$$K(\bar{x}, \bar{y}) := \max\left\{ \left(\frac{1}{2} C_{\frac{1}{2},d} + 3^{-\frac{1}{2}} C_{1,d}\right) \bar{x}, (1 + L_{\sigma,z}) \left(\frac{1}{2} C_{\frac{1}{2},d} + \frac{2}{\sqrt{3}} C_{1,d}\right) \bar{y} \right\}.$$

Proof. Let (X^*, Y^*, Z^*) be a solution of the FBSDE on I with initial condition x and terminal condition ζ. We apply Proposition 3.2 to the Ito process coefficients DX^*, $\mathfrak{D}X^*$, DY^* and $\mathfrak{D}Y^*$, respectively. Let $U \in \mathcal{P}_n^{\frac{1}{2}}(I)$ such that for all $s \in [a, b]$ we have

$$\sqrt{E(|DX_s^* - U_s|^2)} \leq C_{\frac{1}{2},d} \sqrt{s - a} \|DX^*\|_{\frac{1}{2}}.$$

Then we further obtain

$$\begin{aligned}
\left(E\left[|I| \int_a^b |DX_s^* - U_s|^2 ds\right]\right)^{\frac{1}{2}} &= |I|^{\frac{1}{2}} \left(\int_a^b E|DX_s^* - U_s|^2 ds\right)^{\frac{1}{2}} \\
&\leq |I|^{\frac{1}{2}} C_{\frac{1}{2},d} \|DX^*\|_{\frac{1}{2}} \left(\int_a^b (s - a) ds\right)^{\frac{1}{2}} \\
&= \frac{1}{2} C_{\frac{1}{2},d} \|DX^*\|_{\frac{1}{2}} |I|^{\frac{3}{2}}.
\end{aligned}$$

Let $V \in \mathcal{P}_{n\times d}^1(I)$ such that for all $s \in [a, b]$ we have

$$\sqrt{E(|\mathfrak{D}X_s^* - V_s|^2)} \leq C_{1,d}(s - a) \|\mathfrak{D}X^*\|_1.$$

This further yields

$$\begin{aligned}
\left(E\left[\int_a^b |\mathfrak{D}X_s^* - V_s|^2 ds\right]\right)^{\frac{1}{2}} &= \left(\int_a^b E|\mathfrak{D}X_s^* - V_s|^2 ds\right)^{\frac{1}{2}} \\
&\leq C_{1,d} \|\mathfrak{D}X^*\|_1 \left(\int_a^b (s - a)^2 ds\right)^{\frac{1}{2}} = 3^{-\frac{1}{2}} C_{1,d} \|\mathfrak{D}X^*\|_1 |I|^{\frac{3}{2}}.
\end{aligned}$$

Similarly, we choose $R \in \mathcal{P}_m^{\frac{1}{2}}(I)$ and $Q \in \mathcal{P}_{m\times d}^1(I)$ such that

$$\left(E\left[|I| \int_a^b |DY_s^* - R_s|^2 ds\right]\right)^{\frac{1}{2}} \leq \frac{1}{2} C_{\frac{1}{2},d} \|DY^*\|_{\frac{1}{2}} |I|^{\frac{3}{2}},$$

$$\left(E\left[\int_a^b |\mathfrak{D}Y_s^* - Q_s|^2 ds\right]\right)^{\frac{1}{2}} \leq 3^{-\frac{1}{2}} C_{1,d} \|\mathfrak{D}Y^*\|_1 |I|^{\frac{3}{2}}.$$

Define $X_t = x + \int_a^t U_s ds + \int_a^t V_s dW_s$ and $Y_t = Y_a^* + \int_a^t R_s ds + \int_a^t Q_s dW_s$. Notice that X and Y are Itô processes in $\mathcal{P}_n^{\frac{3}{2}}$ and $\mathcal{P}_m^{\frac{3}{2}}$ respectively.

The Minkowski and the Cauchy-Schwarz inequlities entail

$$\begin{aligned}
\left(E(Y_b^* - Y_b)^2\right)^{\frac{1}{2}} &\leq \left(E\left[|I|\int_a^b |DY_s^* - R_s|^2 ds\right]\right)^{\frac{1}{2}} + \left(E\left[\int_a^b |\mathfrak{D}Y_s^* - Q_s|^2 ds\right]\right)^{\frac{1}{2}} \\
&\leq \left(\frac{1}{2}C_{\frac{1}{2},d}\|DY^*\|_{\frac{1}{2}} + 3^{-\frac{1}{2}}C_{1,d}\|\mathfrak{D}Y^*\|_1\right)|I|^{\frac{3}{2}} \\
&\leq \left(\frac{1}{2}C_{\frac{1}{2},d} + 3^{-\frac{1}{2}}C_{1,d}\right)\|Y^*\|_{\frac{3}{2}}|I|^{\frac{3}{2}}.
\end{aligned}$$

The estimates above imply

$$\|(X - X^*, Y - Y^*)\|_1 \leq K(\|X^*\|_{\frac{3}{2}}, \|Y^*\|_{\frac{3}{2}})|I|^{\frac{3}{2}}.$$

Since $(\hat{X}, \hat{Y})$ is a minimizer of $G_{x,I,\zeta}$ in $\mathcal{P}_n^{\frac{3}{2}}(I) \times \mathcal{P}_m^{\frac{3}{2}}(I)$ we have

$$\sqrt{G_{x,I,\zeta}(\hat{X}, \hat{Y})} \leq \sqrt{G_{x,I,\zeta}(X, Y)},$$

and consequently, with Theorem 2.4, we arrive at

$$\sqrt{G_{x,I,\zeta}(\hat{X}, \hat{Y})} \leq C_2C_4\|(X - X^*, Y - Y^*)\|_1 \leq C_2C_4K(\|X^*\|_{\frac{3}{2}}, \|Y^*\|_{\frac{3}{2}})|I|^{\frac{3}{2}}.$$

□

4 A multi-step scheme

On the basis of Theorems 2.4 and 3.3 we develop a scheme for approximating the decoupling field, and in particular the solution, of a fully coupled FBSDE. Let $(\xi, (\mu, \sigma, f))$ satisfy (SLC) and assume that there exists a weakly regular decoupling field u on the whole interval $[0, T]$. To simplify the presentation we assume in addition that $n = m = d = 1$ and also that we are in a Markovian setting, i.e. $(\xi, (\mu, \sigma, f))$ do not depend on $\omega \in \Omega$. This implies that the corresponding decoupling field u is deterministic as well.

Our main objective is to approximate the function $u(0, \cdot)$ as well as possible. An approximation v of $x \mapsto u(0, x)$ is obtained iteratively by first obtaining a suitable approximation of $x \mapsto u(t', x)$ for some $t' \in [0, T]$ close to T, using this approximation as a new terminal condition and repeating the process until 0 is reached. More precisely, let $N \in \mathbb{N}$ and consider an equidistant partition $0 = t_0 < t_1 < \ldots < t_N = T$ of the interval $[0, T]$ such that the interval length $h := \frac{T}{N}$ of each subinterval $I_j := [t_j, t_{j+1}]$, $j = 0, \ldots, N-1$, is sufficiently small, i.e. is such that $\kappa(h, c) < 1$, where κ is defined in (10).

Now fix j and assume that we have a sufficiently smooth approximation $\tilde{u}(t_{j+1}, \cdot) : \mathbb{R} \to \mathbb{R}$ of $u(t_{j+1}, \cdot)$ at time t_{j+1}. We wish to obtain a sufficiently smooth approximation $\tilde{u}(t_j, \cdot)$ of $u(t_j, \cdot)$ for the time t_j. To this end we consider the FBSDE on I_j for the terminal condition $\tilde{u}(t_{j+1}, \cdot)$. First, we construct an approximation point-wise: Choosing an arbitrary $x \in \mathbb{R}$ we wish to approximate $v(t_j, x)$, where v is the decoupling field on I_j for terminal condition $\tilde{u}(t_{j+1}, \cdot)$. We do so by minimizing the error functional $G_{x,I_j,\tilde{u}(t_{j+1},\cdot)}$. We choose $\mathcal{P}^{\frac{3}{2}} \times \mathcal{P}^{\frac{3}{2}}$ as

the space of admissible approximations. In other words, we seek to approximate the correct solution (X^*, Y^*) of the FBSDE with initial condition x via approximations of the form

$$X_s^\alpha = \alpha_1 + \alpha_2(s - t_j) + \alpha_3(W_s - W_{t_j}) + \alpha_4 \int_{t_j}^s \left(W_r - W_{t_j}\right) dW_r,$$

$$Y_s^\beta = \beta_1 + \beta_2(s - t_j) + \beta_3(W_s - W_{t_j}) + \beta_4 \int_{t_j}^s \left(W_r - W_{t_j}\right) dW_r,$$

a.s. for all $s \in I_j$, where $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ and $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)$ are real-valued vectors.

The one-step scheme essentially consists of choosing a vector $(\alpha^*, \beta^*) \in \mathbb{R}^8$ such that $G_{x,I_j,\tilde{u}(t_{j+1},\cdot)}(X^{\alpha^*}, Y^{\beta^*})$ is minimal. We comment on how to practically calculate such a minimizer in Remark 4.1 below. For now we simply assume that a global minimizer (α^*, β^*) of the functional $(\alpha, \beta) \mapsto G_{x,I_j,\tilde{u}(t_{j+1},\cdot)}(X^\alpha, Y^\beta)$ is available and is calculated for an arbitrary $x \in \mathbb{R}$. We define $w(t_j, x) := \beta_1$. This yields a function $w(t_j, \cdot) : \mathbb{R} \to \mathbb{R}$ which can be evaluated (or approximated) at any given point.

If I_j is sufficiently small in the sense of Section 2, then the distance between $w(t_j, \cdot)$ and $v(t_j, \cdot)$ in the supremum norm can be nicely controlled. In general one can not compute $w(t_j, x)$ for all $x \in \mathbb{R}$ in practice, but only finitely many. We assume, therefore, that $w(t_j, \cdot)$ itself is approximated with a sufficiently smooth function $\tilde{u}(t_j, \cdot)$. Such an approximation may be obtained e.g. via a smooth interpolation.

Repeating the above one-step scheme N times in total yields a sufficiently smooth approximation $\tilde{u}(0, \cdot)$ of $u(0, \cdot)$.

Remark 4.1. A key component of our scheme is the minimization of the error functional $G_{x,I,\zeta}$ on the finite dimensional space $\mathcal{P}^{\frac{3}{2}} \times \mathcal{P}^{\frac{3}{2}}$ for a fixed initial value $x \in \mathbb{R}$, where $I := [a, b] = I_j$, $\zeta := \tilde{u}(t_{j+1}, \cdot)$ for some $j \in \{0, \dots, N-1\}$. This means that we need to choose vectors $\alpha, \beta \in \mathbb{R}^4$ such that

$$\begin{aligned} G(\alpha, \beta) &:= G_{x,I_j,\tilde{u}(t_{j+1},\cdot)} \\ &= \mathbb{E}\Bigg[|\alpha_1 - x|^2 + |I| \int_a^b |\alpha_2 - \mu(r, \Theta_r)|^2 \,\mathrm{d}r + \int_a^b |\alpha_3 + \alpha_4(W_r - W_a) - \sigma(r, \Theta_r)|^2 \,\mathrm{d}r \\ &\quad + |I| \int_a^b |\beta_2 - f(r, \Theta_r)|^2 \,\mathrm{d}r + |Y_b - \zeta(X_b^\alpha)|^2 \Bigg], \end{aligned} \tag{11}$$

where $\Theta_r = \Theta_r^{\alpha,\beta} := (X_r^\alpha, Y_r^\beta, \beta_3 + \beta_4(W_r - W_a))$, is as small as possible. The most suitable technique to obtain such a minimizer appears to be the Newton method, which converges very fast and requires the first derivative $\frac{\mathrm{d}}{\mathrm{d}(\alpha,\beta)} G(\alpha, \beta) \in \mathbb{R}^{1\times 8}$ and the second derivative $\frac{\mathrm{d}^2}{(\mathrm{d}(\alpha,\beta))^2} G(\alpha, \beta) \in \mathbb{R}^{1\times 8\times 8}$ of G. The first derivative can be reinterpreted as a vector in $\mathbb{R}^8$ and the second as an $\mathbb{R}^{8\times 8}$ - matrix which is to be inverted in the course of the implementation of the Newton method. Using the chain rule it is straightforward to deduce explicit expressions for $\frac{\mathrm{d}}{\mathrm{d}(\alpha,\beta)} G(\alpha, \beta)$ and $\frac{\mathrm{d}^2}{(\mathrm{d}(\alpha,\beta))^2} G(\alpha, \beta)$ based on (11). We omit these expressions for simplicity.

Note that since μ, σ, f, ζ are non-linear functions in general the expression for $G(\alpha, \beta)$ cannot be straightforwardly evaluated and the same applies to its derivatives. It is, however, natural to allow some error in the evaluation of G and its derivatives: As we will see later (see Proposition 5.2) there is an over-all error of order $|I|^{\frac{3}{2}}$ from the application of our one-step-scheme, even if there is no error in the evaluation of G. Thus, if we additionally allow an

error of similar magnitude in the evaluation and, therefore, in the minimization of $\sqrt{G}$, then, in light of Corollary 2.5, the convergence of the scheme, once proven under nice assumptions, should not be affected by the evaluation error. We propose to approximate G by a functional $\tilde{G}$ defined via

$$\tilde{G}(\alpha,\beta) := \mathbb{E}\Bigg[|\alpha_1 - x|^2 + |I|\int_a^b |\alpha_2 - \tilde{\mu}(r,\Theta_r)|^2\,\mathrm{d}r + \int_a^b |\alpha_3 + \alpha_4(W_r - W_a) - \tilde{\sigma}(r,\Theta_r)|^2\,\mathrm{d}r \\ + |I|\int_a^b |\beta_2 - \tilde{f}(r,\Theta_r)|^2\,\mathrm{d}r + |Y_b - \tilde{\zeta}(X_b^\alpha)|^2\Bigg], \tag{12}$$

where $\tilde{\mu},\tilde{\sigma},\tilde{f},\tilde{\zeta}$ are suitably chosen *polynomial approximations* of μ,σ,f,ζ. Note that unlike G, the functional $\tilde{G}$ and also its first and second derivative can be calculated explicitly in closed form using the coefficients of α,β and the polynomial coefficients of $\tilde{\mu},\tilde{\sigma},\tilde{f},\tilde{\zeta}$. This is because polynomials of stochastic polynomials are again stochastic polynomials and because expectations of stochastic polynomials can be calculated explicitly.

We reserve the subject of obtaining such polynomials $\tilde{\mu},\tilde{\sigma},\tilde{f},\tilde{\zeta}$ to future research. Let us merely comment that it appears sufficient to use the Taylor expansions of μ,σ,f,ζ at suitably chosen points and of suitably chosen order. For instance, we propose to approximate $\zeta = \zeta(\overline{x})$, $\overline{x} \in \mathbb{R}$, by

$$\tilde{\zeta}(\overline{x}) := \zeta(x) + \zeta'(x)(\overline{x} - x) + \frac{1}{2}\zeta''(x)(\overline{x} - x)^2,$$

yielding a local approximation of ζ at the point x of order $|\overline{x} - x|^3$. This requires ζ to be C^2 with known first and second derivatives at $x \in \mathbb{R}$, which is the initial value fixed in advance.

It is worth noting that with the polynomial approximations it is not necessary to use Monte-Carlo simulation or the law of large numbers to evaluate or minimize G. In general, our scheme does not require any simulation of any processes or random variables at all, which makes it very different from existing approaches to numerics of BSDE or FBSDE. It is also a key advantage of our method as Monte-Carlo simulation is usually a major bottle neck in calculations.

5 A convergence result

In the following we provide sufficient conditions for the scheme of Section 4 to be well-defined and to converge. For this we need to require the terminal condition, the decoupling field and related objects to be sufficiently smooth. To make this notion precise we introduce the following definition.

Definition 5.1. Let $k \in \mathbb{N}$ and $\underline{c}_i, \overline{c}_i : [0,T] \to \mathbb{R}$, $i = 1,\dots,k$ be such that $\underline{c}_i \le \overline{c}_i$ everywhere and $c_1 := \sup|\underline{c}_1| \vee \sup|\overline{c}_1| < L_{\sigma,z}^{-1}$. An arbitrary function $\zeta : \mathbb{R} \to \mathbb{R}$ is called $(k,\underline{c},\overline{c})$-*smooth for time* s, where $s \in [0,T]$, if ζ is k times weakly differentiable such that its i-th derivative $\zeta^{(i)}$ assumes values in $[\underline{c}_i(s), \overline{c}_i(s)]$ almost everywhere for all $i \in \{1,\dots,k\}$.

In the following we assume that there exists a triplet $(k,\underline{c},\overline{c})$ such that $(\xi,(\mu,\sigma,f))$ satisfy the following three regularity conditions:

(R1) Controlled solvability: The terminal condition ξ is $(k,\underline{c},\overline{c})$-smooth for T. In addition, there exists a $\delta_1 > 0$ such that for any $s \in [0,T]$ and any mapping $\zeta : \mathbb{R} \to \mathbb{R}$

that is $(k, \underline{c}, \bar{c})$-smooth for time s, there exists a weakly regular deterministic decoupling field v for the FBSDE on $[s', s]$, where $s' := 0 \vee (s - \delta_1)$, with terminal condition $v(s, \cdot) = \zeta$, such that $v(s', \cdot) : \mathbb{R} \to \mathbb{R}$ is $(k, \underline{c}, \bar{c})$-smooth for time s'.

(R2) Local boundedness w.r.t. $\|\cdot\|_{\frac{3}{2}}$: There exists a $\delta_2 \in (0, \delta_1]$ and a constant $\Gamma_1 > 0$ such that if ζ is $(k, \underline{c}, \bar{c})$-smooth for some time $s \in (0, T]$, then the processes X, Y associated with the decoupling field v and solving the FBSDE given by the initial condition $X_{s'} = x$, $s' := 0 \vee (s - \delta_2)$, and terminal condition $Y_s = \zeta(X_s)$, are such that $\|X\|_{\frac{3}{2}}, \|Y\|_{\frac{3}{2}} \leq \Gamma_1$ regardless of $x \in \mathbb{R}$.

(R3) Stability w.r.t. the terminal condition: There exists a $\delta_3 \in (0, \delta_1]$ and a constant $\Gamma_2 > 0$ such that if two terminal conditions ζ, $\overline{\zeta}$ are $(k, \underline{c}, \bar{c})$-smooth for some time $s \in (0, T]$, then the associated decoupling fields v, $\overline{v}$ on $[s', s]$, where $s' := 0 \vee (s - \delta_3)$, satisfy $\|v(r, \cdot) - \overline{v}(r, \cdot)\|_\infty \leq (1 + \Gamma_2(s - r))\|\zeta - \overline{\zeta}\|_\infty$ for all $r \in [s', s]$.

Moreover, we suppose that there exists an algorithm for approximating the functions $v(t_j, \cdot)$ with sufficiently smooth functions $\tilde{u}(t_j, \cdot)$. More precisely, we make the following assumption on this step in the scheme:

(SA) Selection of smooth approximations: There exists a constant $\gamma > 0$ and a calculation algorithm which assigns a $(k, \underline{c}, \bar{c})$-smooth $\tilde{u}(t_j, \cdot)$ to every $w(t_j, \cdot)$ such that

$$|v(t_j, \cdot) - \tilde{u}(t_j, \cdot)| \leq (1 + \gamma)\, C_1 \cdots C_4\, K(\Gamma_1, \Gamma_1) h^{\frac{3}{2}}, \tag{13}$$

where the constants $C_1, \ldots, C_4$ and the function K are defined as in Section 2.

Notice that (R1) is stronger than the assumption that there exists a weakly regular decoupling field u on $[0, T]$. It further entails that u is deterministic and that $u(s, \cdot)$ is $(k, \underline{c}, \bar{c})$-smooth for time s, where $s \in [0, T]$ is arbitrary.

The key to showing convergence of the scheme defined in the previous section is the following statement about the approximation w obtained from minimizing the error functional:

Proposition 5.2. *Assume (R1) and (R2). Let $I = [a, b] \subseteq [0, T]$ be such that $|I| \leq \delta_1, \delta_2$ is sufficiently small. Then*

$$|v(a, x) - w(a, x)| \leq C_1 C_2 C_3 C_4 K(\Gamma_1, \Gamma_1) |I|^{\frac{3}{2}}, \tag{14}$$

for all $x \in \mathbb{R}$.

Proof. Consider the decoupling field v for the FBSDE on I with terminal condition $\tilde{\xi}$. Let $x \in \mathbb{R}$ be arbitrary and consider the associated solution processes X^*, Y^*. Denote by $\hat{X}, \hat{Y}$ the processes which minimize $G_{x,I,\tilde{\xi}}$.

Notice that

$$\begin{aligned} |v(a, x) - w(a, x)| = |\hat{Y}_a - Y^*_a| &\leq \|(\hat{X}, \hat{Y}, \mathfrak{D}\hat{Y}) - (X^*, Y^*, \mathfrak{D}Y^*)\|_\infty \\ &\leq \|(\hat{X}, \hat{Y}) - (X^*, Y^*)\|_1. \end{aligned} \tag{15}$$

By Theorem 2.4 we have

$$\|(\hat{X}, \hat{Y}) - (X^*, Y^*)\|_1 \leq C_3 C_1 \sqrt{G_{x,I,\tilde{\xi}}(\hat{X}, \hat{Y})}. \tag{16}$$

According to Theorem 3.3 we have

$$\sqrt{G_{x,I,\tilde{\xi}}(\hat{X},\hat{Y})} \leq C_2 C_4 K(\|X^*\|_{\frac{3}{2}}, \|Y^*\|_{\frac{3}{2}})|I|^{\frac{3}{2}}. \tag{17}$$

Notice that by (R2) we have $K(\|X^*\|_{\frac{3}{2}}, \|Y^*\|_{\frac{3}{2}}) \leq K(\Gamma_1, \Gamma_1)$. By combining (15), (16) and (17) we arrive at (14). □

The associated multi-step scheme yields a mapping $\tilde{u}(0,\cdot)$, which depends on N and in fact converges to $u(0,\cdot)$ for $N \to \infty$ as the following result shows:

Theorem 5.3. *Under assumptions (R1)-(R3) and (SA) we have*

$$\|\tilde{u}(t_j,\cdot) - u(t_j,\cdot)\|_\infty \leq \frac{\Gamma_4(N-j)}{N^{\frac{3}{2}}},$$

for all $j = 1,\dots,N$, where Γ_4 is some constant which does not depend on N or j and where N is sufficiently large.

Proof. For the remainder of the proof we use the definition $C := (1+\gamma)\, C_1 \cdots C_4\, K(\Gamma_1,\Gamma_1)$. The triangle inequality, (SA) and (R3) imply

$$\begin{aligned}\|\tilde{u}(t_j,\cdot) - u(t_j,\cdot)\|_\infty &\leq \|\tilde{u}(t_j,\cdot) - v(t_j,\cdot)\|_\infty + \|v(t_j,\cdot) - u(t_j,\cdot)\|_\infty \\ &\leq Ch^{\frac{3}{2}} + (1+\Gamma_2 h)\|\tilde{u}(t_{j+1},\cdot) - u(t_{j+1},\cdot)\|_\infty.\end{aligned}$$

By applying the same estimate to $\|\tilde{u}(t_{j+1},\cdot) - u(t_{j+1},\cdot)\|_\infty$ we obtain

$$\|\tilde{u}(t_j,\cdot) - u(t_j,\cdot)\|_\infty \leq Ch^{\frac{3}{2}} + (1+\Gamma_2 h)Ch^{\frac{3}{2}} + (1+\Gamma_2 h)^2\|\tilde{u}(t_{j+2},\cdot) - u(t_{j+2},\cdot)\|_\infty.$$

Via induction one arrives at

$$\|\tilde{u}(t_j,\cdot) - u(t_j,\cdot)\|_\infty \leq Ch^{\frac{3}{2}} \sum_{k=0}^{N-1-j} (1+\Gamma_2 h)^k.$$

Note that $\sum_{k=0}^{N-1-j}(1+\Gamma_2 h)^k = \frac{1-(1+\Gamma_2 h)^{N-j}}{1-(1+\Gamma_2 h)} = \frac{(1+\Gamma_2 h)^{N-j}-1}{\Gamma_2 h} \leq \frac{1}{\Gamma_2 h}\left(\exp\left(\Gamma_2 T\frac{N-j}{N}\right) - 1\right)$. By using the inequality $e^{\alpha\beta} - 1 \leq \beta e^\alpha$ for all $\alpha \geq 0$ and $\beta \in [0,1]$ (the inequality follows from $e^{\alpha\beta} \leq \beta e^\alpha + (1-\beta)$, which can be shown with Jensen's inequality), we can further estimate $\exp\left(\Gamma_2 T\frac{N-j}{N}\right) - 1 \leq (e^{\Gamma_2 T} - 1)\frac{N-j}{N}$, and hence we obtain

$$\|\tilde{u}(t_j,\cdot) - u(t_j,\cdot)\|_\infty \leq \frac{C}{\Gamma_2}(e^{\Gamma_2 T} - 1)\frac{N-j}{N}h^{\frac{1}{2}}.$$

□

Let us now briefly comment on how to verify conditions (R1) to (R3) in practice:

Showing condition (R1) is basically the central topic of the so-called method of decoupling fields, which is used to show solvability of coupled FBSDE. The key idea of the method is to deduce the backward dynamics of the spatial derivative $u_x(s, X_s)$ of the decoupling field u and then show that if the terminal condition ξ is such that ξ' is in some bounded set, then $u_x(s, X_s)$ will also be in some bounded set depending on s. This method of decoupling fields

can be generalized to higher spatial derivatives of u as well (see [Fro18]), which allows to control these higher spatial derivatives.

Condition (R2) appears to be a natural consequence of the forward equation assuming the respective decoupling field is deterministic and sufficiently often differentiable with bounded derivatives: In this case X satisfies a standard SDE, which allows to control its $L^{2,\infty}$ norm. At the same time, since DX and $\mathfrak{D}X$ are sufficiently smooth functions of X, we can control $\|X\|_{\frac{3}{2}}$ by means of several applications of the Itô formula. Using the decoupling condition and additional applications of the Itô formula one can control $\|Y\|_{\frac{3}{2}}$ as well.

Verifying condition (R3) is also based on smoothness properties of the respective decoupling field, but may be more complicated to conduct, even though it appears to be a natural stability property of the considered system: Without this property there would be no control on the error propagation and small errors in the calculation might compound in a way which would render all approximations useless, regardless of the actual method used to obtain approximations on small intervals.

6 Examples of FBSDEs satisfying (R1)-(R3)

Before verifying conditions (R1)-(R3) let us point out that these conditions are likely to be required by any practically implementable and convergent scheme for FBSDE: Condition (R1) essentially means that the problem is solvable in the first place. Because of approximation errors vested in any numerical approximation algorithm, solvability must be required for a whole range of terminal conditions, which is why condition (R1) is stronger than the more standard notions of solvability.

Condition (R2) is needed for the existence of local approximations of order $\frac{3}{2}$ for the backward and the forward processes. This in turn is required for the overall convergence of the multi-step scheme since a local approximation of this order results in a global rate of convergence of $\frac{1}{2}$. Finally, as already mentioned, condition (R3) is necessary to make sure that error propagation does not lead to divergence.

However, this observation that conditions (R1)-(R3) are likely to be satisfied by any numerically treatable problem, does not mean that these conditions are straightforward to prove, especially in a general strongly coupled and/or generalized multi-dimensional setting. Their verification actually requires a rather deep knowledge of the behavior of the decoupling field u and its spatial derivatives.

In this section we verify (R1)-(R3) for two different classes of coupled FBSDE. Throughout the section we only consider $(\xi, (\mu, \sigma, f))$ satisfying (SLC). We further introduce the following assumptions:

(A1) $n = m = d = 1$ and all functions are deterministic;

(A2) μ, σ, f do not depend on z, but only on (t, x, y);

(A3) μ, σ, f are bounded;

(A4) μ, σ, f are thrice continuously differentiable with bounded derivatives.

In the following we show conditions (R1) to (R3) in two different settings: For the decoupled case and a coupled case under monotonicity assumptions.

6.1 Verification of (R1)

Lemma 6.1. *Assume (A1)-(A4), that μ and σ do not depend on y and that ξ is twice continuously differentiable with Lipschitz continuous derivatives. Then there exist constants $r_i, R_i \geq 0$, $i = 1,2,3$, such that $(\xi, (\mu, \sigma, f))$ satisfies condition (R1) w.r.t. $(3, \underline{c}, \overline{c})$, where $\underline{c}_i(s) := -R_i e^{r_i(T-s)}$ and $\overline{c}_i(s) := R_i e^{r_i(T-s)}$ for $i = 1,2,3$.*

Proof. According to Section 6 of [Fro18] (see Theorems 6.4. and 6.5.) there exists an interval $J \subseteq [0,T]$, $T \in J$, on which there is a strongly regular decoupling field u, such that the first derivative u_x w.r.t. x and the second derivative u_{xx} exist and are Lipschitz continuous w.r.t. $x \in \mathbb{R}$. Moreover, according to these results, J can be chosen such that for each $t \in J$ and each $x \in \mathbb{R}$ there exist processes $X, Y, Z, \check{Y}^{(i)}, \check{Z}^{(i)}$, $i = 1,2$, defined on $[t,T]$ such that

- X, Y, Z satisfy the FBSDE with initial conditoin $X_t = x$ and terminal condition ξ,
- $u(s, X_s) = Y_s$, $u_x(s, X_s) = \check{Y}_s^{(1)}$, $u_{xx}(s, X_s) = \check{Y}_s^{(2)}$, a.s. for all $s \in [t,T]$,
- $\check{Y}^{(1)}, \check{Z}^{(1)}$ satisfy the BSDE on $[t,T]$ with generator $(s, \check{y}^{(1)}, \check{z}^{(1)}) \mapsto \varphi^{(1)}(s, X_s, Y_s, \check{y}^{(1)}, \check{z}^{(1)})$, where

$$\varphi^{(1)}(s,x,y,\check{y}^{(1)},\check{z}^{(1)}) = \left(f_x + \check{y}^{(1)}\left(f_y - \mu_x\right) - \check{z}^{(1)}\sigma_x\right)(s,x,y); \tag{18}$$

  and $\check{Y}^{(2)}, \check{Z}^{(2)}$ the BSDE with generator $(s, \check{y}^{(2)}, \check{z}^{(2)}) \mapsto \varphi^{(2)}(s, X_s, Y_s, \check{Y}^{(1)}, \check{Z}^{(1)}, \check{y}^{(2)}, \check{z}^{(2)})$, where

$$\begin{aligned}\varphi^{(2)}(s,x,y,\check{y}^{(1)},\check{z}^{(1)},\check{y}^{(2)},\check{z}^{(2)}) &= \left(\varphi_x^{(1)} + \check{y}^{(1)}\varphi_y^{(1)}\right)(s,x,y,\check{y}^{(1)},\check{z}^{(1)}) \\ &+ \left(\check{y}^{(2)}\left(f_y - 2\mu_x\right) - \left(2\check{z}^{(2)} + \check{y}^{(2)}\sigma_x\right)\sigma_x\right)(s,x,y)\end{aligned} \tag{19}$$

  (cf. with Section 5 in [Fro18]).

Let us first inspect the dynamics of $\check{Y}^{(1)}$. Clearly, $\check{Y}^{(1)}$ satisfies a linear BSDE and, therefore, grows at most exponentially backwards in time. The growth rate r_1 is determined by $\|f_y - \mu_x\|_\infty < \infty$ and the initial bound R_1 by $\|\xi'\|_\infty$ and $T\|f_x\|_\infty$. In particular, $\check{Y}^{(1)}$ is bounded independently of t, x.

Because of uniform boundedness of $\check{Y}^{(1)}$ established above, $\check{Y}^{(2)}$ now also satisfies a linear BSDE, which yields constants r_2, R_2, similarly to r_1, R_1. As a consequence, $\check{Y}^{(2)}$ is also uniformly bounded.

Now according to Corollary 6.11. of [Fro18] we may choose $J = [0,T]$. In particular, the weak derivative u_{xxx} is uniformly bounded. More precisely, Corollary 6.11. of [Fro18] is based on Theorem 6.5. of [Fro18] and according to the final remarks of the proof of this Theorem u_{xxx} can be controlled more explicitly based on the bounds for other processes: We can choose t sufficiently close to T such that $\partial_x X$ does not vanish and we can define a process $\check{Y}^{(3)}$ via $\check{Y}^{(3)} := \frac{\partial_x \check{Y}^{(2)}}{\partial_x X}$. Note that $\check{Y}^{(3)} = u_{xxx}(s, X_s)$ almost everywhere using the chain rule. According to the proof of Theorem 6.5. of [Fro18] the backward process $\check{Y}^{(3)}$ grows at most exponentially backwards in time starting with some value bounded by $\|\xi^{(3)}\|_\infty$. This yields constants r_3, R_3, which control the growth of u_{xxx} backwards in time.

The same line of reasoning can be applied to any sufficiently small interval $[a,b] \subseteq [0,T]$, and hence we obtain condition (R1) in its full generality. □

Lemma 6.2. *Assume (A1)-(A4), that σ does not depend on y and that ξ is twice continuously differentiable with Lipschitz continuous derivatives. Assume further that $\xi' \geq 0$, $\mu_y \leq 0$ and $f_x \geq 0$ everywhere. Then there exist constants $r_i, R_i \geq 0$, $i = 1, 2, 3$, such that $(\xi, (\mu, \sigma, f))$ satisfies condition (R1) w.r.t. $(3, \underline{c}, \overline{c})$, where $\underline{c}_1(s) := 0$, $\overline{c}_1(s) := R_1 e^{r_1(T-s)}$, $\underline{c}_i(s) := -R_i e^{r_i(T-s)}$ and $\overline{c}_i(s) := R_i e^{r_i(T-s)}$ for $i = 2, 3$.*

Proof. According to Section 6 of [Fro18] (see Theorems 6.4. and 6.5.) there exists an interval $J \subseteq [0, T]$, $T \in J$, on which there is a strongly regular decoupling field u, such that the first derivative u_x w.r.t. x and the second derivative u_{xx} exist and are Lipschitz continuous w.r.t. $x \in \mathbb{R}$. Moreover, according to these results, J can be chosen such that for each $t \in J$ and each $x \in \mathbb{R}$ there exist processes $X, Y, Z, \check{Y}^{(i)}, \check{Z}^{(i)}$, $i = 1, 2$, such that

- $u(s, X_s) = Y_s$, $u_x(s, X_s) = \check{Y}_s^{(1)}$, $u_{xx}(s, X_s) = \check{Y}_s^{(2)}$, a.s. for all $s \in [0, T]$,
- X, Y, Z satisfy the FBSDE with initial conditoin $X_t = x$ and terminal condition ξ,
- $\check{Y}^{(1)}, \check{Z}^{(1)}$ and $\check{Y}^{(2)}, \check{Z}^{(2)}$ satisfy BSDEs given by $D\check{Y}_s^{(1)} = \varphi^{(1)}(s, X_s, Y_s, \check{Y}^{(1)}, \check{Z}^{(1)})$ and $D\check{Y}_s^{(2)} = \varphi^{(2)}(s, X_s, Y_s, \check{Y}^{(1)}, \check{Z}^{(1)}, \check{Y}^{(2)}, \check{Z}^{(2)})$, where $\check{Z}^{(i)} = \mathfrak{D}\check{Y}^{(i)}$, $i = 1, 2$, and where $\varphi^{(1)}$, $\varphi^{(2)}$ are calculated using recursion (10) of [Fro18].

Adapted to our setting we have

$$\varphi^{(1)}(s, x, y, \check{y}^{(1)}, \check{z}^{(1)}) = \left(f_x + \check{y}^{(1)}\left(f_y - \mu_x - \check{y}^{(1)}\mu_y\right) - \check{z}^{(1)}\sigma_x\right)(s, x, y) \tag{20}$$

and

$$\begin{aligned}\varphi^{(2)}(s, x, y, \check{y}^{(1)}, \check{z}^{(1)}, \check{y}^{(2)}, \check{z}^{(2)}) &= \left(\varphi_x^{(1)} + \check{y}^{(1)}\varphi_y^{(1)}\right)(s, x, y, \check{y}^{(1)}, \check{z}^{(1)}) \\ &+ \left(\check{y}^{(2)}\left(f_y - 2\mu_x - 3\check{y}^{(1)}\mu_y\right) - \left(2\check{z}^{(2)} + \check{y}^{(2)}\sigma_x\right)\sigma_x\right)(s, x, y).\end{aligned}$$

Let us inspect the dynamics of $\check{Y}^{(1)}$: Firstly, observe that the summand $-\check{z}^{(1)}\sigma_x$ can be effectively removed by performing a Girsanov type measure change and then considering the dynamics of $\check{Y}^{(1)}$ under the new measure. Secondly, using that $\mu_y \leq 0$ and $f_x \geq 0$ a comparison theorem implies that $\check{Y}^{(1)}$ is non-negative and grows at most exponentially backwards in time, where the growth rate is determined by $f_y - \mu_x$ and where for the constant R_1 we can take $\|\xi'\|_\infty$.

Let us now inspect $\check{Y}^{(2)}$: Clearly, once uniform boundedness of $\check{Y}^{(1)}$ is established (which is done above), $\check{Y}^{(2)}$ satisfies a linear BSDE, s.t. it grows at most exponentially backwards in time, which implies corresponding constants r_2, R_2. As a consequence, $\check{Y}^{(2)}$ is also uniformly bounded.

Now according to Corollary 6.11. of [Fro18] we may choose $J = [0, T]$. In particular, the weak derivative u_{xxx} is uniformly bounded. More precisely, Corollary 6.11. of [Fro18] is based on Theorem 6.5. of [Fro18] and according to the final remarks of the proof of this Theorem u_{xxx} can be controlled more explicitely based on the bounds for other processes: We can choose t sufficiently close to T such that $\partial_x X$ does not vanish and we can define a process $\check{Y}^{(3)}$ via $\check{Y}^{(3)} := \frac{\partial_x \check{Y}^{(2)}}{\partial_x X}$. Note that $\check{Y}^{(3)} = u_{xxx}(s, X_s)$ almost everywhere using the chain rule. According to the proof of Theorem 6.5. of [Fro18] the backward process $\check{Y}^{(3)}$ grows at most exponentially backwards in time starting with some value bounded by $\|\xi^{(3)}\|_\infty$. This yields constants r_3, R_3, which control the growth of u_{xxx} backwards in time.

The same line of reasoning can be applied to any sufficiently small interval $[a, b] \subseteq [0, T]$, and hence condition (R1) is satisfied in its full generality. □

6.2 Verification of (R2)

Lemma 6.3. *Assume that μ, σ, f, ξ either satisfy the conditions of Lemma 6.1 or of Lemma 6.2. Let $(3, \underline{c}, \overline{c})$ be the respective triplet w.r.t. which condition (R1) is satisfied. Then condition (R2) is also satisfied w.r.t. the same triplet.*

Proof. As we have seen in the proofs of Lemma 6.1 and Lemma 6.2 condition (R1) is indeed satified and, in addition, we have rather explicit knowledge of the dynamics of the spatial derivatives of the associated decoupling field: Consider an interval $[a, b] \subseteq [0, T]$, a for time b $(3, \underline{c}, \overline{c})$ - smooth ζ and the associated X, Y, Z on $[a, b]$ satisfying the FBSDE with initial condition $X_a = x \in \mathbb{R}$ and the decoupling condition $Y_t = v(t, X_t)$, $t \in [a, b]$. Then the decoupling field v is twice classically differentible w.r.t. x and the processes $\check{Y}_t^{(1)} = v_x(t, X_t)$, $\check{Y}_t^{(2)} = v_{xx}(t, X_t)$ satisfy backward SDEs given by the generators $\varphi^{(1)}$, $\varphi^{(2)}$ mentioned in the proofs of Lemmas 6.1, 6.2. We have uniform boundedness of $\check{Y}^{(1)}$, $\check{Y}^{(2)}$, and also of the associated control processes $\check{Z}^{(1)}$, $\check{Z}^{(2)}$ (cf. Theorem 6.4 in [Fro18]). We can apply an Itô formula to v and v_x to obtain $Z_t = \mathfrak{D}Y_t = v_x(t, X_t)\sigma(t, X_t, Y_t)$ and $\check{Z}_t^{(1)} = \mathfrak{D}\check{Y}_t^{(1)} = v_{xx}(t, X_t)\sigma(t, X_t, Y_t)$.

Now to verify (R2) we must control the $\|\cdot\|_{\frac{3}{2}}$ - norm of X and Y:
Observe that $\|X\|_{\frac{3}{2}}$ is defined as the maximum of the five values $\|DDX\|_{L^2}$, $\|\mathfrak{D}DX\|_{L^{2,\infty}}$, $\|D\mathfrak{D}X\|_{L^{2,\infty}}$, $\|D\mathfrak{D}\mathfrak{D}X\|_{L^2}$ and $\|\mathfrak{D}\mathfrak{D}\mathfrak{D}X\|_{L^{2,\infty}}$. We have $DX_s = \mu(s, X_s, Y_s)$ and $\mathfrak{D}X_s = \sigma(s, X_s, Y_s)$. Using the Itô formula we also obtain

$$\begin{aligned} DDX_t =& \mu_t(t, X_t, Y_t) + \mu_x(t, X_t, Y_t)\mu(t, X_t, Y_t) + \mu_y(t, X_t, Y_t)f(t, X_t, Y_t) \\ &+ \frac{1}{2}\left(\mu_{xx}(t, X_t, Y_t)\sigma^2(t, X_t, Y_t) + \mu_{yy}(t, X_t, Y_t)Z_t^2 + 2\mu_{xy}(t, X_t, Y_t)\sigma(t, X_t, Y_t)Z_t\right) \end{aligned}$$

and a similar expression for $D\mathfrak{D}X$. Both expressions are uniformly bounded, because $Z_t = v_x(t, X_t)\sigma(t, X_t, Y_t)$ is bounded. This yields uniform boundedness of $\|DDX\|_{L^2}$ and $\|D\mathfrak{D}X\|_{L^{2,\infty}}$. We obtain a more simple expression for $\mathfrak{D}DX$:

$$\mathfrak{D}DX_t = \mu_x(t, X_t, Y_t)\sigma(t, X_t, Y_t) + \mu_y(t, X_t, Y_t)Z_t,$$

which is also uniformly bounded. It remains to consider the two expressions $D\mathfrak{D}\mathfrak{D}X$ and $\mathfrak{D}\mathfrak{D}\mathfrak{D}X$. We first calculate $\mathfrak{D}\mathfrak{D}X$:

$$\mathfrak{D}\mathfrak{D}X_t = \sigma_x(t, X_t, Y_t)\sigma(t, X_t, Y_t) + \sigma_y(t, X_t, Y_t)Z_t.$$

Now note that Z is also an Itô process: $Z_t = v_x(t, X_t)\sigma(t, X_t, Y_t) = \check{Y}_t^{(1)}\mathfrak{D}X_t$, where $\check{Y}^{(1)}$ satisfies a BSDE with generator $\varphi^{(1)}(s, X_s, Y_s, \check{y}^{(1)}, \check{z}^{(1)})$, where φ^1 is defined by (18) and (20), respectively. We, thus, obtain

$$\begin{aligned} \mathfrak{D}\mathfrak{D}X_t =& \left(\sigma_x(t, X_t, Y_t) + \sigma_y(t, X_t, Y_t)\check{Y}_t^{(1)}\right)\mathfrak{D}X_t, \\ \mathfrak{D}\mathfrak{D}\mathfrak{D}X_t =& \Big(\sigma_{xx}(t, X_t, Y_t)\mathfrak{D}X_t + \sigma_{xy}(t, X_t, Y_t)Z_t \\ &+ \left(\sigma_{yx}(t, X_t, Y_t)\mathfrak{D}X_t + \sigma_{yy}(t, X_t, Y_t)Z_t\right)\check{Y}_t^{(1)} + \sigma_y(t, X_t, Y_t)\check{Z}_t^{(1)}\Big)\mathfrak{D}X_t \\ &+ \left(\sigma_x(t, X_t, Y_t) + \sigma_y(t, X_t, Y_t)\check{Y}_t^{(1)}\right)\mathfrak{D}\mathfrak{D}X_t \end{aligned}$$

and a somewhat longer expression for $D\mathfrak{D}\mathfrak{D}X_t$ involving third derivatives of σ and φ^1. Note that all expressions are bounded, because σ and its derivatives are bounded and Z and $\check{Z}^{(1)}$ are uniformly bounded.

It remains to uniformly control $\|Y\|_{\frac{3}{2}}$: To this end we need to control the five values $\|DDY\|_{L^2}$, $\|\mathfrak{D}DY\|_{L^{2,\infty}}$, $\|D\mathfrak{D}Y\|_{L^{2,\infty}}$, $\|D\mathfrak{D}\mathfrak{D}Y\|_{L^2}$ and $\|\mathfrak{D}\mathfrak{D}\mathfrak{D}Y\|_{L^{2,\infty}}$. Controlling $\|DDY\|_{L^2}$ and $\|\mathfrak{D}DY\|_{L^{2,\infty}}$ is analogous to the above, because $DY_t = f(t, X_t, Y_t)$ and because f has the same smoothness properties as μ. Next note that $\mathfrak{D}Y_t = Z_t = \check{Y}_t^{(1)}\mathfrak{D}X_t$. We obtain

$$D\mathfrak{D}Y_t = D\check{Y}_t^{(1)}\mathfrak{D}X_t + \check{Y}_t^{(1)}D\mathfrak{D}X_t + \check{Z}_t^{(1)}\mathfrak{D}\mathfrak{D}X_t,$$

which is uniformly bounded. Furthermore,

$$\mathfrak{D}\mathfrak{D}Y_t = \check{Z}_t^{(1)}\mathfrak{D}X_t + \check{Y}_t^{(1)}\mathfrak{D}\mathfrak{D}X_t,$$

which is also bounded. More importantly, the Itô formula yields $\check{Z}_t^{(1)} = \check{Y}_t^{(2)}\mathfrak{D}X_t$, such that

$$\mathfrak{D}\mathfrak{D}Y_t = \check{Y}_t^{(2)}\mathfrak{D}X_t\mathfrak{D}X_t + \check{Y}_t^{(1)}\mathfrak{D}\mathfrak{D}X_t.$$

This allows to calculate $D\mathfrak{D}\mathfrak{D}Y_t$ and $\mathfrak{D}\mathfrak{D}\mathfrak{D}Y_t$ explicitly using the product rule. The resulting expressions are uniformly bounded due to the uniform boundedness of $\|X\|_{\frac{3}{2}}$ shown above and the uniform boundedness of $\check{Y}^{(1)}$, $\check{Y}^{(2)}$, $\check{Z}^{(1)}$, $\check{Z}^{(2)}$.

□

6.3 Verification of (R3)

The objective of this section is to verify assumption (R3) in a setting that is more general than that of Lemma 6.1 and 6.2. For simplicity we consider an interval $[t, T]$ with two terminal conditions $\xi, \tilde{\xi}$. We essentially assume that at least one of the terminal conditions is sufficiently smooth, while the other is merely Lipschitz continuous.

For fixed μ, σ, f we consider two different terminal conditions $\xi, \tilde{\xi} : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$, such that both $(\xi, (\mu, \sigma, f))$ and $(\tilde{\xi}, (\mu, \sigma, f))$ satisfy SLC. Assume further that there exist two weakly regular decoupling fields $u, \tilde{u} : [t, T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^m$ for the two problems such that we have $u(T, \cdot) = \xi$ and $\tilde{u}(T, \cdot) = \tilde{\xi}$. Our objective is to control

$$\|u - \tilde{u}\|_{t,\infty} := \sup_{s\in[t,T]} \operatorname{ess\,sup}_{\omega\in\Omega} \sup_{x\in\mathbb{R}^n} |u(s, \omega, x) - \tilde{u}(s, \omega, x)|$$

by $\|\xi - \tilde{\xi}\|_\infty$, the interval length $T - t$ and the Lipschitz constants for $(\xi, (\mu, \sigma, f))$ and $(\tilde{\xi}, (\mu, \sigma, f))$.

For a fixed initial condition $x \in \mathbb{R}^n$ consider the associated solution triplets (X, Y, Z) and $(\tilde{X}, \tilde{Y}, \tilde{Z})$ on $[t, T]$. We have $Y_s = u(s, X_s)$ and $\tilde{Y}_s = \tilde{u}(s, \tilde{X}_s)$ a.s. for all $s \in [t, T]$. Also, we have $X_t = \tilde{X}_t = x$.

We first derive an estimate for the distance between Y_t and $\tilde{Y}_t$ by using the backward equations.

Lemma 6.4. *Assume (A2) and (A3). Assume further that ξ is deterministic and thrice weakly differentiable with derivatives bounded by some constant $B \in [0, \infty)$. Then there exists*

a constant $C > 0$, depending only on $T - t$, B, $\|\mu\|_\infty$, $\|\sigma\|_\infty$ and the Lipschitz constant of (μ, σ, f), such that for all $s \in [t, T]$ we have

$$|Y_t - \tilde{Y}_t| \leq \|\xi - \tilde{\xi}\|_\infty + C(T-t)\left(\sup_{s\in[t,T]} \left\|\mathbb{E}\left[|X_s - \tilde{X}_s| \middle| \mathcal{F}_t\right]\right\|_\infty + \|u - \tilde{u}\|_{t,\infty}\right). \tag{21}$$

Proof. Notice that

$$Y_t - \tilde{Y}_t = \xi(X_T) - \tilde{\xi}(\tilde{X}_T) - \int_t^T (Z_s - \tilde{Z}_s)\,\mathrm{d}W_s - \int_t^T \Big(f(s, X_s, Y_s) - f(s, \tilde{X}_s, \tilde{Y}_s)\Big)\,\mathrm{d}s. \tag{22}$$

By taking the conditional expectation w.r.t. $\mathcal{F}_t$ the Brownian integral part drops out, and with the triangle inequality we further obtain

$$\left|Y_t - \tilde{Y}_t\right| \leq \left|\mathbb{E}\left[\xi(X_T) - \tilde{\xi}(\tilde{X}_T) \middle| \mathcal{F}_t\right]\right| + \left|\mathbb{E}\left[\int_t^T \Big(f(s, X_s, Y_s) - f(s, \tilde{X}_s, \tilde{Y}_s)\Big)\,\mathrm{d}s \middle| \mathcal{F}_t\right]\right|. \tag{23}$$

The first term of the RHS of (23) satisfies

$$\begin{aligned}\left|\mathbb{E}\left[\xi(X_T) - \tilde{\xi}(\tilde{X}_T) \middle| \mathcal{F}_t\right]\right| &\leq \left|\mathbb{E}\left[\xi(X_T) - \xi(\tilde{X}_T) \middle| \mathcal{F}_t\right]\right| + \left|\mathbb{E}\left[\xi(\tilde{X}_T) - \tilde{\xi}(\tilde{X}_T) \middle| \mathcal{F}_t\right]\right| \\ &\leq \left|\mathbb{E}\left[\xi(X_T) - \xi(\tilde{X}_T) \middle| \mathcal{F}_t\right]\right| + \|\xi - \tilde{\xi}\|_\infty.\end{aligned}$$

With the Itô formula we obtain

$$\begin{aligned}\xi(X_T) - \xi(\tilde{X}_T) = &\int_t^T \Big(\xi'(X_s)\mu(s, X_s, Y_s) - \xi'(\tilde{X}_s)\mu(s, \tilde{X}_s, \tilde{Y}_s)\Big)\,ds + M_{t,T} \\ &+ \frac{1}{2}\int_t^T \Big(\xi''(X_s)\sigma^2(s, X_s, Y_s) - \xi''(\tilde{X}_s)\sigma^2(s, \tilde{X}_s, \tilde{Y}_s)\Big)\,ds,\end{aligned}$$

where $M_{t,T}$ is Brownian integral satisfying $\mathbb{E}[M_{t,T}|\mathcal{F}_t] = 0$. By taking the conditional expectation w.r.t. $\mathcal{F}_t$, and applying the triangle and Jensen inequality we thus get

$$\begin{aligned}\Big|\mathbb{E}[\xi(X_T) - \xi(\tilde{X}_T)|\mathcal{F}_t]\Big| \leq &\int_t^T \mathbb{E}\left[\Big|\xi'(X_s)\mu(s, X_s, Y_s) - \xi'(\tilde{X}_s)\mu(s, \tilde{X}_s, \tilde{Y}_s)\Big| \,\Big|\mathcal{F}_t\right] ds \\ &+ \frac{1}{2}\int_t^T \mathbb{E}\left[\Big|\xi''(X_s)\sigma^2(s, X_s, Y_s) - \xi''(\tilde{X}_s)\sigma^2(s, \tilde{X}_s, \tilde{Y}_s)\Big| \,\Big|\mathcal{F}_t\right] ds.\end{aligned} \tag{24}$$

Note that

$$\begin{aligned}&\Big|\xi'(X_s)\mu(s, X_s, Y_s) - \xi'(\tilde{X}_s)\mu(s, \tilde{X}_s, \tilde{Y}_s)\Big| \\ \leq &\|\mu\|_\infty|\xi'(X_s) - \xi'(\tilde{X}_s)| + \|\xi'\|_\infty(L_{\mu,x}|X_s - \tilde{X}_s| + L_{\mu,y}|Y_s - \tilde{Y}_s|) \\ \leq &\|\mu\|_\infty L_{\xi'}|X_s - \tilde{X}_s| + \|\xi'\|_\infty(L_{\mu,x}|X_s - \tilde{X}_s| + L_{\mu,y}|u(s, X_s) - \tilde{u}(s, \tilde{X}_s)|).\end{aligned}$$

Moreover,

$$\begin{aligned}|u(s, X_s) - \tilde{u}(s, \tilde{X}_s)| &\leq |u(s, X_s) - u(s, \tilde{X}_s)| + |u(s, \tilde{X}_s) - \tilde{u}(s, \tilde{X}_s)| \\ &\leq L_{u,x}|X_s - \tilde{X}_s| + \|u - \tilde{u}\|_{t,\infty},\end{aligned} \tag{25}$$

and hence

$$\begin{aligned}&\left|\xi'(X_s)\mu(s,X_s,Y_s)-\xi'(\tilde{X}_s)\mu(s,\tilde{X}_s,\tilde{Y}_s)\right|\\ \leq&(\|\mu\|_\infty L_{\xi'}+\|\xi'\|_\infty L_{\mu,x}+L_{\mu,y}L_{u,x})|X_s-\tilde{X}_s|+L_{\mu,y}\|u-\tilde{u}\|_{t,\infty}.\end{aligned}\tag{26}$$

Similarly, we obtain the estimate

$$\begin{aligned}&\left|\xi''(X_s)\sigma^2(s,X_s,Y_s)-\xi''(\tilde{X}_s)\sigma^2(s,\tilde{X}_s,\tilde{Y}_s)\right|\\ \leq&(\|\sigma\|_\infty^2 L_{\xi''}+\|\xi''\|_\infty L_{\sigma^2,x}+L_{\sigma^2,y}L_{u,x})|X_s-\tilde{X}_s|+L_{\sigma^2,y}\|u-\tilde{u}\|_{t,\infty}.\end{aligned}\tag{27}$$

Observe that $L_{\sigma^2,x}$ and $L_{\sigma^2,y}$ are finite, because σ is Lipschitz contiuous and bounded. In the following $K_1, K_2, \ldots$ denote constants depending only on $T-t$, B, $\|\mu\|_\infty$, $\|\sigma\|_\infty$ and the Lipschitz constant of (μ,σ,f).

By combining estimates (26), (27) and (24) we obtain, for constants K_1 and K_2, such that

$$\begin{aligned}\left|\mathbb{E}[\xi(X_T)-\xi(\tilde{X}_T)|\mathcal{F}_t]\right| &\leq \int_t^T \left(K_1\mathbb{E}\left[|X_s-\tilde{X}_s|\big|\mathcal{F}_t\right]+K_2\|u-\tilde{u}\|_{t,\infty}\right)ds\\ &\leq (T-t)\left(K_1\sup_{s\in[t,T]}\left\|\mathbb{E}\left[|X_s-\tilde{X}_s|\big|\mathcal{F}_t\right]\right\|_\infty+K_2\|u-\tilde{u}\|_{t,\infty}\right).\end{aligned}\tag{28}$$

With similar arguments one can show that for two constants K_3 and K_4 we have

$$\begin{aligned}&\left|\mathbb{E}\left[\int_t^T\left(f(s,X_s,Y_s)-f(s,\tilde{X}_s,\tilde{Y}_s)\right)\mathrm{d}s\big|\mathcal{F}_t\right]\right|\\ &\leq (T-t)\left(K_3\sup_{s\in[t,T]}\left\|\mathbb{E}\left[|X_s-\tilde{X}_s|\big|\mathcal{F}_t\right]\right\|_\infty+K_4\|u-\tilde{u}\|_{t,\infty}\right).\end{aligned}\tag{29}$$

Combining (22), (28) and (29) we obtain Inequality (21). □

Lemma 6.5. *There exists a constant $C\geq 0$, which depends only on $T-t$, $L_{u,x}$ and the Lipschitz constants of (μ,σ), such that for all $s\in[t,T]$ we have*

$$\operatorname{ess\,sup}_{\omega\in\Omega}\mathbb{E}\left[|X_s-\tilde{X}_s|^2\big|\mathcal{F}_t\right]^{\frac{1}{2}}\leq C\sqrt{s-t}\|u-\tilde{u}\|_{t,\infty}.\tag{30}$$

Proof. Throughout the proof $K_1, K_2, \ldots$ denote constants depending only on $L_{u,x}$, $T-t$ and the Lipschitz constants of μ and σ.

Using the forward equations for X and $\tilde{X}$ we have

$$X_s-\tilde{X}_s=\int_t^s\left(\mu(r,X_r,Y_r)-\mu(r,\tilde{X}_r,\tilde{Y}_r)\right)\mathrm{d}r+\int_t^s\left(\sigma(r,X_r,Y_r)-\sigma(r,\tilde{X}_r,\tilde{Y}_r)\right)\mathrm{d}W_r.$$

With the Minkowski inequality for conditional expectations, the Itô isometry and the Cauchy-

Schwarz inequality we obtain

$$\begin{aligned}\mathbb{E}\left[|X_s-\tilde{X}_s|^2\big|\mathcal{F}_t\right]^{\frac{1}{2}} \leq& \left(\mathbb{E}\left[\left|\int_t^s \left(\mu(r,X_r,Y_r)-\mu(r,\tilde{X}_r,\tilde{Y}_r)\right)\,\mathrm{d}r\big|\mathcal{F}_t\right|^2\right]\right)^{\frac{1}{2}}\\ &+\left(\mathbb{E}\left[\int_t^s \left|\sigma(r,X_r,Y_r)-\sigma(r,\tilde{X}_r,\tilde{Y}_r)\right|^2\,\mathrm{d}s\big|\mathcal{F}_t\right]\right)^{\frac{1}{2}}\\ \leq& \left(\mathbb{E}\left[(s-t)\int_t^s \left|\mu(r,X_r,Y_r)-\mu(r,\tilde{X}_r,\tilde{Y}_r)\right|^2\,\mathrm{d}r\big|\mathcal{F}_t\right]\right)^{\frac{1}{2}}\\ &+\left(\mathbb{E}\left[\int_t^s \left|\sigma(r,X_r,Y_r)-\sigma(r,\tilde{X}_s,\tilde{Y}_r,\tilde{Z}_r)\right|^2\,\mathrm{d}r\big|\mathcal{F}_t\right]\right)^{\frac{1}{2}}.\end{aligned}$$

Now we exploit the Lipschitz continuity of μ and σ in combination with the Minkowski inequality to obtain

$$\mathbb{E}\left[|X_s-\tilde{X}_s|^2\big|\mathcal{F}_t\right]^{\frac{1}{2}} \leq K_1\left(\mathbb{E}\left[\int_t^s |X_r-\tilde{X}_r|^2\,\mathrm{d}r\big|\mathcal{F}_t\right]\right)^{\frac{1}{2}}+K_2\left(\mathbb{E}\left[\int_t^s |Y_r-\tilde{Y}_r|^2\,\mathrm{d}r\big|\mathcal{F}_t\right]\right)^{\frac{1}{2}}.$$

By using the decoupling condition and the estimate (25) we have $|Y_r-\tilde{Y}_r|\leq \|u-\tilde{u}\|_{t,\infty}+L_{u,x}|X_r-\tilde{X}_r|$. Moreover, with the Minkowski inequality we obtain

$$\mathbb{E}\left[\int_t^s |Y_r-\tilde{Y}_r|^2\,\mathrm{d}r\big|\mathcal{F}_t\right]^{\frac{1}{2}} \leq \mathbb{E}\left[\int_t^s \|u-\tilde{u}\|_{t,\infty}^2\,\mathrm{d}r\big|\mathcal{F}_t\right]^{\frac{1}{2}}+L_{u,x}\mathbb{E}\left[\int_t^s |X_r-\tilde{X}_r|^2\,\mathrm{d}r\big|\mathcal{F}_t\right]^{\frac{1}{2}},$$

which further implies

$$\mathbb{E}\left[|X_s-\tilde{X}_s|^2\big|\mathcal{F}_t\right]^{\frac{1}{2}} \leq K_3\mathbb{E}\left[\int_t^s |X_r-\tilde{X}_r|^2\,\mathrm{d}r\big|\mathcal{F}_t\right]^{\frac{1}{2}}+K_4\sqrt{s-t}\|u-\tilde{u}\|_{t,\infty}.$$

Next we square both sides of the previous inequality and then apply the Young inequality, i.e. $ab\leq \frac{1}{2}a^2+\frac{1}{2}b^2$, to the right-hand side to arrive at

$$\mathbb{E}\left[|X_s-\tilde{X}_s|^2\big|\mathcal{F}_t\right] \leq K_5\mathbb{E}\left[\int_t^s |X_r-\tilde{X}_r|^2\,\mathrm{d}r\right]+K_6(s-t)\|u-\tilde{u}\|_{t,\infty}^2.$$

Finally, the Gronwall lemma yields

$$\mathbb{E}\left[|X_s-\tilde{X}_s|^2\big|\mathcal{F}_t\right] \leq e^{K_5(s-t)}K_6(s-t)\|u-\tilde{u}\|_{t,\infty}^2.$$

Notice that the right-hand side of the previous inequality does not depend on ω; hence it yields (30) after taking the square root on both sides. □

Lemma 6.6. *Assume that $\xi,\tilde{\xi},\mu,\sigma,f$ satisfy the assumptions of Lemma 6.4. Then there exists a constant $C>0$, depending only on $T-t$, B, $\|\mu\|_\infty$, $\|\sigma\|_\infty$ and the Lipschitz constant of (μ,σ,f), such that the condition $T-t\leq\frac{1}{C}$ implies*

$$\|u-\tilde{u}\|_{t,\infty}\leq \|\xi-\tilde{\xi}\|_\infty\left(1+C(T-t)\right).$$

Proof. Notice that Lemma 6.4 and Lemma 6.5 imply that there exists a constant $K_1 > 0$ such that

$$|u(t,x) - \tilde{u}(t,x)| = |Y_t - \tilde{Y}_t| \leq \|\xi - \tilde{\xi}\|_\infty + K_1(T-t)\|u - \tilde{u}\|_{t,\infty}.$$

Let $r \in [t,T]$. By using the same arguments to the FBSDE solved on the interval $[r,T]$ with initial condition $X_r = x$ we can derive the estimate

$$|u(r,x) - \tilde{u}(r,x)| \leq \|\xi - \tilde{\xi}\|_\infty + K_1(T-r)\|u - \tilde{u}\|_{r,\infty} \leq \|\xi - \tilde{\xi}\|_\infty + K_1(T-t)\|u - \tilde{u}\|_{t,\infty}.$$

Consequently,

$$\|u - \tilde{u}\|_{t,\infty} \leq \|\xi - \tilde{\xi}\|_\infty + K_1(T-t)\|u - \tilde{u}\|_{t,\infty}.$$

If $T - t \leq \frac{1}{2K_1}$, then

$$\|u - \tilde{u}\|_{t,\infty} \leq \frac{1}{1 - K_1(T-t)}\|\xi - \tilde{\xi}\|_\infty \leq (1 + 2K_1(T-t))\|\xi - \tilde{\xi}\|_\infty,$$

and hence the claim is true for $C = 2K_1$. □

References

[Ant93] Fabio Antonelli, *Backward-forward stochastic differential equations*, Ann. Appl. Probab. **3** (1993), no. 3, 777–793.

[BS12] Christian Bender and Jessica Steiner, *Least-squares Monte Carlo for backward SDEs*, Numerical methods in finance, Springer Proc. Math., vol. 12, Springer, Heidelberg, 2012, pp. 257–289. MR 3287774

[BT04] Bruno Bouchard and Nizar Touzi, *Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations*, Stochastic Process. Appl. **111** (2004), no. 2, 175–206. MR 2056536

[BZ08] Christian Bender and Jianfeng Zhang, *Time discretization and Markovian iteration for coupled FBSDEs*, The Annals of Applied Probability **18** (2008), no. 1, 143–177.

[CM14] Dan Crisan and Konstantinos Manolarakis, *Second order discretization of backward sdes and simulation with the cubature method*, The Annals of Applied Probability **24** (2014), no. 2, 652–678.

[Del02] François Delarue, *On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case*, Stochastic Process. Appl. **99** (2002), no. 2, 209–286.

[Fro15] Alexander Fromm, *Theory and applications of decoupling fields for forward-backward stochastic differential equations*, Ph.D. thesis, Humboldt-Universität zu Berlin, 2015.

[Fro18] ———, *The method of decoupling fields generalized to higher spatial derivatives*, Preprint arXiv:1804.10970 (2018).

[MPY94] Jin Ma, Philip Protter, and Jiongmin Yong, *Solving forward-backward stochastic differential equations explicitly – a four step scheme*, Probab. Theory Relat. Fields **98** (1994), no. 3, 339–359.

[MWZZ15] Jin Ma, Zhen Wu, Detao Zhang, and Jianfeng Zhang, *On Wellposedness of Forward-Backward SDEs — A Unified Approach*, Ann. Appl. Probab. **25** (2015), no. 4.

[MY99] Jin Ma and Jiongmin Yong, *Forward-backward stochastic differential equations and their applications*, Lecture Notes in Mathematics, vol. 1702, Springer-Verlag, Berlin, 1999. MR 1704232

[PT99] Etienne Pardoux and Shanjian Tang, *Forward-backward stochastic differential equations and quasilinear parabolic PDEs*, Probab. Theory Relat. Fields **114** (1999), no. 2, 123–150.

[Zha06] Jianfeng Zhang, *The wellposedness of FBSDEs*, Discrete Contin. Dyn. Syst. Ser. B **6** (2006), no. 4, 927–940. MR 2223916