Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Time discretization and Markovian iteration for coupled FBSDEs

Christian Bender[1], Jianfeng Zhang[2],

submitted: 25 August 2006

[1] Weierstrass Institute
for Applied Analysis and Stochastics
Mohrenstr. 39
D-10117 Berlin
Germany
E-mail: bender@wias-berlin.de

[2] Department of Mathematics
University of Southern California
Los Angeles, California 90089
USA
Email: jianfenz@usc.edu

No. 1160
Berlin 2006

2000 *Mathematics Subject Classification.* 65C30, 60H10, 60H30, 65C05.

Key words and phrases. Forward backward SDE, numerics, time discretization, Monte Carlo simulation.

J.Z. is supported in part by NSF grant DMS-0403575.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

In this paper we lay the foundation for a numerical algorithm to simulate high-dimensional coupled FBSDEs under weak coupling or monotonicity conditions. In particular we prove convergence of a time discretization and a Markovian iteration. The iteration differs from standard Picard iterations for FBSDEs in that the dimension of the underlying Markovian process does not increase with the number of iterations. This feature seems to be indispensable for an efficient iterative scheme from a numerical point of view. We finally suggest a fully explicit numerical algorithm and present some numerical examples with up to 10-dimensional state space.

1 Introduction

Motivated by the aim to simulate high dimensional coupled forward backward stochastic differential equations (FBSDEs) we study a time discretization and a Markovian iteration for equations of the form

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s)ds + \int_0^t \sigma(s, X_s, Y_s)dW_s; \\ Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s; \end{cases} \tag{1.1}$$

where b, σ, f, g are deterministic and Lipschitz continuous functions which are additionally supposed to satisfy some weak coupling or monotonicity condition. Note that (1.1) is not in its most general form, since Z does not couple into the forward SDE.

Most of the numerical algorithms for coupled FBSDEs, with the notably exception of Delarue and Menozzi (2006), exploit the relation to quasi-linear parabolic PDEs via the four-step-scheme (Ma et al., 1994). Under appropriate conditions (X, Y, Z) are connected by

$$Y_t = u(t, X_t); \quad Z_t = v(t, X_t) \stackrel{\triangle}{=} u_x(t, X_t)\sigma(t, X_t, u(t, X_t)). \tag{1.2}$$

where u is a classical solution of the PDE

$$\begin{cases} u_t + \frac{1}{2}\text{trace}(\sigma\sigma^*(t, x, u)u_{xx}) + u_x b(t, x, u) + f(t, x, u, u_x\sigma(t, x, u)) = 0; \\ u(T, x) = g(x). \end{cases} \tag{1.3}$$

The main focus in these approaches is on the numerical solution of the PDE (1.3), see Douglas et al. (1996), Milstein and Tretyakov (2006), and Ma et al. (2006). Since the PDE approach requires existence of a classical solution to (1.3), there is typically need for some smoothness, boundedness, and regularity conditions such as uniform ellipticity of the differential operator. Moreover, solving (1.3) numerically by standard PDE techniques becomes more difficult, if not impossible, with increasing spatial dimension. To overcome these limitations it seems necessary to tackle the FBSDE (1.1) directly by probabilistic means.

A natural time discretization of equation (1.1) is

$$\begin{cases} X_0^n \triangleq x; \\ X_{i+1}^n \triangleq X_i^n + b(t_i, X_i^n, Y_i^n)h + \sigma(t_i, X_i^n, Y_i^n)\Delta W_{i+1}; \\ Y_n^n \triangleq g(X_n^n); \\ \hat{Z}_i^n \triangleq \dfrac{1}{h} E_{t_i}\{Y_{i+1}^n \Delta W_{i+1}\}; \\ Y_i^n \triangleq E_{t_i}\{Y_{i+1}^n + f(t_i, X_i^n, Y_{i+1}^n, \hat{Z}_i^n)h\}; \end{cases} \tag{1.4}$$

where $h \triangleq \frac{T}{n}$ and $t_i \triangleq ih, i = 0, 1 \cdots, n$, and $\Delta W_{i+1} \triangleq W_{t_{i+1}} - W_{t_i}$. Here, of course, E_{t_i} denotes the conditional expectation $E\{\cdot|\mathcal{F}_{t_i}\}$. This time discretization was investigated in detail by Zhang (2004) for decoupled FBSDEs. However, since X is discretized forwardly and Y is discretized backwardly, (1.4) is by no means an explicit discretization in the present situation due to the coupling. Note that one can rewrite

$$Y_i^n = u_i^n(X_i^n); \quad \hat{Z}_i^n = v_i^n(X_i^n), \tag{1.5}$$

where

$$\begin{cases} u_n^n(x) \triangleq g(x); \\ X_{i+1}^{n,i,x} \triangleq x + b(t_i, x, u_i^n(x))h + \sigma(t_i, x, u_i^n(x))\Delta W_{i+1}; \\ Y_{i+1}^{n,i,x} \triangleq u_{i+1}^n(X_{i+1}^{n,i,x}); \\ v_i^n(x) \triangleq \frac{1}{h} E\{Y_{i+1}^{n,i,x} \Delta W_{i+1}\}; \\ u_i^n(x) \triangleq E\Big\{Y_{i+1}^{n,i,x} + f(t_i, x, Y_{i+1}^{n,i,x}, v_i^n(x))h\Big\}. \end{cases} \tag{1.6}$$

Equation (1.6) is still implicit in u_i^n, but truly backwards in time. Combined with a local updating technique it serves as starting point for the probabilistic scheme in Delarue and Menozzi (2006). This type of scheme requires, however, apart from estimating the expectations, a discretization of the state space. Such space discretization may again become prohibitive, when the dimension increases.

We, hence, propose to combine the time discretization (1.4) with an iterative scheme. Indeed, it is known from results by Antonelli (1993) and Pardoux and Tang (1999) that under weak coupling or monotonicity conditions (1.1) has a unique solution (X,Y,Z) which can be constructed via a Picard iteration

$$\begin{cases} \check{X}_t^m = x + \int_0^t b(s,\check{X}_s^m,\check{Y}_s^{m-1})ds + \int_0^t \sigma(s,\check{X}_s^m,\check{Y}_s^{m-1})dW_s; \\ \check{Y}_t^m = g(\check{X}_T^m) + \int_t^T f(s,\check{X}_s^m,\check{Y}_s^m,\check{Z}_s^m)ds - \int_t^T \check{Z}_s^m dW_s. \end{cases} \tag{1.7}$$

The drawback of (1.7) is that the dimension of the underlying Markovian process $(\check{X}^1,\dots,\check{X}^m)$ increases with the number of iterations, and, consequently, $\check{Y}_t^m$ is a function of time and $(\check{X}^1,\dots,\check{X}^m)$. This renders a combination of (1.4) with a Picard iteration like (1.7), which was recently suggested by Riviere (2005) in theory, impractical from a numerical point of view. The stochastic control approach in Cvitanić and Zhang (2005) faces the same kind of difficulty.

In this paper we introduce an alternative iteration in a way that the dimension of the underlying Markovian process does not change in the number of iterations. It reads, in discretized form, $u_i^{n,0}(x)=0$, and

$$\begin{cases} X_0^{n,m} \triangleq x; \\ X_{i+1}^{n,m} \triangleq X_i^{n,m} + b(t_i,X_i^{n,m},u_i^{n,m-1}(X_i^{n,m}))h + \sigma(t_i,X_i^{n,m},u_i^{n,m-1}(X_i^{n,m}))\Delta W_{i+1}; \\ Y_n^{n,m} \triangleq g(X_n^{n,m}); \\ \hat{Z}_i^{n,m} \triangleq \frac{1}{h}E_{t_i}\left\{Y_{i+1}^{n,m}\Delta W_{i+1}\right\}; \\ Y_i^{n,m} \triangleq E_{t_i}\{Y_{i+1}^{n,m} + f(t_i,X_i^{n,m},Y_{i+1}^{n,m},\hat{Z}_i^{n,m})h\}; \\ u^{n,m}(X_i^{n,m}) = Y_i^{n,m}. \end{cases} \tag{1.8}$$

The main advantage is that here $Y_i^{n,m}$ is a function of time and $X_i^{n,m}$ but does not depend on $(X_i^{n,\mu},\ \mu=1,\dots,m-1)$. Establishing the convergence of this new 'Markovian' iteration turns out to be more involved than for the standard Picard iteration, because controlling the Lipschitz constant and the linear growth of $u_i^{n,m}(x)$ uniformly in i,n,m becomes crucial. This is indeed the reason, why we cannot allow Z to couple in the forward SDE at the current state of our research.

We also indicate how this discretized Markovian iteration may be transformed into a viable numerical scheme, replacing the conditional expectations by simulation based least squares regression and estimating $u^{n,m}$ this way. Such estimator was introduced by Carrière (1996) and Longstaff and Schwartz (2001) in the context of American options and is applied by Gobet et al. (2005) and Bender and Denk (2005) for decoupled FBSDEs. Although a convergence analysis for this estimator in the present context of a coupled FBSDE is beyond the scope of this paper, we illustrate by some examples with up to 10-dimensional state space that the proposed numerical algorithm works in practice.

The paper is organized as follows: In Section 2 we state the main results on convergence of the discretized Markovian iteration. The proof is given in several steps in Sections 3–5, where we establish the control of the Lipschitz constant, of the linear growth, and the convergence of $u^{n,m}$ to u^n respectively. In Section 6 we investigate the error due to the time discretization. To the best of our knowledge our convergence theorem is the first of this type for coupled FBSDEs which also holds for a degenerate diffusion coefficient σ. In Section 7 we spell out the proposed numerical scheme and present some numerical examples in Section 8.

2 Notations and Main Results

The main results of this paper estimate the error of the discretized Markovian iteration (1.8) as the number of time steps n and the number of iterations m tend to infinity. Before we can state these results, we need to fix some notations and discuss some assumptions. From now on we suppose, in the theoretical part, that all processes are one-dimensional. This is only to ease the notation and the attentive reader will easily see that all results hold true for the multi-dimensional case as well. The augmented filtration generated by the Brownian motion is denoted by $\mathbf{F} = \{\mathcal{F}_t,\ 0 \le t \le T\}$.

The first assumption concerns the Lipschitz continuity and monotonicity of the coefficients. It will be in force throughout the whole paper without further notice.

Assumption 1 *(i) There exist (possibly negative) constants k_b, k_f such that*

$$[b(t,x_1,y) - b(t,x_2,y)][x_1 - x_2] \le k_b|x_1 - x_2|^2;$$
$$[f(t,x,y_1,z) - f(t,x,y_2,z)][y_1 - y_2] \le k_f|y_1 - y_2|^2.$$

(ii) The coefficients b, σ, f, g are uniformly Lipschitz continuous with respect to (x,y,z). In particular, there are constants $K, b_y, \sigma_x, \sigma_y, f_x, f_z$, and g_x such that

$$|b(t,x_1,y_1) - b(t,x_2,y_2)|^2 \le K|x_1 - x_2|^2 + b_y|y_1 - y_2|^2;$$
$$|\sigma(t,x_1,y_1) - \sigma(t,x_2,y_2)| \le \sigma_x|x_1 - x_2|^2 + \sigma_y|y_1 - y_2|^2;$$
$$|f(t,x_1,y_1,z_1) - f(t,x_2,y_2,z_2)| \le f_x|x_1 - x_2|^2 + K|y_1 - y_2|^2 + f_z|z_1 - z_2|^2;$$
$$|g(x_1) - g(x_2)|^2 \le g_x|x_1 - x_2|^2.$$

(iii) $b(t,0,0), \sigma(t,0,0), f(t,0,0,0)$ are bounded. In particular, there are constants

b_0, σ_0, f_0, and g_0 such that

$$
\begin{aligned}
&|b(t,x,y)|^2 \le b_0 + K|x|^2 + b_y|y|^2;\\
&|\sigma(t,x,y)|^2 \le \sigma_0 + \sigma_x|x|^2 + \sigma_y|y|^2;\\
&|f(t,x,y,z)|^2 \le f_0 + f_x|x|^2 + K|y|^2 + f_z|z|^2;\\
&|g(x)|^2 \le g_0 + g_x|x|^2.
\end{aligned}
$$

We emphasize that here b_y et al are constants, not partial derivatives. Indeed, we will not assume any differentiability conditions throughout this paper. For convenience we also suppose that K is an upper bound for all the constants above.

For results concerning the error due to the time discretization we require the following assumption.

Assumption 2 *The coefficients (b, σ, f) are uniformly Hölder-$\frac{1}{2}$ continuous with respect to t.*

If Assumption 2 is in force, we use the same constant K to denote the square of the Hölder constants.

To ensure that the iteration converges we further need to impose conditions which guarantee that we are in one of the following five cases:

1. *Small time duration*, i.e. T is small.
2. *Weak coupling of Y into the forward SDE*, i.e. b_y and σ_y are small. In particular, if $b_y = \sigma_y = 0$, then the forward equation in (1.1) does not depend on the backward one and thus (1.1) is decoupled.
3. *Weak coupling of X into the backward SDE*, i.e. f_x and g_x are small. In particular, if $f_x = g_x = 0$, then the backward equation in (1.1) does not depend on the forward one and thus (1.1) is also decoupled. In fact in this case $Z = 0$ and (1.1) reduces to a decoupled system of ordinary differential equations.
4. *f is strongly decreasing in y*, i.e. k_f is very negative.
5. *b is strongly decreasing in x*, i.e. k_b is very negative.

The above conditions will be made precise later. Generically, we will derive the following theorems. The first theorem concerns the convergence of the iteration as m tends to infinity.

Theorem 2.1 *Under Assumption 1 let one of the conditions 1.–5. hold true. Then, for sufficiently small h, (1.6) has an 'essentially' unique solution u^n with linear growth and there are constants $C > 0$ and $0 < c < 1$ such that*

$$\max_{0 \le i \le n} |u_i^{n,m}(x) - u_i^n(x)|^2 \le C(|x|^2 + m)c^m,$$

where $u^{n,m}$ is given by (1.8).

Concerning the error due to the time discretization we obtain:

Theorem 2.2 *Suppose Assumptions 1, 2 and one of the conditions 1.–5. is in force. Then equation (1.3) admits a viscosity solution $u(t,x)$ with linear growth and there is a constant $C > 0$ such that for sufficiently small h,*

$$\max_{0 \le i \le n} |u_i^n(x) - u(t_i, x)|^2 \le C(1 + |x|^2)h$$

Combining these two theorems one can derive with a little extra effort:

Theorem 2.3 *Under the assumptions of Theorem 2.2 FBSDE (1.1) has a unique solution (X, Y, Z) and there are constants $C > 0$ and $0 < c < 1$ such that for sufficiently small h,*

$$\sup_{1 \le i \le n} E\Big\{ \sup_{t \in [t_{i-1}, t_i]} [|X_t - X_{i-1}^{n,m}|^2 + |Y_t - Y_{i-1}^{n,m}|^2] \Big\} + \sum_{i=1}^{n} E\Big\{ \int_{t_{i-1}}^{t_i} |Z_t - \hat{Z}_{i-1}^{n,m}|^2 dt \Big\}$$
$$\le C(1 + |x|^2)[mc^m + h].$$

These generic results will be made precise in Theorems 5.1, 6.4, and 6.6 below.

We emphasize that none of the above theorems requires non-degeneracy of σ and, in principle, X and W can have different dimensions. Moreover, we do not suppose any smoothness or boundedness conditions. However, we also underline again that FBSDE (1.1) does not allow coupling through the control part Z.

Before we turn to the proofs in next sections we first explain an additional difficulty that one faces when proving convergence of this Markovian iteration. In a standard Picard iteration, like (1.7), one estimates $|\check{Y}^{m+1} - \check{Y}^m|$ in terms of $|\check{X}^{m+1} - \check{X}^m|$ and then $|\check{X}^{m+1} - \check{X}^m|$ in terms of $|\check{Y}^m - \check{Y}^{m-1}|$. However, applying similar techniques to (1.8) yields only estimates of $|X^{n,m+1} - X^{n,m}|$ in terms of $|u^{n,m}(X^{n,m+1}) - u^{n,m-1}(X^{n,m})|$. Since $Y^{n,m} = u^{n,m}(X^{n,m})$, it seems unavoidable to control the Lipschitz constant of $u^{n,m}$ to obtain estimates in terms of $|Y^{n,m} - Y^{n,m-1}|$.

To study the behavior of the functions $u^{n,m}$, we introduce an important operator F_n for each n. For any measurable functions $\varphi = \{\varphi_i\}_{0\le i\le n-1}$, define ψ and Φ as follows.

$$
\begin{cases}
\Phi_n(x) \triangleq g(x); \\
X_{i+1}^{\varphi,i,x} \triangleq x + b(t_i, x, \varphi_i(x))h + \sigma(t_i, x, \varphi_i(x))\Delta W_{i+1}; \\
Y_{i+1}^{\varphi,i,x} \triangleq \Phi_{i+1}(X_{i+1}^{\varphi,i,x}); \\
\psi_i(x) \triangleq \frac{1}{h} E\Big\{ Y_{i+1}^{\varphi,i,x} \Delta W_{i+1} \Big\}; \\
\Phi_i(x) \triangleq E\Big\{ Y_{i+1}^{\varphi,i,x} + f(t_i, x, Y_{i+1}^{\varphi,i,x}, \psi_i(x))h \Big\}.
\end{cases}
\tag{2.1}
$$

We then set $F_n(\varphi) \triangleq \Phi$. It is then obvious that $u^{n,m} = F_n(u^{n,m-1})$, and $F_n(u^n) = u^n$ if (1.6) has a solution u^n. We also point out that $Y^{n,m}$, given by (1.8), can be expressed in the form

$$
Y_i^{n,m} = Y_{i+1}^{n,m} + f(t_i, X_i^{n,m}, Y_{i+1}^{n,m}, \hat{Z}_i^{n,m})h - \int_{t_i}^{t_{i+1}} Z_t^{n,m} dW_t \tag{2.2}
$$

thanks to the martingale representation theorem. The analogous expression holds for Y^n defined in (1.4).

3 Lipschitz Continuity

In this section we obtain a uniform (in i, n, m) Lipschitz constant of $u_i^{n,m}(x)$. To this end we first investigate the Lipschitz continuity of $F_n(\varphi)$. Given Lipschitz continuous φ, let $L(\varphi_i)$ denote the square of a Lipschitz constant of φ_i, and $L(\varphi) \triangleq \sup_i L(\varphi_i)$. Our aim is to derive the following theorem:

Theorem 3.1 *Denote*

$$
\begin{aligned}
L_0 &\triangleq [b_y + \sigma_y][g_x + f_x T] T e^{[b_y+\sigma_y][g_x+f_xT]T + [2k_b+2k_f+2+\sigma_x+f_z]T}; \\
L_1 &\triangleq [g_x + f_x T]\Big[e^{[b_y+\sigma_y][g_x+f_xT]T + [2k_b+2k_f+2+\sigma_x+f_z]T+1} \vee 1 \Big];
\end{aligned}
\tag{3.1}
$$

If

$$
L_0 < e^{-1}, \tag{3.2}
$$

then for any $L > L_1$ and for h small enough, we have

$$
L(u^{n,m}) \le L, \quad \forall m.
$$

Notice that (3.2) holds true in all five cases of Section 2. We prepare the proof of Theorem 3.1 with several lemmas.

Lemma 3.2 *Fix i and for $l = 1, 2$, let*

$$X^l_{i+1} \stackrel{\triangle}{=} X^l_i + b(t_i, X^l_i, \varphi^l(X^l_i))h + \int_{t_i}^{t_{i+1}} \alpha^l_t dt + [\sigma(t_i, X^l_i, \varphi^l(X^l_i))\Delta W_{i+1} + \int_{t_i}^{t_{i+1}} \beta^l_t dW_t,$$

where X^l_i is $\mathcal{F}_{t_i}$-measurable and α^l_t, β^l_t are $\mathbf{F}$-adapted. Assume φ^1 is uniformly Lipschitz continuous. Then for any $\lambda_j > 0$,

$$\begin{aligned} E_{t_i}\{|\Delta X_{i+1}|^2\} \quad \leq \quad & [1 + A_1 h + (1+\lambda_2)A_2 h L(\varphi^1)]|\Delta X_i|^2 + (1+\lambda_2^{-1})A_2 h|\Delta\varphi(X^2_i)|^2 \\ & +2(1+\lambda_1^{-1})E_{t_i}\Big\{\int_{t_i}^{t_{i+1}} [|\Delta\alpha_t|^2 + |\Delta\beta_t|^2]dt\Big\} \end{aligned}$$

where

$$\Delta X \stackrel{\triangle}{=} X^1 - X^2; \quad \Delta\alpha \stackrel{\triangle}{=} \alpha^1 - \alpha^2; \quad \Delta\beta \stackrel{\triangle}{=} \beta^1 - \beta^2; \quad |\Delta\varphi| \stackrel{\triangle}{=} |\varphi^1 - \varphi^2|.$$

and

$$\begin{aligned} A_1 &\stackrel{\triangle}{=} \lambda_1 + (1+\lambda_1 h)(2k_b + 1 + Kh) + (1+\lambda_1)\sigma_x; \\ A_2 &\stackrel{\triangle}{=} (1+\lambda_1 h)b_y + Kh + (1+\lambda_1)\sigma_y. \end{aligned} \tag{3.3}$$

Proof. Denote, for $\phi = b, \sigma$,

$$\Delta\phi \stackrel{\triangle}{=} \phi(t_i, X^1_i, \varphi^1(X^1_i)) - \phi(t_i, X^2_i, \varphi^2(X^2_i)).$$

Then

$$\Delta X_{i+1} = \Delta X_i + \Delta b h + \int_{t_i}^{t_{i+1}} \Delta\alpha_t dt + \Delta\sigma\Delta W_{i+1} + \int_{t_i}^{t_{i+1}} \Delta\beta_t dW_t.$$

Thus, for h small enough,

$$\begin{aligned}
& E_{t_i}\{|\Delta X_{i+1}|^2\} \\
&= E_{t_i}\Big\{(\Delta X_i + \Delta b h)^2 + |\int_{t_i}^{t_{i+1}} \Delta\alpha_t dt|^2 + 2(\Delta X_i + \Delta b h)\int_{t_i}^{t_{i+1}} \Delta\alpha_t dt \\
&\quad + \int_{t_i}^{t_{i+1}} |\Delta\sigma + \Delta\beta_t|^2 dt + 2\Delta\sigma\Delta W_{i+1}\int_{t_i}^{t_{i+1}} \Delta\alpha_t dt \\
&\quad +2\int_{t_i}^{t_{i+1}} \Delta\beta_t dW_t \int_{t_i}^{t_{i+1}} \Delta\alpha_t dt\Big\} \\
&\le (1+\lambda_1 h)(\Delta X_i + \Delta b h)^2 + (1 + \frac{\lambda_1}{2} + \frac{\lambda_1}{2})|\Delta\sigma|^2 h \\
&\quad +E_{t_i}\Big\{[1 + \frac{1}{\lambda_1 h} + \frac{2}{\lambda_1} + 1]|\int_{t_i}^{t_{i+1}} \Delta\alpha_t dt|^2 + [1 + \frac{2}{\lambda_1} + 1]\int_{t_i}^{t_{i+1}} |\Delta\beta_t|^2 dt\Big\} \\
&\le (1+\lambda_1 h)[|\Delta X_i|^2 + 2\Delta X_i \Delta b h + |\Delta b h|^2] + (1+\lambda_1)|\Delta\sigma|^2 h \\
&\quad +2(1+\lambda_1^{-1})E_{t_i}\Big\{\int_{t_i}^{t_{i+1}} [|\Delta\alpha_t|^2 + |\Delta\beta_t|^2] dt\Big\} \\
&\le (1+\lambda_1 h)\Big[|\Delta X_i|^2 + 2k_b|\Delta X_i|^2 h + |\Delta X_i|^2 h + b_y|\varphi^1(X_i^1) - \varphi^2(X_i^2)|^2 h \\
&\quad +K[|\Delta X_i|^2 + |\varphi^1(X_i^1) - \varphi^2(X_i^2)|^2]h^2\Big] \\
&\quad +(1+\lambda_1)\Big[\sigma_x|\Delta X_i|^2 + \sigma_y|\varphi^1(X_i^1) - \varphi^2(X_i^2)|^2\Big]h \\
&\quad +2(1+\lambda_1^{-1})E_{t_i}\Big\{\int_{t_i}^{t_{i+1}} [|\Delta\alpha_t|^2 + |\Delta\beta_t|^2] dt\Big\} \\
&= [1 + A_1 h]|\Delta X_i|^2 + A_2 h|\varphi^1(X_i^1) - \varphi^2(X_i^2)|^2 \\
&\quad +2(1+\lambda_1^{-1})E_{t_i}\Big\{\int_{t_i}^{t_{i+1}} [|\Delta\alpha_t|^2 + |\Delta\beta_t|^2] dt\Big\}.
\end{aligned}$$

Note that

$$|\varphi^1(X_i^1) - \varphi^2(X_i^2)| \le |\varphi^1(X_i^1) - \varphi^1(X_i^2)| + |\varphi^1(X_i^2) - \varphi^2(X_i^2)|.$$

Then

$$|\varphi^1(X_i^1) - \varphi^2(X_i^2)|^2 \le [1+\lambda_2]L(\varphi^1)|\Delta X_i|^2 + [1+\lambda_2^{-1}]|\Delta\varphi(X_i^2)|^2.$$

Hence

$$\begin{aligned}
E_{t_i}\{|\Delta X_{i+1}|^2\} \quad \le \quad & [1 + A_1 h + (1+\lambda_2)A_2 h L(\varphi^1)]|\Delta X_i|^2 + (1+\lambda_2^{-1})A_2 h|\Delta\varphi(X_i^2)|^2 \\
& +2(1+\lambda_1^{-1})E\Big\{\int_{t_i}^{t_{i+1}} [|\Delta\alpha_t|^2 + |\Delta\beta_t|^2] dt\Big\},
\end{aligned}$$

and the lemma is proved. ■

Lemma 3.3 *Fix i and for $l = 1, 2$, let*

$$Y_i^l = Y_{i+1}^l + f(t_i, X_i^l, Y_{i+1}^l, \hat{Z}_i^l)h + \gamma_{i+1}^1 h - \int_{t_i}^{t_{i+1}} Z_t^l dW_t;$$

where

$$\hat{Z}_i^l \triangleq \frac{1}{h} E_{t_i}\Big\{ Y_{i+1}^l \Delta W_{i+1} \Big\}.$$

Then for any $\lambda_j > 0$,

$$\begin{aligned}
&|\Delta Y_i|^2 + (1 - A_3)h|\Delta \hat{Z}_i|^2 \\
&\le E_{t_i}\Big\{ (1 + A_4 h)|\Delta Y_{i+1}|^2 + A_5 h|\Delta X_i|^2 + (\lambda_1^{-1} + h + 2\lambda_3^{-1}h)h|\Delta\gamma_{i+1}|^2 \Big\};
\end{aligned}$$

where

$$\Delta X \triangleq X^1 - X^2; \quad \Delta Y \triangleq Y^1 - Y^2; \quad \Delta\hat{Z} \triangleq \hat{Z}^1 - \hat{Z}^2; \quad \Delta\gamma \triangleq \gamma^1 - \gamma^2;$$

and

$$\begin{aligned}
A_3 &\triangleq \lambda_3 + (1 + \lambda_1 h)\lambda_4 + (1 + \lambda_1 h + 2\lambda_3^{-1})Kh; \\
A_4 &\triangleq \lambda_1 + (1 + \lambda_1 h)(2k_f + 1 + \lambda_4^{-1} f_z) + (1 + \lambda_1 h + 2\lambda_3^{-1})Kh; \\
A_5 &\triangleq (1 + \lambda_1 h)f_x + (1 + \lambda_1 h + 2\lambda_3^{-1})Kh.
\end{aligned} \tag{3.4}$$

Proof. Denote

$$\Delta Z \triangleq Z^1 - Z^2; \quad \Delta f \triangleq f(t_i, X_i^1, Y_{i+1}^1, \hat{Z}_i^1) - f(t_i, X_i^2, Y_{i+1}^2, \hat{Z}_i^2).$$

Then

$$\Delta Y_i + \int_{t_i}^{t_{i+1}} \Delta Z_t dW_t = \Delta Y_{i+1} + \Delta f h + h\Delta\gamma_{i+1}.$$

Thus

$$\begin{aligned}
&|\Delta Y_{t_i}|^2 + E_{t_i}\Big\{ \int_{t_i}^{t_{i+1}} |\Delta Z_t|^2 dt \Big\} \\
&\le E_{t_i}\Big\{ (1 + \lambda_1 h)[\Delta Y_{t_{i+1}} + \Delta f h]^2 + (1 + \frac{1}{\lambda_1 h})|h\Delta\gamma_{i+1}|^2 \Big\} \\
&\le E_{t_i}\Big\{ (1 + \lambda_1 h)[|\Delta Y_{t_{i+1}}|^2 + 2\Delta Y_{t_{i+1}}\Delta f h + |\Delta f|^2 h^2] + (\lambda_1^{-1} + h)h|\Delta\gamma_{i+1}|^2 \Big\}.
\end{aligned}$$

Note that

$$E_{t_i}\Big\{ \int_{t_i}^{t_{i+1}} |\Delta Z_t|^2 dt \Big\} \ge \frac{1}{h} |E_{t_i}\Big\{ \int_{t_i}^{t_{i+1}} \Delta Z_t dt \Big\}|^2;$$

and that

$$\begin{aligned}
E_{t_i}\Big\{ \int_{t_i}^{t_{i+1}} \Delta Z_t dt \Big\} &= E_{t_i}\Big\{ \Big[\Delta Y_{t_{i+1}} + \Delta f h + h\Delta\gamma_{i+1}\Big]\Delta W_{i+1} \Big\} \\
&= h\Delta\hat{Z}_i + hE_{t_i}\{[\Delta f + \Delta\gamma_{i+1}]\Delta W_{i+1}\Big\}.
\end{aligned}$$

Then

$$E_{t_i}\Big\{\int_{t_i}^{t_{i+1}}|\Delta Z_t|^2dt\Big\} \geq (1-\lambda_3)h|\Delta\hat{Z}_i|^2 - \frac{h}{\lambda_3}|E_{t_i}\Big\{[\Delta f+\Delta\gamma_{i+1}]\Delta W_{i+1}\Big\}|^2$$
$$\geq (1-\lambda_3)h|\Delta\hat{Z}_i|^2 - 2\lambda_3^{-1}h^2E_{t_i}\Big\{|\Delta f|^2+|\Delta\gamma_{i+1}|^2\Big\}.$$

Thus

$$\begin{aligned}
&|\Delta Y_i|^2+(1-\lambda_3)h|\Delta\hat{Z}_i|^2\\
&\leq E_{t_i}\Big\{(1+\lambda_1 h)[|\Delta Y_{i+1}|^2+2\Delta Y_{i+1}\Delta f h]+(1+\lambda_1 h+2\lambda_3^{-1})h^2|\Delta f|^2\\
&\quad+(\lambda_1^{-1}+h+2\lambda_3^{-1}h)h|\Delta\gamma_{i+1}|^2\Big\}\\
&\leq E_{t_i}\Big\{(1+\lambda_1 h)\Big[|\Delta Y_{i+1}|^2+2k_f h|\Delta Y_{i+1}|^2+|\Delta Y_{i+1}|^2h+f_x|\Delta X_i|^2h+\lambda_4|\Delta\hat{Z}_i|^2h\\
&\quad+\lambda_4^{-1}f_z|\Delta Y_{i+1}|^2h\Big]+(1+\lambda_1 h+2\lambda_3^{-1})h^2K[|\Delta X_i|^2+|\Delta Y_{i+1}|^2+|\Delta\hat{Z}_i|^2]\\
&\quad+(\lambda_1^{-1}+h+2\lambda_3^{-1}h)h|\Delta\gamma_{i+1}|^2\Big\};
\end{aligned}$$

which implies the lemma immediately. ■

With these lemmas at hand we can study the Lipschitz continuity of $F_n(\varphi)$ given Lipschitz continuous φ.

Theorem 3.4 *For any Lipschitz continuous φ, we have*

$$L(F_n(\varphi)) \leq [g_x+A_5T]\Big[\exp\Big([A_1+A_4+A_1A_4h]T+[A_2+A_2A_4h]TL(\varphi)\Big)\vee 1\Big];$$

where $\lambda_1=\lambda_2=0$ and $\lambda_3,\lambda_4>0$ are chosen such that

$$A_3 \leq 1. \tag{3.5}$$

Proof. Recall (2.1). Fix i and x_1,x_2. Denote

$$\Delta x \stackrel{\triangle}{=} x_1-x_2;\quad \Delta X \stackrel{\triangle}{=} X^{\varphi,i,x_1}-X^{\varphi,i,x_2};\quad \Delta Y \stackrel{\triangle}{=} Y^{\varphi,i,x_1}-Y^{\varphi,i,x_2};$$
$$\Delta\Phi_i \stackrel{\triangle}{=} \Phi_i(x_1)-\Phi_i(x_2);\quad \Delta\psi_i \stackrel{\triangle}{=} \psi_i(x_1)-\psi_i(x_2).$$

We apply Lemmas 3.2 and 3.3, setting $\lambda_1=\lambda_2=0$, and obtain

$$E\{|\Delta X_{i+1}|^2\} \leq [1+A_1h+A_2hL(\varphi)]|\Delta x|^2;$$
$$|\Delta\Phi_i|^2+(1-A_3)h|\Delta\psi_i|^2 \leq (1+A_4h)E\{|\Delta Y_{i+1}|^2\}+A_5h|\Delta x|^2;$$

By (3.5) we have

$$|\Delta\Phi_i|^2 \leq [1+A_4h]L(\Phi_{i+1})E\{|\Delta X_{i+1}|^2\}+A_5h|\Delta x|^2$$
$$\leq [1+A_4h][1+A_1h+A_2hL(\varphi)]L(\Phi_{i+1})|\Delta x|^2+A_5h|\Delta x|^2.$$

Thus

$$L(\Phi_i) \le [1 + A_4 h][1 + A_1 h + A_2 h L(\varphi)]L(\Phi_{i+1}) + A_5 h$$
$$\triangleq [1 + \tilde{A}h]L(\Phi_{i+1}) + A_5 h \le [1 + \tilde{A}^+ h]L(\Phi_{i+1}) + A_5 h; \quad (3.6)$$

where $\tilde{A}^+ \triangleq \tilde{A} \vee 0$ and

$$\tilde{A} \triangleq A_1 + A_4 + A_1 A_4 h + [A_2 + A_2 A_4 h]L(\varphi). \quad (3.7)$$

Note that $L(\Phi_n) = g_x$. Hence, we can apply the discrete Gronwall inequality to (3.6) and get

$$L(\Phi) \le e^{\tilde{A}^+ T}[g_x + A_5 T] = [g_x + A_5 T][e^{\tilde{A}T} \vee 1],$$

which, combined with (3.7), yields the assertion. ∎

We are now in the position to give the proof of Theorem 3.1.

Proof of Theorem 3.1. First by induction one can easily show that $L_m \triangleq L(u^{n,m}) < \infty$ for each (n, m). Due to Theorem 3.4 we have

$$L_m \le [g_x + A_5 T]\Big[\exp\Big([A_1 + A_4 + A_1 A_4 h]T + [A_2 + A_2 A_4 h]T L_{m-1}\Big) \vee 1\Big];$$

for $\lambda_1 = \lambda_2 = 0$ and any $\lambda_3, \lambda_4 > 0$ satisfying (3.5).

Introducing

$$\tilde{L}_m \triangleq [A_2 + A_2 A_4 h]T L_m,$$

we get

$$\tilde{L}_m \le [A_2 + A_2 A_4 h][g_x + A_5 T]T\Big[e^{[A_1 + A_4 + A_1 A_4 h]T} e^{\tilde{L}_{m-1}} \vee 1\Big] \quad (3.8)$$
$$\le [A_2 + A_2 A_4 h][g_x + A_5 T]T\Big[e^{[A_1 + A_4 + A_1 A_4 h]T} e^{\tilde{L}_{m-1}} + 1\Big].$$

Denote

$$L_0(\lambda, h) \triangleq [A_2 + A_2 A_4 h][g_x + A_5 T]T \times \exp\Big([A_2 + A_2 A_4 h][g_x + A_5 T]T + [A_1 + A_4 + A_1 A_4 h]T\Big). \quad (3.9)$$

Obviously, $\tilde{L}_0 = 0$. If

$$L_0(\lambda, h) \le e^{-1}, \quad (3.10)$$

then, by induction, one can easily show that

$$\tilde{L}_m \le [A_2 + A_2 A_4 h][g_x + A_5 T]T + 1, \quad \forall m.$$

We plug this into the right side of (3.8) to obtain

$$\tilde{L}_m \le [A_2 + A_2 A_4 h][g_x + A_5 T]T\Big[e^{[A_1 + A_4 + A_1 A_4 h]T + [A_2 + A_2 A_4 h][g_x + A_5 T]T + 1} \vee 1\Big].$$

Thus

$$L_m \le [g_x + A_5 T]\Big[e^{[A_1+A_4+A_1A_4h]T+[A_2+A_2A_4h][g_x+A_5T]T+1} \vee 1\Big] \triangleq L_1(\lambda,h). \tag{3.11}$$

So we want to choose λ_3, λ_4 and h which satisfy (3.5) and minimize $L_0(\lambda, h)$. Recall again that $\lambda_1 = \lambda_2 = 0$. In dependence of h we set, for small h,

$$\lambda_3(h) \triangleq \sqrt{h}, \quad \lambda_4(h) \triangleq 1 - \sqrt{h} - Kh - 2K\sqrt{h}. \tag{3.12}$$

Then $A_3 = 1$ and

$$\begin{aligned}
&\lim_{h\downarrow 0} L_0(\lambda(h), h) \\
&= [b_y + \sigma_y][g_x + f_x T]T \exp\Big([b_y + \sigma_y][g_x + f_x T]T + [2k_b + 2k_f + 2 + \sigma_x + f_z]T\Big) \\
&= L_0.
\end{aligned}$$

Note also that

$$\lim_{h\downarrow 0} L_1(\lambda(h), h) = L_1.$$

Suppose now that (3.2) holds true. Then for any $L > L_1$, we obtain $L_0(\lambda(h), h) \le e^{-1}$ and $L_1(\lambda(h), h) \le L$ provided h is small enough. In view of (3.11) the theorem is proved. ∎

4 Linear Growth

This section is devoted to studying the linear growth of the functions $u_i^{n,m}(x)$. Given linear growing functions φ_i, assume

$$|\varphi_i(x)|^2 \le G(\varphi_i)|x|^2 + H(\varphi_i), \quad \forall x;$$

and let

$$G(\varphi) \triangleq \sup_i G(\varphi_i); \quad H(\varphi) \triangleq \sup_i H(\varphi_i).$$

To state the main result of this section we first introduce the functions

$$\Gamma_0(x) \triangleq \frac{e^x - 1}{x}; \quad \Gamma_1(x,y) \triangleq \sup_{0<\theta<1} e^{\theta x}\Gamma_0(\theta y); \quad \forall x, y \in \mathbb{R}; \tag{4.1}$$

and for $G > 0$,

$$
\begin{aligned}
c_0(G) &\triangleq [b_y + \sigma_y]T \times \\
&\Big[g_x\Gamma_1\Big([2k_f + 1 + f_z]T, [(2k_b + 1 + \sigma_x) + (b_y + \sigma_y)G]T\Big) \\
&+ f_x T\Gamma_0([2k_f + 1 + f_z]T)\Gamma_0\Big([2k_b + 1 + \sigma_x]T + [b_y + \sigma_y]GT\Big)\Big]; \\
L_2(G) &\triangleq \Big[e^{[2k_f+1+f_z]T} \vee 1\Big]g_0 + f_0 T\Gamma_0\Big([2k_f + 1 + f_z]T\Big) + [b_0 + \sigma_0]T \times \\
&\Big[g_x\Gamma_1\Big([2k_f + 1 + f_z]T, [(2k_b + 1 + \sigma_x) + (b_y + \sigma_y)G]T\Big) \\
&+ f_x T\Gamma_0([2k_f + 1 + f_z]T)\Gamma_0\Big([2k_b + 1 + \sigma_x]T + [b_y + \sigma_y]GT\Big)\Big].
\end{aligned}
$$

Theorem 4.1 *Assume (3.2) holds true and*

$$c_0(L_1) < 1; \tag{4.2}$$

For any $G > L_1$, $c_0(L_1) < c_0 < 1$, $L_2 > L_2(L_1)$, and for h small enough we have

$$G(u^{n,m}) \leq G; \quad H(u^{n,m}) \leq \frac{L_2}{1 - c_0}; \quad \forall m.$$

Notice that

$$\lim_{x \to -\infty} \Gamma_0(x) = 0; \quad \lim_{x \to -\infty} \Gamma_1(x, y) = 0; \quad \lim_{y \to -\infty} \Gamma_1(x, y) = 0. \tag{4.3}$$

Hence, (4.2) is satisfied in Cases 1–5 of Section 2.

Lemma 4.2 *Assume*

$$X_{i+1} = X_i + b(t_i, X_i, \varphi(X_i))h + \sigma(t_i, X_i, \varphi(X_i))\Delta W_{i+1}.$$

Then

$$E_{t_i}\{|X_{i+1}|^2\} \leq [1 + A_1 h + A_2 h G(\varphi)]|X_i|^2 + [B_1 + A_2 H(\varphi)]h;$$

where $\lambda_1 = \lambda_2 = 0$ and

$$B_1 \triangleq b_0 + \sigma_0 + K b_0 h. \tag{4.4}$$

Proof. Denote

$$b_i \triangleq b(t_i, X_i, \varphi(X_i)); \quad \sigma_i \triangleq \sigma(t_i, X_i, \varphi(X_i)).$$

Then

$$
\begin{aligned}
E_{t_i}\{|X_{i+1}|^2\} &= (X_i + b_i h)^2 + \sigma_i^2 h = X_i^2 + 2X_i b_i h + b_i^2 h^2 + \sigma_i^2 h \\
&\leq X_i^2 + 2k_b X_i^2 h + X_i^2 h + [b_0 + b_y|\varphi(X_i)|^2]h \\
&\quad + K[b_0 + X_i^2 + |\varphi(X_i)|^2]h^2 + [\sigma_0 + \sigma_x X_i^2 + \sigma_y|\varphi(X_i)|^2]h \\
&= [1 + A_1 h]|X_i|^2 + A_2 h|\varphi(X_i)|^2 + B_1 h \\
&\leq [1 + A_1 h]|X_i|^2 + A_2 h[G(\varphi)|X_i|^2 + H(\varphi)] + B_1 h;
\end{aligned}
$$

which proves the lemma. ■

Following the arguments in Lemma 3.3, one can easily prove

Lemma 4.3 *Assume*

$$Y_i = Y_{i+1} + f(t_i, X_i, Y_{i+1}, \hat{Z}_i)h - \int_{t_i}^{t_{i+1}} Z_t dW_t;$$

where

$$\hat{Z}_i = \frac{1}{h} E_{t_i}\{Y_{i+1}\Delta W_{i+1}\}.$$

Then

$$|Y_i|^2 + (1 - A_3)h|\hat{Z}_i|^2 \le [1 + A_4 h]E_{t_i}\{|Y_{i+1}|^2\} + A_5 h|X_i|^2 + B_2 h;$$

where $\lambda_1 = \lambda_2 = 0$ *and*

$$B_2 \triangleq f_0 + K f_0 h. \tag{4.5}$$

To derive bounds for the linear growth of $F_n(\varphi)$, we define discrete time versions of Γ_0 and Γ_1 by

$$\Gamma_0^i(x) \triangleq \frac{(1+xh)^i - 1}{x}; \quad \Gamma_1^n(x,y) \triangleq \sup_{0\le i\le n}(1+xh)^i \Gamma_0^i(y); \tag{4.6}$$

and discrete time versions of $c_0(G), L_2(G)$ by

$$\begin{aligned} c_0(\lambda,h,G) &\triangleq A_2\Big[g_x\Gamma_1^n(A_4, A_1+A_2G) + A_5\Gamma_0^n(A_4)\Gamma_0^n(A_1+A_2G)\Big]; \\ L_2(\lambda,h,G) &\triangleq B_1\Big[g_x\Gamma_1^n(A_4, A_1+A_2G) + A_5\Gamma_0^n(A_4)\Gamma_0^n(A_1+A_2G)\Big] \\ &\quad +[e^{A_4T}\vee 1]g_0 + B_2\Gamma_0^n(A_4). \end{aligned} \tag{4.7}$$

Theorem 4.4 *For any linear growing* φ*,*

$$G(F_n(\varphi)) \le [g_x + A_5 T]\Big[e^{[A_1+A_4+A_1A_4h]T + [A_2+A_2A_4h]TG(\varphi)} \vee 1\Big]; \tag{4.8}$$

$$H(F_n(\varphi)) \le c_0(\lambda,h,G(\varphi))H(\varphi) + L_2(\lambda,h,G(\varphi)). \tag{4.9}$$

where $\lambda_1 = \lambda_2 = 0$ *and* $\lambda_3, \lambda_4 > 0$ *are supposed to fulfill (3.5).*

Proof. Denote $\Phi \triangleq F_n(\varphi)$. Fix (i_0, x) and define, for $i = i_0, \cdots, n-1$,

$$\begin{cases} X_{i_0} \triangleq x; \\ X_{i+1} \triangleq X_i + b(t_i, X_i, \varphi_i(X_i))h + \sigma(t_i, X_i, \varphi_i(X_i))\Delta W_{i+1}; \\ Y_n \triangleq g(X_n); \\ \hat{Z}_i \triangleq \frac{1}{h}E_{t_i}\{Y_{i+1}\Delta W_{i+1}\}; \\ Y_i \triangleq Y_{i+1} + f(t_i, X_i, Y_{i+1}, \hat{Z}_i)h - \int_{t_i}^{t_{i+1}} Z_t dW_t. \end{cases}$$

Then obviously $Y_{i_0} = \Phi_{i_0}(x)$. Since $\lambda_1 = \lambda_2 = 0$ we obtain from Lemma 4.2 that

$$E\{|X_{i+1}|^2\} \leq [1 + A_1 h + A_2 h G(\varphi)] E\{|X_i|^2\} + [B_1 + A_2 H(\varphi)]h.$$

Then

$$\begin{aligned} E\{|X_i|^2\} &\leq [1 + A_1 h + A_2 h G(\varphi)]^{i-i_0} E\{|X_{i_0}|^2\} \\ &\quad + [B_1 + A_2 H(\varphi)] h \sum_{j=i_0}^{i-1} [1 + A_1 h + A_2 h G(\varphi)]^{j-i_0} \\ &= [1 + A_1 h + A_2 h G(\varphi)]^{i-i_0} |x|^2 + [B_1 + A_2 H(\varphi)] \Gamma_0^{i-i_0}(A_1 + A_2 G(\varphi)). \end{aligned}$$

Next, applying Lemma 4.3 and by (3.5) we have

$$E\{|Y_i|^2\} \leq [1 + A_4 h] E\{|Y_{i+1}|^2\} + A_5 h E\{|X_i|^2\} + B_2 h.$$

Then

$$\begin{aligned} & |\Phi_{i_0}(x)|^2 = |Y_{i_0}|^2 \\ & \leq (1 + A_4 h)^{n-i_0} E\{|Y_n|^2\} + A_5 h \sum_{i=i_0}^{n-1} (1 + A_4 h)^{i-i_0} E\{|X_i|^2\} \\ & \quad + B_2 h \sum_{i=i_0}^{n-1} (1 + A_4 h)^{i-i_0} \\ & \leq (1 + A_4 h)^{n-i_0} [g_0 + g_x E\{|X_n|^2\}] + A_5 h \sum_{i=i_0}^{n-1} (1 + A_4 h)^{i-i_0} E\{|X_i|^2\} \\ & \quad + B_2 \Gamma_0^{n-i_0}(A_4) \\ & \leq (1 + A_4 h)^{n-i_0} g_0 + B_2 \Gamma_0^{n-i_0}(A_4) + (1 + A_4 h)^{n-i_0} g_x \times \\ & \quad \Big[[1 + A_1 h + A_2 h G(\varphi)]^{n-i_0} |x|^2 + [B_1 + A_2 H(\varphi)] \Gamma_0^{n-i_0}(A_1 + A_2 G(\varphi)) \Big] \\ & \quad + A_5 h \sum_{i=i_0}^{n-1} (1 + A_4 h)^{i-i_0} \Big[[1 + A_1 h + A_2 h G(\varphi)]^{i-i_0} |x|^2 \\ & \quad + [B_1 + A_2 H(\varphi)] \Gamma_0^{i-i_0}(A_1 + A_2 G(\varphi)) \Big]. \end{aligned}$$

This implies

$$\begin{aligned} G(\Phi_{i_0}) &\leq (1 + A_4 h)^{n-i_0} g_x [1 + A_1 h + A_2 h G(\varphi)]^{n-i_0} \\ &\quad + A_5 h \sum_{i=i_0}^{n-1} (1 + A_4 h)^{i-i_0} [1 + A_1 h + A_2 h G(\varphi)]^{i-i_0}; \\ H(\Phi_{i_0}) &\leq (1 + A_4 h)^{n-i_0} g_0 + B_2 \Gamma_0^{n-i_0}(A_4) \\ &\quad + \Big[g_x (1 + A_4 h)^{n-i_0} \Gamma^{n-i_0}(A_1 + A_2 G(\varphi)) \\ &\quad + A_5 h \sum_{i=i_0}^{n-1} (1 + A_4 h)^{i-i_0} \Gamma^{i-i_0}(A_1 + A_2 G(\varphi)) \Big] [B_1 + A_2 H(\varphi)]. \end{aligned}$$

Note that, for $0 \le i \le n$,

$$(1+xh)^i \le e^{xT} \vee 1; \quad \Gamma_0^i(x) \le \Gamma_0^n(x); \quad (1+xh)^i \Gamma_0^i(y) \le \Gamma_1^n(x,y). \tag{4.10}$$

Then

$$\begin{aligned} G(\Phi_{i_0}) &\le [g_x + A_5 T]\Big[e^{[A_1+A_4+A_1A_4h]T+[A_2+A_2A_4h]TG(\varphi)} \vee 1\Big]; \\ H(\Phi_{i_0}) &\le [e^{A_4T} \vee 1]g_0 + B_2\Gamma_0^n(A_4) \\ &\quad + \Big[g_x\Gamma_1^n(A_4, A_1 + A_2G(\varphi)) + A_5\Gamma_0^n(A_4)\Gamma_0^n(A_1 + A_2G(\varphi))\Big] \\ &\quad \times [B_1 + A_2H(\varphi)]. \end{aligned}$$

Since the right hand side does not depend on i_0, the assertion is proved. ■

After these preparations we give the proof of Theorem 4.1:

Proof of Theorem 4.1. Denote $G_m \triangleq G(u^{n,m}), H_m \triangleq H(u^{n,m})$. Obviously, $G_0 = H_0 = 0$. We may now conclude from Theorem 4.4 that under (3.5)

$$G_m \le [g_x + A_5T]\Big[e^{[A_1+A_4+A_1A_4h]T+[A_2+A_2A_4h]TG_{m-1}} \vee 1\Big]; \tag{4.11}$$

$$H_m \le c_0(\lambda, h, G_m)H_{m-1} + L_2(\lambda, h, G_m). \tag{4.12}$$

We now choose $\lambda_3(h)$ and $\lambda_4(h)$ as in (3.12) for small h. Since (3.2) holds true, for any $G > L_1$, we can follow the same arguments as in Theorem 3.1 and get $G(u^{n,m}) \le G$ from (4.11). Note that

$$\begin{aligned} &\lim_{n\to\infty} \Gamma_0^n(x) = T\Gamma_0(xT); \\ &\lim_{n\to\infty} \Gamma_1^n(x,y) = T\Gamma_1(xT, yT); \\ &\lim_{h\downarrow 0} c_0(\lambda(h), h, G) = c_0(G); \\ &\lim_{h\downarrow 0} L_2(\lambda(h), h, G) = L_2(G). \end{aligned}$$

For any c_0, $c_0(L_1) < c_0 < 1$, and L_2, $L_2(L_1) < L_2$, we can choose $G > L_1$ such that $c_0(G) < c_0$ and $L_2(G) < L_2$. Then, for sufficiently small h, it holds that $c_0(\lambda, h, G) \le c_0$ and $L_2(\lambda, h, G) \le L_2$. Now by (4.12) we get

$$H_m \le c_0 H_{m-1} + L_2,$$

which implies the result. ■

5 Convergence of the Markovian Iteration

We now make the assumptions of Theorem 2.1 precise and prove convergence of the Markovian iteration as the number of iteration steps tends to infinity.

To this end we first introduce

$$
\begin{aligned}
&c_1(\lambda_2, L) \triangleq (1+\lambda_2^{-1})[b_y+\sigma_y]T \times \\
&\quad \Big[g_x\Gamma_1\Big([2k_f+1+f_z]T, [2k_b+1+\sigma_x+(1+\lambda_2)[b_y+\sigma_y]L]T\Big) \\
&\quad +f_xT\Gamma_0\Big([2k_f+1+f_z]T\Big)\Gamma_0\Big([2k_b+1+\sigma_x+(1+\lambda_2)[b_y+\sigma_y]L]T\Big)\Big]; \\
&c_2(\lambda_2, L, G) \triangleq \Big[e^{[2k_b+1+\sigma_x+[b_y+\sigma_y]G]T} \vee 1\Big] c_1(\lambda_2, L); \\
&c_2(L, G) \triangleq \inf_{\lambda_2>0} c_2(\lambda_2, L, G).
\end{aligned}
$$

We are going to prove the following theorem:

Theorem 5.1 *Assume (3.2) and*

$$
c_2(L_1, L_1) < 1. \tag{5.1}
$$

(i) For any $\bar{L} > L_1, \bar{G} > L_1, L_2 > L_2(L_1), c_0(L_1) < c_0 < 1$, there exists a solution u^n to (1.6) such that

$$
L(u^n) \le \bar{L}; \quad G(u^n) \le \bar{G}; \quad H(u^n) \le \bar{H} \triangleq \frac{L_2}{1-c_0}, \tag{5.2}
$$

if h is small enough.

(ii) For any $c_2(L_1, L_1) < c_2 < 1$, we may find $c_1 \le c_2$ such that for h small enough,

$$
G(u^{n,m}-u^n) \le \frac{3\bar{G}}{(1-\sqrt{c_2})^2} c_2^m; \tag{5.3}
$$

$$
\begin{aligned}
H(u^{n,m}-u^n) \le \; & \frac{3\bar{H}}{(1-\sqrt{c_1})^2} c_1^m \\
& + \frac{3}{(1-\sqrt{c_2})^4}\Big[[b_0+\sigma_0]+[b_y+\sigma_y]\bar{H}\Big] T\bar{G} m c_2^m.
\end{aligned} \tag{5.4}
$$

(iii) Fix $G > 0$ and suppose $\tilde{u}^n$ is another solution to (1.6) with linear growth such that $G(\tilde{u}^n) \le G$. Then $\tilde{u}^n = u^n$, if h (depending on G) is small enough.

Remark 5.2 *(i) In view of (4.3), it is straightforward to see that (5.1) is also satisfied in Cases 1–5 of Section 2.*

(ii) One can recover $c_0(L)$ from $c_1(\lambda_2, L)$ by formally replacing λ_2 and λ_2^{-1} by zero. Consequently, we have for all $L > 0$,

$$
c_0(L) \le \inf_{\lambda_2>0} c_1(\lambda_2, L) \le c_2(L, L).
$$

In particular, condition (5.1) implies (4.2).

Again we first study the operator F_n to prepare the proof of Theorem 5.1.

Theorem 5.3 *Assume φ^1, φ^2 have linear growth and φ^1 is Lipschitz continuous. Then*

$$\begin{aligned}
G(F_n(\varphi^1) - F_n(\varphi^2)) &\le c_2(\lambda_2, h, L(\varphi^1), G(\varphi^2))G(\Delta\varphi)\\
H(F_n(\varphi^1) - F_n(\varphi^2)) &\le c_1(\lambda_2, h, L(\varphi^1))H(\Delta\varphi)\\
&\quad + c_2(\lambda_2, h, L(\varphi^1), G(\varphi^2))[B_1 + A_2 H(\varphi^2)]TG(\Delta\varphi).
\end{aligned}$$

where $\lambda_1 = 0$, λ_3, λ_4 are chosen such that (3.5) holds, and

$$\begin{aligned}
c_1(\lambda_2, h, L) &\triangleq (1 + \lambda_2^{-1})A_2\Big[g_x\Gamma_1^n(A_4, A_1 + (1+\lambda_2)A_2 L)\\
&\qquad + A_5\Gamma_0^n(A_4)\Gamma_0^n(A_1 + (1+\lambda_2)A_2 L)\Big];\\
c_2(\lambda_2, h, L, G) &\triangleq \Big[e^{[A_1 + A_2 G]T} \vee 1\Big]c_1(\lambda_2, h, L).
\end{aligned}$$

Proof. For $l = 1, 2$, denote $\Phi^l \triangleq F_n(\varphi^l)$. Fix (i_0, x) and define, for $i = i_0, \cdots, n-1$,

$$\begin{cases}
X_{i_0}^l \triangleq x;\\
X_{i+1}^l \triangleq X_i^l + b(t_i, X_i^l, \varphi_i^l(X_i^l))h + \sigma(t_i, X_i^l, \varphi_i^l(X_i^l))\Delta W_{i+1};\\
Y_n^l \triangleq g(X_n^l);\\
\hat{Z}_i^l \triangleq \dfrac{1}{h}E_{t_i}\{Y_{i+1}^l \Delta W_{i+1}\};\\
Y_i^l \triangleq Y_{i+1}^l + f(t_i, X_i^l, Y_{i+1}^l, \hat{Z}_i^l)h - \displaystyle\int_{t_i}^{t_{i+1}} Z_t^l dW_t.
\end{cases}$$

Then obviously $Y_{i_0}^l = \Phi_{i_0}^l(x)$.

Denote $L \triangleq L(\varphi_1)$, and

$$\Delta X \triangleq X^1 - X^2; \quad \Delta Y \triangleq Y^1 - Y^2; \quad \Delta\hat{Z} \triangleq \hat{Z}^1 - \hat{Z}^1; \quad \Delta\varphi \triangleq \varphi^1 - \varphi^2; \quad \Delta\Phi \triangleq \Phi^1 - \Phi^2.$$

Application of Lemma 3.2, with $\lambda_1 = 0$, yields

$$\begin{aligned}
&E\{|\Delta X_{i+1}|^2\} \le E\Big\{[1 + A_1 h + (1+\lambda_2)A_2 hL]|\Delta X_i|^2 + (1+\lambda_2^{-1})A_2 h|\Delta\varphi(X_i^2)|^2\Big\}\\
&\le E\Big\{[1 + A_1 h + (1+\lambda_2)A_2 hL]|\Delta X_i|^2 + (1+\lambda_2^{-1})A_2 h\Big[G(\Delta\varphi)|X_i^2|^2 + H(\Delta\varphi)\Big]\Big\}\\
&\le [1 + A_1 h + (1+\lambda_2)A_2 hL]E\{|\Delta X_i|^2\}\\
&\quad + (1+\lambda_2^{-1})A_2 h\Big[G(\Delta\varphi)\sup_{i_0 \le j \le n} E\{|X_j^2|^2\} + H(\Delta\varphi)\Big].
\end{aligned}$$

Note that $\Delta X_{i_0} = 0$. Therefore,

$$\begin{aligned}
&\sup_{i_0 \le i \le n} E\{|\Delta X_i|^2\} \le (1+\lambda_2^{-1})A_2 h \times \\
&\qquad \Big[G(\Delta\varphi) \sup_{i_0 \le j \le n} E\{|X_j^2|^2\} + H(\Delta\varphi)\Big] \sum_{i=i_0}^{n-1} [1 + A_1 h + (1+\lambda_2)A_2 h L]^{i-i_0} \\
&= (1+\lambda_2^{-1})A_2 \Big[G(\Delta\varphi) \sup_{i_0 \le j \le n} E\{|X_j^2|^2\} + H(\Delta\varphi)\Big] \Gamma_0^{n-i_0}(A_1 + (1+\lambda_2)A_2 L).
\end{aligned}$$

Applying Lemma 4.2, we have

$$\sup_{i_0 \le i \le n} E\{|X_i^2|^2\} \le \Big[|x|^2 + [B_1 + A_2 H(\varphi^2)]T\Big]\Big[e^{[A_1 + A_2 G(\varphi^2)]T} \vee 1\Big] \triangleq \tilde{A}. \qquad (5.5)$$

Thus

$$\sup_{i_0 \le i \le n} E\{|\Delta X_i|^2\} \le (1+\lambda_2^{-1})A_2 \Gamma_0^{n-i_0}(A_1 + (1+\lambda_2)A_2 L)\Big[G(\Delta\varphi)\tilde{A} + H(\Delta\varphi)\Big]. \quad (5.6)$$

Furthermore, we obtain from Lemma 3.3 and (3.5),

$$E\{|\Delta Y_i|^2\} \le [1 + A_4 h] E\{|\Delta Y_{i+1}|^2\} + A_5 h E\{|\Delta X_i|^2\}.$$

Hence

$$\begin{aligned}
&|\Delta \Phi_{i_0}(x)|^2 = |\Delta Y_{i_0}|^2 \\
&\le (1+A_4 h)^{n-i_0} E\{|\Delta Y_n|^2\} + A_5 \Gamma_0^{n-i_0}(A_4) \sup_{i_0 \le i \le n} E\{|\Delta X_i|^2\} \\
&\le \Big[(1+A_4 h)^{n-i_0} g_x + A_5 \Gamma_0^{n-i_0}(A_4)\Big] \sup_{i_0 \le i \le n} E\{|\Delta X_i|^2\} \\
&\le (1+\lambda_2^{-1})A_2 \Big[G(\Delta\varphi)\tilde{A} + H(\Delta\varphi)\Big]\Big[g_x (1+A_4 h)^{n-i_0} \Gamma_0^{n-i_0}(A_1 + (1+\lambda_2)A_2 L) \\
&\quad + A_5 \Gamma_0^{n-i_0}(A_4) \Gamma_0^{n-i_0}(A_1 + (1+\lambda_2)A_2 L)\Big].
\end{aligned}$$

In view of (4.10), we get

$$\begin{aligned}
&\sup_i |\Delta \Phi_i(x)|^2 \\
&\le \Big[G(\Delta\varphi)\tilde{A} + H(\Delta\varphi)\Big] c_1(\lambda_2, h, L) \\
&= c_1(\lambda_2, h, L)\Big\{G(\Delta\varphi)\Big[e^{[A_1 + A_2 G(\varphi^2)]T} \vee 1\Big]\Big[|x|^2 + [B_1 + A_2 H(\varphi^2)]T\Big] + H(\Delta\varphi)\Big\};
\end{aligned}$$

which implies the theorem. ■

We can apply this theorem to estimate the distance between $u^{n,m}$ and $u^{n,m-1}$.

Theorem 5.4 *Assume that $L(u^{n,m}) \le \bar{L}$, $G(u^{n,m}) \le \bar{G}$ and $H(u^{n,m}) \le \bar{H}$ for all $m \in \mathbb{N}$ and sufficiently small h. Moreover let*

$$c_2(\bar{L}, \bar{G}) < 1.$$

Then for any $c_2(\bar{L}, \bar{G}) < c_2 < 1$, we may find $c_1 \le c_2$ such that for h small enough,

$$\begin{aligned} &G(u^{n,m} - u^{n,m-1}) \le \bar{G} c_2^{m-1}; \\ &H(u^{n,m} - u^{n,m-1}) \le \bar{H} c_1^{m-1} + \Big[[b_0 + \sigma_0] + [b_y + \sigma_y]\bar{H} \Big] T\bar{G}(m-1)c_2^{m-1}. \end{aligned}$$

Proof. Let $\lambda_1 = 0$ and choose $\lambda_3,\ \lambda_4$ depending on h as in (3.12). Note that with this choice

$$\begin{aligned} &\lim_{h \to 0} \Gamma_1^n(x, y) = T\Gamma_1(xT, yT); \\ &\lim_{h \to 0} c_1(\lambda_2, h, L) = c_1(\lambda_2, L); \\ &\lim_{h \to 0} c_2(\lambda_2, h, L, G) = c_2(\lambda_2, L, G). \end{aligned}$$

Hence we may find an appropriate λ_2 such that for h small enough,

$$c_2(\lambda_2, h, \bar{L}, \bar{G}) < c_2.$$

Since

$$c_1(\lambda_2, h, \bar{L}) \le c_2(\lambda_2, h, \bar{L}, \bar{G}),$$

we may find an $c_1 \le c_2$ such that for small h

$$c_1(\lambda_2, h, \bar{L}) \le c_1.$$

Applying Theorem 5.3, we get, for small h,

$$G(u^{n,m} - u^{n,m-1}) \le c_2 G(u^{n,m-1} - u^{n,m-2}); \tag{5.7}$$

$$\begin{aligned} H(u^{n,m} - u^{n,m-1}) \le\ & c_1 H(u^{n,m-1} - u^{n,m-2}) \\ &+ c_2(\lambda_2, h, \bar{L}, \bar{G})[B_1 + A_2\bar{H}]TG(u^{n,m-1} - u^{n,m-2}). \end{aligned} \tag{5.8}$$

Note that

$$G(u^{n,1} - u^{n,0}) = G(u^{n,1}) \le \bar{G}.$$

By (5.7) we therefore get

$$G(u^{n,m} - u^{n,m-1}) \le \bar{G} c_2^{m-1}.$$

Moreover, for h small enough we may also assume that

$$c_2(\lambda_2, h, L, G)[B_1 + A_2\bar{H}] \le c_2 \Big[[b_0 + \sigma_0] + [b_y + \sigma_y]\bar{H} \Big].$$

Then by (5.8) we get

$$
\begin{aligned}
&H(u^{n,m}-u^{n,m-1}) \le c_1^{m-1}H(u^{n,1}-u^{n,0})\\
&\quad + c_2\Big[[b_0+\sigma_0]+[b_y+\sigma_y]\bar H\Big]Tc_1^{m-1}\sum_{i=1}^{m-1}\frac{G(u^{n,i}-u^{n,i-1})}{c_1^i}\\
&\le c_1^{m-1}H(u^{n,1})+c_2\Big[[b_0+\sigma_0]+[b_y+\sigma_y]\bar H\Big]Tc_1^{m-1}\bar G\sum_{i=1}^{m-1}\frac{c_2^{i-1}}{c_1^i}\\
&\le \bar H c_1^{m-1}+\Big[[b_0+\sigma_0]+[b_y+\sigma_y]\bar H\Big]T\bar G[m-1]c_2^{m-1}.
\end{aligned}
$$

The proof is complete now. ■

Theorem 5.1 can now be proved by iterating the above theorem.

Proof of Theorem 5.1. Assume $\bar G, \bar L, L_2, c_0, c_2$ satisfy the conditions specified in the theorem. Without loss of generality we assume $c_2(\bar L, \bar G) < c_2$. Recall that (5.1) implies (4.2). Hence by Theorems 3.1 and 4.1, we get for h small enough

$$L(u^{n,m}) \le \bar L; \quad G(u^{n,m}) \le \bar G; \quad H(u^{n,m}) \le \bar H.$$

Hence, (i) will follow directly from (ii).

To prove (ii), we denote

$$\tilde L \triangleq \Big[[b_0+\sigma_0]+[b_y+\sigma_y]\bar H\Big]T\bar G.$$

Applying Theorem 5.4, we get

$$|u_i^{n,m}(x)-u_i^{n,m-1}(x)|^2 \le \bar G|x|^2 c_2^{m-1}+\bar H c_1^{m-1}+\tilde L[m-1]c_2^{m-1}.$$

Then

$$|u_i^{n,m}(x)-u_i^{n,m-1}(x)| \le \sqrt{\bar G}|x|c_2^{\frac{m-1}{2}}+\sqrt{\bar H}c_1^{\frac{m-1}{2}}+\sqrt{\tilde L[m-1]}c_2^{\frac{m-1}{2}}.$$

Thus for any $m_1 > m$,

$$
\begin{aligned}
&|u_i^{n,m}(x)-u_i^{n,m_1}(x)| \le \sum_{j=m}^{\infty}\Big[\sqrt{\bar G}|x|c_2^{\frac{j}{2}}+\sqrt{\bar H}c_1^{\frac{j}{2}}+\sqrt{\frac{\tilde L}{m}}jc_2^{\frac{j}{2}}\Big]\\
&\le \sqrt{\bar G}|x|\frac{c_2^{\frac{m}{2}}}{1-\sqrt{c_2}}+\sqrt{\bar H}\frac{c_1^{\frac{m}{2}}}{1-\sqrt{c_1}}+\sqrt{\frac{\tilde L}{m}}\frac{m(1-\sqrt{c_2})+\sqrt{c_2}}{(1-\sqrt{c_2})^2}c_2^{\frac{m}{2}}.
\end{aligned}
$$

Note that the right side above converges to 0 as $m\to\infty$. Then $u_i^{n,m}(x)$ is a Cauchy sequence and hence converges to some $u_i^n(x)$. Moreover,

$$
\begin{aligned}
&|u_i^{n,m}(x)-u_i^n(x)|^2\\
&\le 3\Big[\bar G|x|^2\frac{c_2^m}{(1-\sqrt{c_2})^2}+\bar H\frac{c_1^m}{(1-\sqrt{c_1})^2}+\frac{\tilde L m}{(1-\sqrt{c_2})^4}c_2^m\Big];
\end{aligned}
$$

which leads to (5.3) and (5.4) and thus proves (ii).

It remains to prove (iii). For any $G > 0$, assume $\tilde{u}^n$ is another solution to (1.6) with linear growth such that $G(\tilde{u}^n) \le G$. Then $F_n(\tilde{u}^n) = \tilde{u}^n$. Note that $\tilde{u}^n_n = g = u^n_n$. Assume $\tilde{u}^n_{i+1} = u^n_{i+1}$. We now apply a local version of Theorem 5.3. That is, we consider (2.1) only on the interval $[t_i, t_{i+1}]$ with terminal condition $\Phi_{i+1}(x) \stackrel{\triangle}{=} u^n_{i+1}(x)$ (instead of on $[0, T]$ with terminal condition $g(x)$). We note that in this case there is only one time subinterval. One can check directly that

$$\Gamma^1_0(x) = h; \quad \Gamma^1_1(x, y) = (1 + xh)h.$$

Setting $\varphi^1 \stackrel{\triangle}{=} u^n, \varphi^2 \stackrel{\triangle}{=} \tilde{u}^n$ we get

$$\begin{aligned} G(u^n_i - \tilde{u}^n_i) &\le \tilde{c}_2(h) G(u^n_i - \tilde{u}^n_i); \\ H(u^n_i - \tilde{u}^n_i) &\le \tilde{c}_1(h) H(u^n_i - \tilde{u}^n_i) + \tilde{c}_2(h)[B_1 + A_2 H(\tilde{u}^n)] T G(u^n_i - \tilde{u}^n_i); \end{aligned}$$

where

$$\begin{aligned} \tilde{c}_1(h) &\stackrel{\triangle}{=} (1 + \lambda_2^{-1}) A_2 [\bar{L}(1 + A_4 h)h + A_5 h^2]; \\ \tilde{c}_2(h) &\stackrel{\triangle}{=} \Big[e^{[A_1 + A_2 G]h} \vee 1\Big] \tilde{c}_1(h). \end{aligned}$$

For any G, we have $\tilde{c}_1(h) \le \tilde{c}_2(h) < 1$, provided h is small enough. Then $G(u^n_i - \tilde{u}^n_i) = 0$ and thus $H(u^n_i - \tilde{u}^n_i) = 0$. Consequently $\tilde{u}^n_i = u^n_i$. Repeating the arguments backwardly we get $\tilde{u}^n = u^n$. ■

6 Convergence of the Time Discretization

We now study the error due to the time discretization. We first introduce a continuous time version of the operator F_n. Suppose φ is a function on $[0, T] \times \mathbb{R}$ which is Lipschitz in the space variable and let $(X^{\varphi,r,x}, Y^{\varphi,r,x}, Z^{\varphi,r,x})$ be the unique solution to the decoupled FBSDE $(0 \le r \le t \le T)$

$$\begin{cases} X^{\varphi,r,x}_t = x + \displaystyle\int_r^t b(s, X^{\varphi,r,x}_s, \varphi(s, X^{\varphi,r,x}_s))ds + \int_r^t \sigma(s, X^{\varphi,r,x}_s, \varphi(s, X^{\varphi,r,x}_s))dW_s; \\ Y^{\varphi,r,x}_t = g(X^{\varphi,r,x}_T) + \displaystyle\int_t^T f(s, X^{\varphi,r,x}_s, Y^{\varphi,r,x}_s, Z^{\varphi,r,x}_s)ds - \int_t^T Z^{\varphi,r,x}_s dW_s; \end{cases} \tag{6.1}$$

We then define $\Phi(t, x) \stackrel{\triangle}{=} Y^{\varphi,t,x}_t$ and $F(\varphi) \stackrel{\triangle}{=} \Phi$. It is known from Pardoux and Peng (1992) that, under Assumption 2 and if φ is additionally continuous as a function in time and space, Φ is a viscosity solution to the following semilinear PDE:

$$\begin{cases} \Phi_t + \frac{1}{2}\sigma^2(t, x, \varphi)\Phi_{xx} + b(t, x, \varphi)\Phi_x + f(t, x, \Phi, \Phi_x \sigma(t, x, \varphi)) = 0; \\ \Phi(T, x) = g(x); \end{cases}$$

We now define recursively $\tilde{u}^0 \stackrel{\triangle}{=} 0$ and $\tilde{u}^m \stackrel{\triangle}{=} F(\tilde{u}^{m-1})$. Then the following theorem can be proved similarly, actually more easily than, Theorem 5.1. We hence postpone the proof to the appendix.

Theorem 6.1 *Assume (3.2) and (5.1) hold true.*

(i) $\tilde{u}^m$ *converges to some function* u *uniformly on compacts.*

(ii) $|u(t,x_1) - u(t,x_2)|^2 \le L_1|x_1 - x_2|^2; \quad |u(t,x)|^2 \le L_1|x|^2 + \frac{L_2(L_1)}{1-c_0(L_1)}$.

(iii) $F(u) = u$. *Moreover, if* $F(\tilde{u}) = \tilde{u}$ *and* $\tilde{u}$ *has linear growth, then* $\tilde{u} = u$.

(iv) Under Assumption 2, u *is a viscosity solution to (1.3).*

From the previous theorem and some arguments similar to those in Delarue (2002) or Zhang (2006) we can derive the following corollary. A detailed proof is again given in the appendix.

Corollary 6.2 *Assume all the conditions in Theorem 6.1 hold true. Then FBSDE (1.1) has a unique solution* (X, Y, Z). *Moreover, it holds that* $Y_t = u(t, X_t)$, *and thus*

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, u(s, X_s))ds + \int_0^t \sigma(s, X_s, u(s, X_s))dW_s \\ Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s; \end{cases} \tag{6.2}$$

From now on we fix some $T_0 > 0$ and always assume $T \le T_0$. Moreover, we denote by C a generic constant which may depend on the coefficients b, σ, f, g, and T_0, but is independent of n, h, T and x. The value of C may vary from line to line.

The decoupling relation $Y_t = u(t, X_t)$ together with the Lipschitz continuity of u enables us to apply some estimates for decoupled FBSDEs directly in the present situation.

Corollary 6.3 *Under the assumptions of Theorem 6.1 the following estimates hold true:*

$$|u(s,x) - u(t,x)|^2 \le C[1 + |x|^2]|s - t|; \tag{6.3}$$

$$E\Big\{\sup_{0\le t\le T}[|X_t|^2 + |Y_t|^2] + \int_0^T |Z_t|^2 dt\Big\} \le C[1 + |x|^2]. \tag{6.4}$$

Moreover, if additionally Assumption 2 is in force, then

$$\sup_i E\Big\{\sup_{t\in[t_i,t_{i+1}]}[|X_t - X_{t_i}|^2 + |Y_t - Y_{t_i}|^2]\Big\} + \sum_i E\Big\{\int_{t_i}^{t_{i+1}} |Z_t - \tilde{Z}_{t_i}|^2 dt\Big\} \le C[1 + |x|^2]h, \tag{6.5}$$

where

$$\tilde{Z}_{t_i} \triangleq \frac{1}{h} E_{t_i}\Big\{ \int_{t_i}^{t_{i+1}} Z_t dt \Big\} \tag{6.6}$$

Proof. We exploit that, by Corollary 6.2, (X, Y, Z) solves (1.1) and (6.2). Since (6.2) is a decoupled FBSDE with Lipschitz coefficients, (6.4) is standard, see, e.g. Lemma 2.4 in Zhang (2004). Moreover, (6.5) will follow from Theorem 3.4.3 in Zhang (2001), once (6.3) is proved.

It thus remains to prove (6.3). Let $(X^{t,x}, Y^{t,x}, Z^{t,x})$ denote the solution to (6.2) with initial time t and initial value x. Then, for $0 \le s < t \le T$,

$$\begin{aligned}
&|u(s,x) - u(t,x)|^2 = |E\Big\{ Y^{s,x}_s - Y^{s,x}_t + u(t, X^{s,x}_t) - u(t,x) \Big\}|^2 \\
&\le CE\Big\{ |\int_s^t f(r, X^{s,x}_r, Y^{s,x}_r, Z^{s,x}_r) dr|^2 + |X^{s,x}_t - X^{s,x}_s|^2 \Big\} \\
&\le CE\Big\{ (t-s) \int_s^t |f(r, X^{s,x}_r, Y^{s,x}_r, Z^{s,x}_r)|^2 dr \\
&\quad + (t-s) \int_s^t |b(r, X^{s,x}_r, Y^{s,x}_r)|^2 dr + \int_s^t |\sigma(r, X^{s,x}_r, Y^{s,x}_r)|^2 dr \Big\} \\
&\le CE\Big\{ 1 + \sup_{s \le r \le t} [|X^{s,x}_r|^2 + |Y^{s,x}_r|^2] + \int_s^t |Z^{s,x}_r|^2 dr \Big\} [t-s] \\
&\le C[1 + |x|^2][t-s];
\end{aligned}$$

thanks to (6.4). ■

With this regularity results for (X, Y, Z) and u at hand, we restate and prove Theorem 2.2.

Theorem 6.4 *Suppose Assumption 2 is in force. Moreover, let (3.2) and (5.1) hold true. Then*

$$|u^n_i(x) - u(t_i, x)|^2 \le C[1 + |x|^2]h.$$

Proof. We treat the case $b_y + \sigma_y \neq 0$ only. Otherwise the FBSDE is decoupled and the Theorem is proved in Zhang (2004). We follow similar arguments as in Theorem 5.3. Fix n and (i_0, x). Consider (1.1) and (1.4) with initial time i_0 and initial value x. For notational simplicity we still denote their solutions as (X, Y, Z) and $(X^n, Y^n, Z^n, \hat{Z}^n)$, respectively. Denote

$$\Delta X^n_i \triangleq X_{t_i} - X^n_i; \quad \Delta Y^n_i \triangleq Y_{t_i} - Y^n_i; \quad \Delta \hat{Z}^n_i \triangleq Z_{t_i} - \hat{Z}^n_i; \quad \Delta u^n_i(x) \triangleq u(t_i, x) - u^n_i(x).$$

For $t \in [t_i, t_{i+1})$, denote

$$\alpha^1_t \triangleq b(t, X_t, Y_t) - b(t_i, X_{t_i}, Y_{t_i}); \quad \beta^1_t \triangleq \sigma(t, X_t, Y_t) - \sigma(t_i, X_{t_i}, Y_{t_i}).$$

Then by Assumption 2 and (6.5) we have

$$E\{|\alpha_t^1|^2+|\beta_t^1|^2\}\le CE\Big\{h+|X_t-X_{t_i}|^2+|Y_t-Y_{t_i}|^2\Big\}\le C[1+|x|^2]h.$$

We thus obtain from Lemma 3.2 (with $\lambda_1>0$) and Theorem 6.1 (ii) that

$$\begin{aligned}
&E\{|\Delta X_{i+1}|^2\}\\
&\le E\Big\{[1+A_1h+(1+\lambda_2)A_2hL_1]|\Delta X_i|^2+(1+\lambda_2^{-1})A_2h|\Delta u^n(X_i^n)|^2\\
&\quad +2(1+\lambda_1^{-1})\int_{t_i}^{t_{i+1}}[|\Delta\alpha_t|^2+|\Delta\beta_t|^2]dt\Big\}\\
&\le E\Big\{[1+A_1h+(1+\lambda_2)A_2hL_1]|\Delta X_i|^2\\
&\quad +(1+\lambda_2^{-1})A_2h\Big[G(\Delta u^n)|X_i^n|^2+H(\Delta u^n)\Big]\Big\}+C(1+\lambda_1^{-1})(1+|x|^2)h^2.
\end{aligned}$$

Note that $\Delta X_{i_0}=0$, and by (5.5), we have

$$\sup_{i_0\le i\le n}E\{|X_i^n|^2\}\le\Big[|x|^2+[B_1+A_2H(u^n)]T\Big]\Big[e^{[A_1+A_2G(u^n)]T}\vee 1\Big]\triangleq\tilde A.$$

Hence, by similar arguments as in Theorem 5.3 we get

$$\begin{aligned}
&\sup_{i_0\le i\le n}E\{|\Delta X_i|^2\}\le\Gamma_0^{n-i_0}(A_1+(1+\lambda_2)A_2L_1)\times\\
&\quad\Big[(1+\lambda_2^{-1})A_2[G(\Delta u^n)\tilde A+H(\Delta u^n)]+C(1+\lambda_1^{-1})[1+|x|^2]h\Big].
\end{aligned}\tag{6.7}$$

Next, denote

$$\gamma_{i+1}^1\triangleq\frac{1}{h}\int_{t_i}^{t_{i+1}}f(t,X_t,Y_t,Z_t)dt-f(t_i,X_{t_i},Y_{t_{i+1}},\hat Z_{t_i}),$$

where

$$\hat Z_{t_i}\triangleq\frac{1}{h}E_{t_i}\Big\{Y_{t_{i+1}}\Delta W_{i+1}\Big\}.\tag{6.8}$$

Then, by (6.5),

$$\begin{aligned}
&E\{|\gamma_{i+1}^1|^2\}\le\frac{1}{h}E\Big\{\int_{t_i}^{t_{i+1}}|f(t,X_t,Y_t,Z_t)-f(t_i,X_{t_i},Y_{t_{i+1}},\hat Z_{t_i})|^2dt\Big\}\\
&\le\frac{C}{h}E\Big\{\int_{t_i}^{t_{i+1}}\Big[h+|X_t-X_{t_i}|^2+|Y_t-Y_{t_i}|^2+|Z_t-\hat Z_{t_i}|^2]dt\Big\}\\
&\le C[1+|x|^2]h+\frac{C}{h}E\Big\{\int_{t_i}^{t_{i+1}}[|Z_t-\tilde Z_{t_i}|^2+|\tilde Z_{t_i}-\hat Z_{t_i}|^2]dt\Big\}\\
&\le C[1+|x|^2]h+CE\Big\{\int_{t_i}^{t_{i+1}}|Z_t|^2dt\Big\}+\frac{C}{h}E\Big\{\int_{t_i}^{t_{i+1}}|Z_t-\tilde Z_{t_i}|^2dt\Big\}.
\end{aligned}\tag{6.9}$$

Here we made use of the estimate

$$
\begin{aligned}
& h^2E\{|\tilde{Z}_{t_i} - \hat{Z}_{t_i}|^2\} \\
= \;& E\Big\{\Big|E_{t_i}\{\int_{t_i}^{t_{i+1}} Z_t dt\} \\
& -E_{t_i}\{[Y_{t_i} - \int_{t_i}^{t_{i+1}} f(t, X_t, Y_t, Z_t)dt + \int_{t_i}^{t_{i+1}} Z_t dW_t]\Delta W_{i+1}\}\Big|^2\Big\} \\
= \;& E\Big\{\Big|E_{t_i}\{\int_{t_i}^{t_{i+1}} f(t, X_t, Y_t, Z_t)dt\Delta W_{t_{i+1}}\Big|^2\Big\} \\
\le \;& E\Big\{\Big|\int_{t_i}^{t_{i+1}} f(t, X_t, Y_t, Z_t)dt\Big|^2\Big\}E\{|\Delta W_{i+1}|^2\} \\
\le \;& h^2E\Big\{\int_{t_i}^{t_{i+1}} |f(t, X_t, Y_t, Z_t)|^2 dt\Big\} \\
\le \;& Ch^2E\Big\{\int_{t_i}^{t_{i+1}} [1 + |X_t|^2 + |Y_t|^2 + |Z_t|^2]dt \\
\le \;& C[1 + |x|^2]h^3 + Ch^2E\Big\{\int_{t_i}^{t_{i+1}} |Z_t|^2 dt\Big\}.
\end{aligned}
$$

Applying Lemma 3.3 we get, under (3.5),

$$
\begin{aligned}
& E\{|\Delta Y_i|^2\} \le E\Big\{[1 + A_4 h]|\Delta Y_{i+1}|^2 + A_5 h|\Delta X_i|^2 \\
& +C[\lambda_1^{-1} + \lambda_3^{-1}h + h]\Big[(1 + |x|^2)h^2 + h\int_{t_i}^{t_{i+1}} |Z_t|^2 dt + \int_{t_i}^{t_{i+1}} |Z_t - \tilde{Z}_{t_i}|^2 dt\Big]\Big\}.
\end{aligned}
$$

Then, by (6.4), (6.5), and (6.7),

$$
\begin{aligned}
&|\Delta u^n_{i_0}(x)|^2 = |\Delta Y_{i_0}|^2\\
&\le (1+A_4h)^{n-i_0}E\{|\Delta Y_n|^2\} + A_5\Gamma_0^{n-i_0}(A_4)\sup_{i_0\le i\le n}E\{|\Delta X_i|^2\}\\
&\quad +C[\lambda_1^{-1}+\lambda_3^{-1}h+h](1+A_4^+h)^{n-i_0}\times\\
&\quad E\Big\{\Big[(1+|x|^2)h + h\int_{t_{i_0}}^T |Z_t|^2dt + \sum_{i=i_0}^{n-1}\int_{t_i}^{t_{i+1}}|Z_t-\tilde Z_{t_i}|^2dt\Big]\Big\}\\
&\le \Big[(1+A_4h)^{n-i_0}g_x + A_5\Gamma_0^{n-i_0}(A_4)\Big]\sup_{i_0\le i\le n}E\{|\Delta X_i|^2\}\\
&\quad +C[\lambda_1^{-1}+\lambda_3^{-1}h+h]e^{A_4^+T}(1+|x|^2)h\\
&\le \Big[(1+\lambda_2^{-1})A_2[G(\Delta u^n)\tilde A + H(\Delta u^n)] + C(1+\lambda_1^{-1})[1+|x|^2]h\Big]\times\\
&\quad \Big[g_x(1+A_4h)^{n-i_0}\Gamma_0^{n-i_0}(A_1+(1+\lambda_2)A_2L_1)\\
&\quad +A_5\Gamma_0^{n-i_0}(A_4)\Gamma_0^{n-i_0}(A_1+(1+\lambda_2)A_2L_1)\Big]\\
&\quad +C[\lambda_1^{-1}+\lambda_3^{-1}h+h]e^{A_4^+T}(1+|x|^2)h\\
&\le (1+\lambda_2^{-1})A_2[G(\Delta u^n)\tilde A + H(\Delta u^n)]C_\lambda\\
&\quad +C\Big[C_\lambda(1+\lambda_1^{-1}) + [\lambda_1^{-1}+\lambda_3^{-1}h+h]e^{A_4^+T}\Big](1+|x|^2)h,
\end{aligned}
$$

with

$$C_\lambda \stackrel{\triangle}{=} g_x(1+A_4h)^{n-i_0}\Gamma_0^{n-i_0}(A_1+(1+\lambda_2)A_2L_1)+A_5\Gamma_0^{n-i_0}(A_4)\Gamma_0^{n-i_0}(A_1+(1+\lambda_2)A_2L_1).$$

By (4.10) we get

$$
\begin{aligned}
&\sup_i |\Delta u^n_i(x)|^2\\
&\le \Big[G(\Delta u^n)\tilde A + H(\Delta u^n)\Big]c_1(\lambda,h,L_1)\\
&\quad +C\Big[(1+\lambda_1^{-1}) + [\lambda_1^{-1}+\lambda_3^{-1}h+h]e^{A_4^+T}\Big](1+|x|^2)h\\
&\le c_1(\lambda,h,L_1)\Big\{G(\Delta u^n)\Big[e^{[A_1+A_2G(u^n)]T}\vee 1\Big]\Big[|x|^2+[B_1+A_2H(u^n)]T\Big] + H(\Delta u^n)\Big\}\\
&\quad +C\Big[\frac{c_1(\lambda,h,L_1)}{b_y+\sigma_y}(1+\lambda_1^{-1}) + [\lambda_1^{-1}+\lambda_3^{-1}h+h]e^{A_4^+T}\Big](1+|x|^2)h.
\end{aligned}
$$

Thus

$$
\begin{aligned}
G(\Delta u^n) &\le c_2(\lambda,h,L_1,G(u^n))G(\Delta u^n)\\
&\quad +C\Big[\tfrac{c_1(\lambda,h,L_1)}{b_y+\sigma_y}(1+\lambda_1^{-1}) + [\lambda_1^{-1}+\lambda_3^{-1}h+h]e^{A_4^+T}\Big]h;\\
H(\Delta u^n) &\le c_1(\lambda,h,L_1)H(\Delta u^n) + c_2(\lambda,h,L_1,G(u^n))[B_1+A_2H(u^n)]TG(\Delta u^n)\\
&\quad +C\Big[\tfrac{c_1(\lambda,h,L_1)}{b_y+\sigma_y}(1+\lambda_1^{-1}) + [\lambda_1^{-1}+\lambda_3^{-1}h+h]e^{A_4^+T}\Big]h.
\end{aligned}
\qquad (6.10)
$$

Fix some c_2, $c_2(L_1, L_1) < c_2 < 1$. In dependence of c_2 we may and do choose some $\lambda_1 > 0$ such that

$$\lim_{h\downarrow 0} c_2(\lambda_1, \lambda_2^*, \lambda_3(h), \lambda_4(h), h, L_1, L_1) < c_2,$$

where λ_2^* is the minimum argument of $c_2(\lambda_2, L_1, L_1)$ and

$$\lambda_3(h) \triangleq \sqrt{h}; \quad \lambda_4(h) \triangleq \frac{1-\sqrt{h}-K(1+\lambda_1 h)h - 2K\sqrt{h}}{1+\lambda_1 h}.$$

With this choice, $A_3 = 1$ holds true, and, for sufficiently small h, we obtain

$$c_1(\lambda, h, L_1) \le c_2(\lambda, h, L_1, G(u^n)) \le c_2 < 1.$$

Consequently (6.10) implies that for sufficiently small h

$$G(\Delta u^n) \le Ch; \quad H(\Delta u^n) \le Ch,$$

and the assertion is proved. ■

As a direct consequence of Theorems 5.1 and 6.4 we have

Theorem 6.5 *Under the assumptions of Theorem 6.4 we have for any* $c_2(L_1, L_1) < c_2 < 1$ *and for* h *small enough*

$$|u_i^{n,m}(x) - u(t_i, x)|^2 \le C(1+|x|^2)[mc_2^m + h].$$

We close the theoretical part of this paper with a precise version of the generic Theorem 2.3.

Theorem 6.6 *Under the assumptions of Theorem 6.4 we have for any* $c_2(L_1, L_1) < c_2 < 1$ *and for* h *small enough*

$$\sup_{1\le i\le n} E\Big\{ \sup_{t\in[t_{i-1},t_i]} [|X_t - X_{i-1}^{n,m}|^2 + |Y_t - Y_{i-1}^{n,m}|^2]\Big\} + \sum_{i=1}^n E\Big\{\int_{t_{i-1}}^{t_i} |Z_t - \hat{Z}_{i-1}^{n,m}|^2 dt\Big\}$$
$$\le C(1+|x|^2)[mc_2^m + h].$$

This theorem follows from the previous one by arguments which are fairly standard. A detailed proof can be found in the appendix.

7 A Numerical Algorithm

We now briefly explain how the discretized Markovian iteration above can be transformed into a numerical algorithm which is viable also for high-dimensional problems. To this end we replace the conditional expectations by a simulation based least squares regression estimator, as was suggested e.g. by Gobet et al. (2005) and Bender and Denk (2005) in the context of decoupled FBSDEs. An alternative estimator based on Malliavin calculus is discussed in Bouchard and Touzi (2004) for decoupled FBSDEs.

For the reader's convenience we spell out our algorithm for the coupled case. While a convergence analysis is out of the scope of the present paper, we will illustrate the algorithm by some numerical examples in the next section.

We assume that the number of time steps n is fixed for the remainder of this section. In the algorithm conditional expectations are first replaced by orthogonal projections on K basis functions. Then the orthogonal projections are approximated by simulating Λ trajectories. Hence, the algorithm can be described for the one-dimensional case iteratively as follows. It is straightforward how this extends to the multi-dimensional case.

- Fix some x_0. Set $\widetilde{u}_i^{n,0,K,\Lambda}(x) = 0$.
- Sample Λ independent copies of the time discretized Brownian motion $W_{t_i}^\lambda$, $i = 0, \ldots, n$, $\lambda = 1, \ldots, \Lambda$ starting in 0 and denote the corresponding increments by ΔW_i^λ.
- Suppose $\widetilde{u}_i^{n,m-1,K,\Lambda}(x)$ is already constructed. Let $\widetilde{X}_0^{n,m,\lambda} = x_0$ and
$$\begin{aligned}\widetilde{X}_{i+1}^{n,m,\lambda} &= \widetilde{X}_i^{n,m,\lambda} + b(t_i, \widetilde{X}_i^{n,m,\lambda}, \widetilde{u}_i^{n,m-1,K,\Lambda}(\widetilde{X}_i^{n,m,\lambda}))h \\ &\quad +\sigma(t_i, \widetilde{X}_i^{n,m,\lambda}, \widetilde{u}_i^{n,m-1,K,\Lambda}(\widetilde{X}_i^{n,m,\lambda}))\Delta W_{i+1}^\lambda,\end{aligned}$$
where – for notational convenience – we suppress the dependence of $\widetilde{X}_i^{n,m,\lambda}$ on K through $\widetilde{u}^{n,m-1,K,\Lambda}$. Note, $\widetilde{X}_i^{n,m,\lambda_0}$ depends on all Brownian increments ΔW_i^λ, $i = 1, \ldots, n$, $\lambda = 1, \ldots, \Lambda$ through $\widetilde{u}_i^{n,m-1,K,\Lambda}$. While we expect, that this dependence will make a convergence analysis difficult, the examples below indicate that the algorithm works without re-simulating the Brownian paths in every iteration step.
- Choose a set of Lipschitz continuous basis functions
$$\mathcal{B}_i^{n,m,K} = \left\{\eta_i^{n,m,k}(x), \quad k = 1, \ldots, K\right\}$$
such that
$$\left\{\eta_i^{n,m,k}(\widetilde{X}_i^{n,m,\lambda}), \quad k = 1, \ldots, K\right\} \tag{7.1}$$
forms a subset of $L^2(\Omega, \widetilde{X}_i^{n,m,\lambda})$. From the construction below, it will become evident, that $\widetilde{u}_i^{n,m,K,\Lambda}(x)$ inherits the Lipschitz continuity from the basis functions. This feature seems to be important to ensure that the discretized forward equations for $\widetilde{X}^{n,m+1,\lambda}$ do not explode.

- Let
$$\mathcal{A}_i^{n,m,K,\Lambda} = \frac{1}{\sqrt{\Lambda}} \left(\eta_i^{n,m,k}(\widetilde{X}_i^{n,m,\lambda}) \right)_{\lambda=1,\ldots,\Lambda,\; k=1,\ldots,K}.$$
Define, for $i = n-1, \ldots, 1$
$$\begin{aligned}
\tilde{u}_n^{n,m,K,L}(x) &= g(x) \\
\tilde{v}_n^{n,m,K,L}(x) &= 0 \\
\widetilde{Y}_{i+1}^{n,m,K,\lambda} &= \tilde{u}_{i+1}^{n,m,K,\Lambda}(\widetilde{X}_{i+1}^{n,m,\lambda}) \\
\widetilde{Z}_{i+1}^{n,m,K,\lambda} &= \tilde{v}_{i+1}^{n,m,K,\Lambda}(\widetilde{X}_{i+1}^{n,m,\lambda}) \\
\beta_{i,\cdot}^{n,m,K,\Lambda} &= \frac{1}{\sqrt{\Lambda}} \left(\mathcal{A}_i^{n,m,K,\Lambda} \right)^{\oplus} \left(\frac{1}{h} \widetilde{Y}_{i+1}^{n,m,K,\cdot} \Delta W_{i+1}^{\cdot} \right) \\
\tilde{v}_i^{n,m,K,\Lambda}(x) &= \sum_{k=1}^{K} \beta_{i,k}^{n,m,K,\Lambda} \eta_i^{n,m,k}(x) \\
\alpha_{i,\cdot}^{n,m,K,\Lambda} &= \frac{1}{\sqrt{\Lambda}} \left(\mathcal{A}_i^{n,m,K,\Lambda} \right)^{\oplus} \left(\widetilde{Y}_{i+1}^{n,m,K,\cdot} + f(t_i, \widetilde{X}_i^{n,m,\cdot}, \widetilde{Y}_{i+1}^{n,m,K,\cdot}, \widetilde{Z}_i^{n,m,K,\cdot}) h \right) \\
\tilde{u}_i^{n,m,K,\Lambda}(x) &= \sum_{k=1}^{K} \alpha_{i,k}^{n,m,K,\Lambda} \eta_i^{n,m,k}(x)
\end{aligned}$$
where the $\oplus$ denotes the pseudo inverse. Recall here that, by the definition of the pseudo inverse, e.g. $\widetilde{Y}_i^{n,m,K,\cdot} = \tilde{u}_i^{n,m,K,\Lambda}(\widetilde{X}_i^{n,m,\cdot})$ satisfies
$$\begin{aligned}
\widetilde{Y}_i^{n,m,K,\cdot} = \operatorname{arginf} &\left\{ \frac{1}{\Lambda} \sum_{\lambda=1}^{\Lambda} \left| \widetilde{Y}_{i+1}^{n,m,K,\lambda} + f(t_i, \widetilde{X}_i^{n,m,\lambda}, \widetilde{Y}_{i+1}^{n,m,K,\lambda}, \widetilde{Z}_i^{n,m,K,\cdot}) h - y^{\lambda} \right|^2 ; \right. \\
&\left. y = U(\widetilde{X}_i^{n,m,\cdot}),\; U \in \mathcal{B}_i^{n,m,K} \right\}.
\end{aligned}$$

- Let
$$\begin{aligned}
\widetilde{Y}_1^{n,m,K,\lambda} &= \tilde{u}_1^{n,m,K,\Lambda}(\widetilde{X}_1^{n,m,\lambda}) \\
\widetilde{Z}_1^{n,m,K,\lambda} &= \tilde{v}_1^{n,m,K,\Lambda}(\widetilde{X}_1^{n,m,\lambda}) \\
\widetilde{Z}_0^{n,m,K,\lambda} &= \frac{1}{\Lambda} \sum_{\bar{\lambda}=1}^{\Lambda} \frac{1}{h} \widetilde{Y}_1^{n,m,K,\bar{\lambda}} \Delta W_1^{\bar{\lambda}} \\
\widetilde{Y}_0^{n,m,K,\lambda} &= \frac{1}{\Lambda} \sum_{\bar{\lambda}=1}^{\Lambda} Y_1^{n,m,K,\bar{\lambda}} + f(0, x_0, \widetilde{Y}_1^{n,m,K,\bar{\lambda}}, \widetilde{Z}_0^{n,m,K,\bar{\lambda}}) h.
\end{aligned}$$

We expect that the thus constructed $(\widetilde{X}^{n,m,\lambda}, \widetilde{Y}^{n,m,K,\lambda}, \widetilde{Z}^{n,m,K,\lambda})$ are 'close' to $(X^{n,m,\lambda}, Y^{n,m,\lambda}, \hat{Z}^{n,m,\lambda})$, the solution of the discretized Markovian iteration (1.8) with the Brownian motion W replaced by W^{λ}, if the basis functions are chosen appropriately and the number Λ of simulated paths is sufficiently large. While an analysis of the error by estimating the conditional expectations is left to future research, the numerical examples in the next section support this conjecture.

8 Numerical Examples

For the simulations we consider the example

$$\begin{cases} X_{d,t} = x_{d,0} + \int_0^t \sigma Y_u dW_{d,u} \\ Y_t = \sum_{d=1}^{D} \sin(X_{d,T}) + \int_t^T -rY_u + \frac{1}{2} e^{-3r(T-u)} \sigma^2 \left(\sum_{d=1}^{D} \sin(X_{d,u}) \right)^3 du \\ \qquad - \int_t^T \sum_{d=1}^{D} Z_{d,u} dW_{d,t}, \end{cases}$$

where $W_{d,t}, d = 1, \ldots, D$ is a D-dimensional Brownian motion and $\sigma > 0$, r, $x_{d,0}$ are constants. Note that the corresponding differential operator degenerates at $y = 0$.

By Itô's formula one can easily check that this FBSDE decouples via the relation

$$Y_t = e^{-r(T-t)} \sum_{d=1}^{D} \sin(X_{d,t}) \tag{8.1}$$

Note, that for small σ the weak coupling condition of Y into X is satisfied, while, for large σ, the monotonicity condition of f can be fulfilled by choosing r large enough.

In the simulations we replace conditional expectations by least squares regression as explained above with the 'canonical' basis functions

$$1, \quad x_d, \; 1 \le d \; \le D, \quad (-R) \vee (x_d x_q) \wedge R, \; 1 \le d \le q \le D,$$

i.e. monomials up to order two in $x = (x_1, \ldots, x_D)$. The truncation constant R guarantees that the basis functions are Lipschitz continuous. We set

$$R = 10, \quad X_{d,0} = \pi, 1 \le d \le D, \quad T = 1, \quad \Lambda = 50000, \quad n = 50.$$

With this initial condition we get $Y_0 = De^{-r(T-t)}$. Recall also that the estimator $\widetilde{Y}_0^{n,m,K,\lambda}$ of Y_0 does not depend on λ. We stop the iteration when two consecutive estimates $\widetilde{Y}_0^{n,m,K} := \widetilde{Y}_0^{n,m,K,\lambda}$ are within a distance of 10^{-4}. This iteration level is denoted m_{stop}.

From (8.1) we can also approximate the true value of X via the usual Euler scheme (applying the same simulated Brownian increments ΔW_i^{λ}). The corresponding approximation along the λth path is denoted $\check{X}_i^{n,\lambda}$, and hence

$$\check{Y}_i^{n,\lambda} = e^{-r(T-t_i)} \sum_{d=1}^{D} \sin(\check{X}_{d,i}^{n,\lambda})$$

may be considered a close approximation of Y_{t_i}. In the figures below we display a comparison between a typical path of $\check{Y}_i^{n,\lambda_0}$ (dashed line) and $\widetilde{Y}_i^{n,m_{stop},K,\lambda_0}$ (solid line) as well as the (absolute) empirical mean square error between $\check{Y}_i^{n,\lambda}$ and $\widetilde{Y}_i^{n,m_{stop},K,\lambda}$, $\lambda = 1, \dots, \Lambda$. Precisely, the figures on the right hand side show

$$\frac{1}{\Lambda} \sum_{\lambda=1}^{\Lambda} |\widetilde{Y}_i^{n,m_{stop},K,\lambda} - \check{Y}_i^{n,\lambda}|^2$$

as function of time. We consider the following cases:
Case 1: $D = 4$, $\sigma = 0.2$, $r = 0$. Then, $m_{stop} = 9$.
Case 2: $D = 4$, $\sigma = 0.4$, $r = 0$. Then, $m_{stop} = 42$.
Case 3: $D = 4$, $\sigma = 0.4$, $r = 1$. Then, $m_{stop} = 11$.
Case 4: $D = 10$, $\sigma = 0.1$, $r = 0$. Then, $m_{stop} = 12$.

The numbers of iteration required for termination and the figures below (for cases 1–3) clearly illustrate how the convergence quality decreases in σ and increases in r, as expected from the above discussion. We also report that for $D = 4$, $\sigma = 1$, $r = 0$ the coupling apparently becomes too strong for the algorithm to converge. Indeed, we obtained values $\widetilde{Y}_0^{n,m,K} \approx 35$ for odd m and $\widetilde{Y}_0^{n,m,K} \approx 0.63$ for even m which did not change significantly with increasing number of iterations. We hence terminated the experiment after 75 iteration steps. However, convergence can be enforced by increasing r. For instance, for $D = 4$, $\sigma = 1$, $r = 2.5$, the algorithm terminated after 30 steps. The estimated value $\widehat{Y}_0^{n,30,K,\lambda}$ is 0.320 while the true value Y_0 is 0.328. To demonstrate that the space dimension four is no limitation for the proposed method, we additionally include case 4 where X takes values in $\mathbb{R}^{10}$.

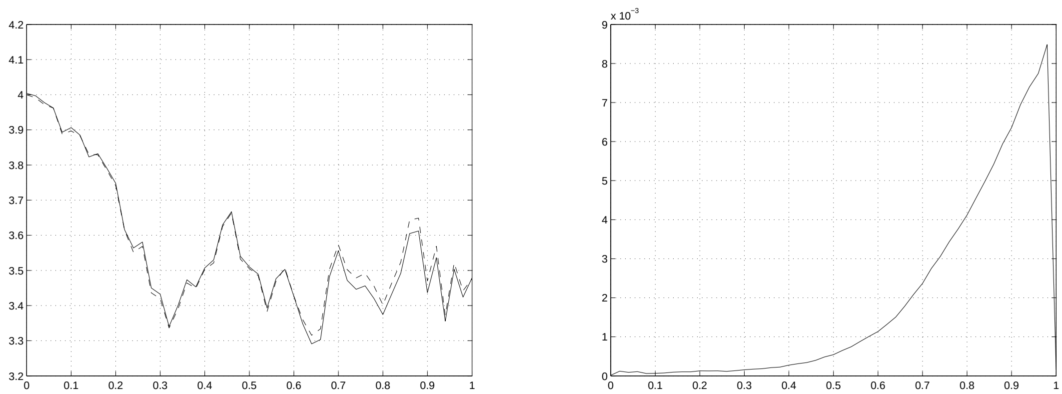

Figure 1: Case 1

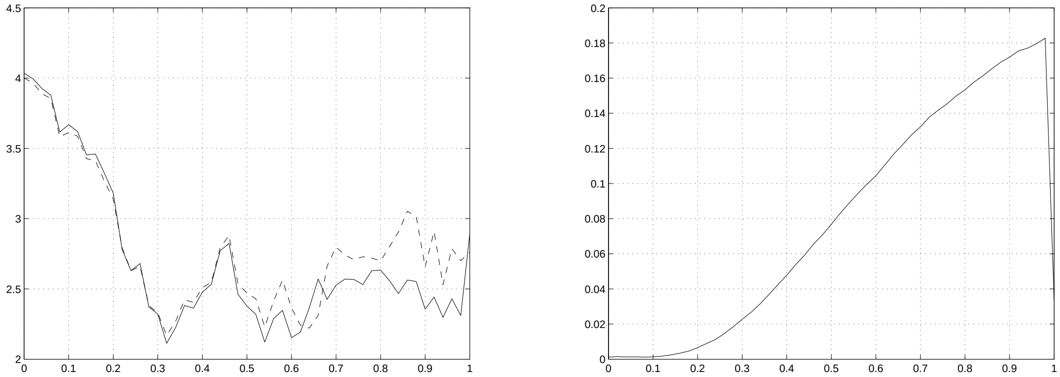

Figure 2: Case 2

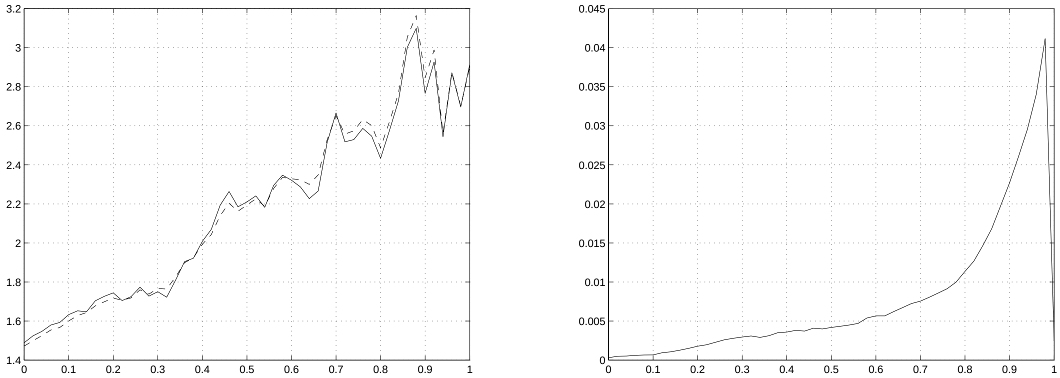

Figure 3: Case 3

A Proofs of Theorem 6.1 and Corollary 6.2

Denote

$$\begin{aligned}
\bar{A}_1 &\stackrel{\triangle}{=} 2k_b + 1 + \sigma_x;\\
\bar{A}_2 &\stackrel{\triangle}{=} b_y + \sigma_y;\\
\bar{A}_3 &\stackrel{\triangle}{=} \lambda_4;\\
\bar{A}_4 &\stackrel{\triangle}{=} 2k_f + 1 + \lambda_4^{-1} f_z;\\
\bar{A}_5 &\stackrel{\triangle}{=} f_x.
\end{aligned} \tag{A.1}$$

A.1 Lipschitz continuity

Lemma A.1 *Given (φ_1, φ_2), (x_1, x_2), and t_0, for $i = 1, 2$ and $\Theta = X, Y, Z$ denote*

$$\Theta^i \stackrel{\triangle}{=} \Theta^{\varphi, t_0, x_i}; \quad \Delta x \stackrel{\triangle}{=} x_1 - x_2; \quad \Delta\varphi \stackrel{\triangle}{=} \varphi^1 - \varphi^2; \quad \Delta\Theta \stackrel{\triangle}{=} \Theta^1 - \Theta^2.$$

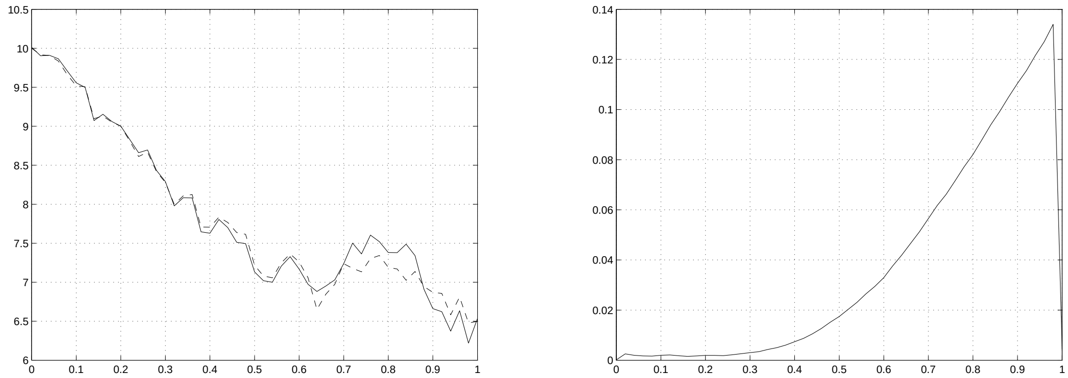

Figure 4: Case 4

Then, for any $\lambda_j \geq 0$,

$$\begin{aligned}
&\frac{d}{dt}E\{|\Delta X_t|^2\} \leq [\bar{A}_1 + (1+\lambda_2)\bar{A}_2 L(\varphi^1)]E\{|\Delta X_t|^2\} \\
&\quad +(1+\lambda_2^{-1})\bar{A}_2 E\{|\Delta\varphi(X_t^2)|^2\}; \\
&-\frac{d}{dt}E\{|\Delta Y_t|^2\} + (1-\bar{A}_3)|\Delta Z_t|^2 \leq \bar{A}_4 E\{|\Delta Y_t|^2\} + \bar{A}_5 E\{|\Delta X_t|^2\};
\end{aligned}$$

Proof. Applying Ito's formula we have

$$\begin{aligned}
d(|\Delta X_t|^2) &= 2\Delta X_t \Delta b dt + 2\Delta X_t \Delta\sigma dW_t + |\Delta\sigma|^2 dt \\
&\leq 2\Delta X_t \Delta\sigma dW_t + \Big[2k_b|\Delta X_t|^2 + |\Delta X_t|^2 + b_y|\varphi_1(X_t^1) - \varphi_2(X_t^2)|^2 \\
&\quad +\sigma_x|\Delta X_t|^2 + \sigma_y|\varphi_1(X_t^1) - \varphi_2(X_t^2)|^2\Big]dt.
\end{aligned}$$

Note that

$$|\varphi_1(X_t^1) - \varphi_2(X_t^2)|^2 \leq (1+\lambda_2)L(\varphi_1)|\Delta X_t|^2 + (1+\lambda_2^{-1})|\Delta\varphi(X_t^2)|^2.$$

Then

$$d(|\Delta X_t|^2) \leq 2\Delta X_t \Delta\sigma dW_t + \Big[[\bar{A}_1 + (1+\lambda_2)\bar{A}_2 L(\varphi^1)]|\Delta X_t|^2 + (1+\lambda_2^{-1})A_2|\Delta\varphi(X_t^2)|^2\Big]dt,$$

which implies the estimate for X.

Moreover,

$$\begin{aligned}
-d|\Delta Y_t|^2 &= -2\Delta Y_t \Delta Z_t dW_t + 2\Delta Y_t \Delta f dt - |\Delta Z_t|^2 dt \\
&\leq -2\Delta Y_t \Delta Z_t dW_t + \Big[2k_f|\Delta Y_t|^2 + |\Delta Y_t|^2 + f_x|\Delta X_t|^2 \\
&\quad +\lambda_4^{-1} f_z|\Delta Y_t|^2 + \lambda_4|\Delta Z_t|^2 - |\Delta Z_t|^2\Big]dt.
\end{aligned}$$

This proves the estimate for Y and Z. ■

Lemma A.2 *For any Lipschitz continuous φ,*

$$L(F(\varphi)) \le [g_x + f_x T]\Big[\exp\Big([\bar{A}_1 + \bar{A}_4]T + \bar{A}_2 T L(\varphi)\Big) \vee 1\Big];$$

provided

$$\bar{A}_3 \le 1. \tag{A.2}$$

Proof. Fix t_0 and (x_1, x_2). Setting $\lambda_2 = 0$ and $\varphi^1 = \varphi^2 = \varphi$, and applying Lemma A.1, we get

$$E\{|\Delta X_t|^2\} \le |\Delta x|^2 + [\bar{A}_1 + \bar{A}_2 L(\varphi^1)] \int_{t_0}^t E\{|\Delta X_s|^2\} ds.$$

Then

$$E\{|\Delta X_t|^2\} \le |\Delta x|^2 \exp\Big([\bar{A}_1 + \bar{A}_2 L(\varphi^1)][t - t_0]\Big).$$

Moreover, since $\bar{A}_3 = \lambda_4 \le 1$, we get

$$-d(e^{\bar{A}_4 t} E\{|\Delta Y_t|^2\}) \le \bar{A}_5 e^{\bar{A}_4 t} E\{|\Delta X_t|^2\} dt.$$

Then

$$\begin{aligned} E\{|\Delta Y_{t_0}|^2\} &\le g_x e^{\bar{A}_4(T-t_0)} |\Delta X_T|^2 + \bar{A}_5 \int_{t_0}^T e^{\bar{A}_4(t-t_0)} E\{|\Delta X_t|^2\} dt \\ &\le \Big[g_x e^{[\bar{A}_1 + \bar{A}_2 L(\varphi^1) + \bar{A}_4](T-t_0)} + \bar{A}_5 \int_{t_0}^T e^{[\bar{A}_1 + \bar{A}_2 L(\varphi^1) + \bar{A}_4](t-t_0)} dt\Big] |\Delta x|^2 \\ &\le [g_x + \bar{A}_5 T]\Big[e^{[\bar{A}_1 + \bar{A}_2 L(\varphi^1) + \bar{A}_4]T} \vee 1\Big] |\Delta x|^2. \end{aligned}$$

Note that $Y^i_{t_0} = F(\varphi)(t_0, x_i)$, we get

$$L(F(\varphi)(t_0, \cdot)) \le [g_x + \bar{A}_5 T]\Big[e^{[\bar{A}_1 + \bar{A}_2 L(\varphi^1) + \bar{A}_4]T} \vee 1\Big].$$

Since the right side at above is independent of t_0, the lemma is proved. ∎

For the remainder of this appendix we shall always assume that $\lambda_4 = \bar{A}_3 = 1$.

Lemma A.3 *Assume (3.2) holds true. Then*

$$L(\tilde{u}^m) \le L_1, \quad \forall m.$$

Proof. By Lemma A.3, with $\lambda_4 = 1$, we have

$$L(\tilde{u}^m) \le [g_x + \bar{A}_5 T]\Big[e^{[\bar{A}_1 + \bar{A}_4]T + \bar{A}_2 L(\tilde{u}^{m-1})T} + 1\Big].$$

Then

$$\bar{A}_2 T L(\tilde{u}^m) \le \bar{A}_2 T[g_x + \bar{A}_5 T]\Big[e^{[\bar{A}_1+\bar{A}_4]T+\bar{A}_2 L(\tilde{u}^{m-1})T} + 1\Big]$$

Note, that $L(\tilde{u}^0) = 0$. If $L_0 \le e^{-1}$, one can easily show by induction that

$$\bar{A}_2 T L(\tilde{u}^m) \le \bar{A}_2 T[g_x + \bar{A}_5 T] + 1.$$

Thus

$$L(\tilde{u}^m) \le [g_x + \bar{A}_5 T]\Big[e^{[\bar{A}_1+\bar{A}_4]T+\bar{A}_2 T[g_x+\bar{A}_5 T]+1} \vee 1\Big].$$

That is, $L(\tilde{u}^m) \le L_1$. ■

A.2 Linear growth

Denote

$$\bar{B}_1 \triangleq b_0 + \sigma_0; \quad \bar{B}_2 \triangleq f_0. \tag{A.3}$$

Lemma A.4 *Given φ, t_0, x, denote $\Theta \triangleq \Theta^{\varphi,t_0,x}$ for $\Theta = X, Y, Z$. Then*

$$\frac{d}{dt}E\{|X_t|^2\} \le [\bar{A}_1 + \bar{A}_2 G(\varphi)]E\{|X_t|^2\} + [\bar{B}_1 + \bar{A}_2 H(\varphi)];$$
$$-\frac{d}{dt}E\{|Y_t|^2\} + (1-\bar{A}_3)|Z_t|^2 \le \bar{A}_4 E\{|Y_t|^2\} + \bar{A}_5 E\{|X_t|^2\} + \bar{B}_2.$$

Proof. By Ito's formula we have

$$\begin{aligned}
&d(|X_t|^2) = 2X_t\sigma dW_t + 2X_t b dt + |\sigma|^2 dt\\
&\le 2X_t\sigma dW_t + \Big[2k_b|X_t|^2 + |X_t|^2 + b_0 + b_y|\varphi(X_t)|^2 + \sigma_0 + \sigma_x|X_t|^2 + \sigma_y|\varphi(X_t)|^2\Big]dt\\
&\le 2X_t\sigma dW_t + \Big[\bar{A}_1|X_t|^2 + \bar{B}_1 + \bar{A}_2[G(\varphi)|X_t|^2 + H(\varphi)]\Big]dt.
\end{aligned}$$

Taking the expectation yields the estimate for X.

Similarly,

$$\begin{aligned}
&-d|Y_t|^2 = -2Y_t Z_t dW_t + 2Y_t f dt - |Z_t|^2 dt\\
&= -2Y_t Z_t dW_t + 2Y_t\Big[[f(t,X_t,Y_t,Z_t) - f(t,X_t,0,Z_t)]\\
&\quad +[f(t,X_t,0,Z_t) - f(t,X_t,0,0)] + f(t,X_t,0,0)\Big]dt - |Z_t|^2 dt\\
&\le -2Y_t Z_t dW_t + \Big[2k_f|Y_t|^2 + \lambda_4^{-1} f_z|Y_t|^2 + \lambda_4|Z_t|^2 + |Y_t|^2 + f_0 + f_x|X_t|^2 - |Z_t|^2\Big]dt.
\end{aligned}$$

The estimate for (Y, Z) follows again by taking expectation. ■

Lemma A.5 *For any linear growing φ,*

$$G(F(\varphi)) \leq [g_x + \bar{A}_5 T]\Big[e^{[\bar{A}_1+\bar{A}_4]T+\bar{A}_2 T G(\varphi)} \vee 1\Big]; \tag{A.4}$$

$$H(F(\varphi)) \leq c_0(G(\varphi))H(\varphi) + L_2(G(\varphi)); \tag{A.5}$$

with c_0 and L_2 as defined at the beginning of Section 4.

Proof. Fix t_0 and x. Applying Lemma A.4 we have

$$E\{|X_t|^2\} \leq |x|^2 e^{[\bar{A}_1+\bar{A}_2 G(\varphi)][t-t_0]} + [\bar{B}_1 + \bar{A}_2 H(\varphi)] \int_{t_0}^{T} e^{[\bar{A}_1+\bar{A}_2 G(\varphi)][s-t_0]} ds.$$

Moreover, since $\lambda_4 = 1$,

$$\begin{aligned}
&|F(\varphi)(t_0,x)|^2 = |Y_{t_0}|^2 \\
&\leq e^{\bar{A}_4(T-t_0)} E\{|Y_T|^2\} + \int_{t_0}^{T} e^{\bar{A}_4[t-t_0]} \Big[\bar{A}_5 E\{|X_t|^2\} + \bar{B}_2\Big] dt \\
&\leq e^{\bar{A}_4(T-t_0)} \Big[g_0 + g_x|x|^2 e^{[\bar{A}_1+\bar{A}_2 G(\varphi)][T-t_0]} \\
&+ g_x[\bar{B}_1 + \bar{A}_2 H(\varphi)] \int_{t_0}^{T} e^{[\bar{A}_1+\bar{A}_2 G(\varphi)][s-t_0]} ds\Big] \\
&+ \bar{A}_5 \int_{t_0}^{T} e^{\bar{A}_4[t-t_0]} \Big[|x|^2 e^{[\bar{A}_1+\bar{A}_2 G(\varphi)][t-t_0]} + [\bar{B}_1 + \bar{A}_2 H(\varphi)] \int_{t_0}^{T} e^{[\bar{A}_1+\bar{A}_2 G(\varphi)][s-t_0]} ds\Big] dt \\
&+ \bar{B}_2 \int_{t_0}^{T} e^{\bar{A}_4[t-t_0]} dt.
\end{aligned}$$

This implies that

$$\begin{aligned}
G(\varphi(t_0,\cdot)) &\leq g_x e^{[\bar{A}_1+\bar{A}_4+\bar{A}_2 G(\varphi)][T-t_0]} + \bar{A}_5 \int_{t_0}^{T} e^{[\bar{A}_1+\bar{A}_4+\bar{A}_2 G(\varphi)][t-t_0]} dt \\
&\leq [g_x + \bar{A}_5 T]\Big[e^{[\bar{A}_1+\bar{A}_4+\bar{A}_2 G(\varphi)]T} \vee 1\Big]; \\
H(\varphi(t_0,\cdot)) &\leq \bar{A}_2 T[g_x + \bar{A}_5 T \Gamma_0(\bar{A}_4 T)]\Gamma_0([\bar{A}_1 + \bar{A}_2 G(\varphi)]T) H(\varphi) \\
&\quad + g_0 e^{\bar{A}_4(T-t_0)} + \bar{B}_1 T[g_x + \bar{A}_5 T \Gamma_0(\bar{A}_4 T)]\Gamma_0([\bar{A}_1 + \bar{A}_2 G(\varphi)]T) \\
&\quad + \bar{B}_2 \Gamma_0(\bar{A}_4 T) T.
\end{aligned}$$

The result now follows immediately. ■

Lemma A.6 *Assume (3.2) and (4.2) hold true. Then*

$$G(\tilde{u}^m) \leq L_1; \quad H(\tilde{u}^m) \leq \frac{L_2(L_1)}{1 - c_0(L_1)}; \quad \forall m.$$

Proof. By Lemma A.5, we have

$$\begin{array}{rcl} G(\tilde{u}^m) & \leq & [g_x + \bar{A}_5 T]\Big[e^{[\bar{A}_1+\bar{A}_4]T+\bar{A}_2 T G(\tilde{u}^{m-1})} \vee 1\Big]; \\ H(\tilde{u}^m) & \leq & c_0(G(\tilde{u}^{m-1}))H(\tilde{u}^{m-1}) + L_2(G(\tilde{u}^{m-1})). \end{array}$$

As in Lemma A.3 we obtain $G(\tilde{u}^m) \leq L_1$. Thus,

$$H(\tilde{u}^m) \leq c_0(L_1)H(\tilde{u}^{m-1}) + L_2(L_1).$$

Therefore, $H(\tilde{u}^m) \leq \frac{L_2(L_1)}{1-c_0(L_1)}$. ■

A.3 Convergence

Lemma A.7 *Assume φ^1, φ^2 have linear growth and φ^1 is Lipschitz continuous. Then*

$$\begin{array}{rcl} G(F(\varphi^1) - F(\varphi^2)) & \leq & c_2(\lambda_2, L(\varphi^1), G(\varphi^2))G(\Delta\varphi) \\ H(F(\varphi^1) - F(\varphi^2)) & \leq & c_1(\lambda_2, L(\varphi^1))H(\Delta\varphi) \\ & & +c_2(\lambda_2, L(\varphi^1), G(\varphi^2))[\bar{B}_1 + \bar{A}_2 H(\varphi^2)]TG(\Delta\varphi). \end{array}$$

Proof. We define $L \triangleq L(\varphi_1), G \triangleq G(\varphi^2)$. Moreover, we fix (t_0, x) and denote, for $l = 1, 2$ and $\Theta = X, Y, Z$

$$\Phi^l \triangleq F(\varphi^l); \quad \Theta^l \triangleq \Theta^{\varphi^l, t_0, x}, \quad \Delta\Theta \triangleq \Theta^1 - \Theta^2; \Delta\varphi \triangleq \varphi^1 - \varphi^2; \quad \Delta\Phi \triangleq \Phi^1 - \Phi^2.$$

Applying Lemma A.1 and noting that $\Delta X_{t_0} = 0$, we get

$$\begin{array}{l} \sup\limits_{t_0 \leq t \leq T} E\{|\Delta X_t|^2\} \\ \leq (1 + \lambda_2^{-1})\bar{A}_2\Big[G(\Delta\varphi) \sup\limits_{t_0 \leq t \leq T} E\{|X_t^2|^2\} + H(\Delta\varphi)\Big] \\ \quad \times (T - t_0)\Gamma_0([\bar{A}_1 + (1 + \lambda_2)\bar{A}_2 L](T - t_0)). \end{array}$$

From Lemma A.4, we obtain

$$\sup_{t_0 \leq t \leq T} E\{|X_t^2|^2\} \leq \Big[|x|^2 + [\bar{B}_1 + \bar{A}_2 H(\varphi^2)]T\Big]\Big[e^{[\bar{A}_1+\bar{A}_2 G]T} \vee 1\Big] \triangleq \tilde{A}.$$

Thus

$$\begin{array}{l} \sup\limits_{t_0 \leq t \leq T} E\{|\Delta X_t|^2\} \\ \leq (1 + \lambda_2^{-1})\bar{A}_2\Big[G(\Delta\varphi)\tilde{A} + H(\Delta\varphi)\Big](T - t_0)\Gamma_0([\bar{A}_1 + (1 + \lambda_2)\bar{A}_2 L][T - t_0]). \end{array}$$

Applying the second inequality of Lemma A.1 with $\lambda_4 = 1$,we get

$$
\begin{aligned}
&|\Delta\Phi(t_0,x)|^2 = |\Delta Y_{t_0}|^2 \\
&\leq e^{\bar{A}_4[T-t_0]}E\{|\Delta Y_T|^2\} + \bar{A}_5(T-t_0)\Gamma_0(\bar{A}_4[T-t_0]) \sup_{t_0\leq t\leq T} E\{|\Delta X_t|^2\} \\
&\leq \Big[e^{\bar{A}_4[T-t_0]}g_x + \bar{A}_5(T-t_0)\Gamma_0(\bar{A}_4[T-t_0])\Big] \times \\
&\quad (1+\lambda_2^{-1})\bar{A}_2\Big[G(\Delta\varphi)\tilde{A} + H(\Delta\varphi)\Big](T-t_0)\Gamma_0([\bar{A}_1 + (1+\lambda_2)\bar{A}_2 L][T-t_0]) \\
&\leq (1+\lambda_2^{-1})\bar{A}_2\Big[G(\Delta\varphi)\tilde{A} + H(\Delta\varphi)\Big]T \times \\
&\quad \Big[g_x\Gamma_1(\bar{A}_4T, [\bar{A}_1 + (1+\lambda_2)\bar{A}_2L]T) + \bar{A}_5T\Gamma_0(\bar{A}_4T)\Gamma_0([\bar{A}_1 + (1+\lambda_2)\bar{A}_2L]T)\Big].
\end{aligned}
$$

We can now plug in the definition of the constants to prove the assertion. ■

Lemma A.8 *Assume (3.2) and (5.1) hold true. Let* $\bar{H} \triangleq \frac{L_2(L_1)}{1-c_0(L_1)}$, $c_2 \triangleq c_2(L_1, L_1)$, *and* $c_1 \triangleq c_1(\lambda_2, L_1)$ *where* λ_2 *is the minimum argument of* $c_2(\lambda_2, L_1, L_1)$. *Then*

$$
\begin{aligned}
&G(\tilde{u}^m - \tilde{u}^{m-1}) \leq L_1 c_2^{m-1}; \\
&H(\tilde{u}^m - \tilde{u}^{m-1}) \leq \bar{H}c_1^{m-1} + \Big[[b_0+\sigma_0] + [b_y+\sigma_y]\bar{H}\Big]TL_1(m-1)c_2^{m-1}.
\end{aligned}
$$

Proof. Let λ_2 denote the minimum argument of $c_2(\lambda_2, L_1, L_1)$. By Lemmas A.2 and A.5 we have

$$
L(\tilde{u}^m) \leq L_1; \quad G(\tilde{u}^m) \leq L_1; \quad H(\tilde{u}^m) \leq \bar{H}.
$$

Thanks to Lemma A.7, we get,

$$
\begin{aligned}
G(\tilde{u}^m - \tilde{u}^{m-1}) &\leq c_2 G(\tilde{u}^{m-1} - \tilde{u}^{m-2}); \\
H(\tilde{u}^m - \tilde{u}^{m-1}) &\leq c_1 H(\tilde{u}^{m-1} - \tilde{u}^{m-2}) + c_2[\bar{B}_1 + \bar{A}_2\bar{H}]TG(\tilde{u}^{m-1} - \tilde{u}^{m-2}).
\end{aligned}
$$

Note that

$$
G(\tilde{u}^1 - \tilde{u}^0) = G(\tilde{u}^1) \leq L_1; \quad H(\tilde{u}^1 - \tilde{u}^0) = H(\tilde{u}^1) \leq \bar{H};.
$$

Therefore, we have

$$
G(\tilde{u}^m - \tilde{u}^{m-1}) \leq L_1 c_2^{m-1};
$$

and then

$$
\begin{aligned}
&H(\tilde{u}^m - \tilde{u}^{m-1}) \leq c_1^{m-1}H(\tilde{u}^1 - \tilde{u}^0) \\
&\quad + c_2\Big[[b_0+\sigma_0] + [b_y+\sigma_y]\bar{H}\Big]Tc_1^{m-1}\sum_{i=1}^{m-1}\frac{G(\tilde{u}^i - \tilde{u}^{i-1})}{c_1^i} \\
&\leq c_1^{m-1}\bar{H} + c_2\Big[[b_0+\sigma_0] + [b_y+\sigma_y]\bar{H}\Big]Tc_1^{m-1}L_1\sum_{i=1}^{m-1}\frac{c_2^{i-1}}{c_1^i} \\
&\leq \bar{H}c_1^{m-1} + \Big[[b_0+\sigma_0] + [b_y+\sigma_y]\bar{H}\Big]TL_1[m-1]c_2^{m-1}.
\end{aligned}
$$

The proof is complete now. ■

A.4 Proof of Theorem 6.1.

Analogous to Theorem 5.1 one can easily prove (i) and (ii). It is obvious that $F(u) = u$.

We now prove the uniqueness. Assume $F(\tilde{u}) = \tilde{u}$ and $\tilde{u}$ has linear growth. To emphasize the dependence of $c_1(\lambda_2, L)$ and $c_2(\lambda_2, L, G)$ on T, we denote them as $c_1(\lambda_2, L, T)$ and $c_2(\lambda_2, L, G, T)$, respectively. Fix $\lambda_2 = 1$. Choose $\delta > 0$ small enough such that

$$c_1(1, L_1, \delta) \leq c_2(1, L_1, G(\tilde{u}), \delta) < 1.$$

Note that F is a local operator and $u(T, x) = g(x) = \tilde{u}(T, x)$. We set $\varphi^1 = u, \varphi^2 = \tilde{u}$ and apply Lemma A.7 over $[T - \delta, T]$, to get $u(T - \delta, x) = \tilde{u}(T - \delta, x)$. Applying Lemma A.7 again, but over $[T - 2\delta, T - \delta]$ with terminal condition $u(T - \delta, x)$, yields $u(T - 2\delta, x) = \tilde{u}(T - 2\delta, x)$. Repeating the arguments backwardly we obtain $\tilde{u} = u$.

Finally, we show that u is a viscosity solution of (1.3). From Corollary 6.3 we know that u is 1/2-Hölder continuous in time and Lipschitz continuous in space. Hence, under Assumption 2, by Pardoux and Peng (1992) we know that $\Phi = F(u)$ is a viscosity solution of

$$\begin{cases} \Phi_t + \frac{1}{2}\sigma^2(t, x, u)\Phi_{xx} + b(t, x, u)\Phi_x + f(t, x, \Phi, \Phi_x\sigma(t, x, u)) = 0; \\ \Phi(T, x) = g(x). \end{cases}$$

Since $F(u) = u$, (iv) is proved. ■

A.5 Proof of Corollary 6.2

We first prove uniqueness. Assume $(X^l, Y^l, Z^l), l = 1, 2$ are two solutions to (1.1). Let u denote the limit function in Theorem 6.1, and $\delta \triangleq \frac{T}{k}$ for some integer $k > 0$ which will be specified later. Note that

$$\begin{cases} X_t^l = X_{(n-1)\delta}^l + \displaystyle\int_{(n-1)\delta}^t b(s, X_s^l, Y_s^l)ds + \int_{(n-1)\delta}^t \sigma(s, X_s^l, Y_s^l)dW_s; \\ Y_t^l = g(X_{n\delta}^l) + \displaystyle\int_t^{n\delta} f(s, X_s^l, Y_s^l, Z_s^l)ds - \int_t^{n\delta} Z_s^l dW_s; \end{cases} \quad t \in [(n-1)\delta, n\delta].$$

Moreover, given $X_{(n-1)\delta}^l$, the above FBSDE has a unique solution by Antonelli (1993). By the Markovian structure of the problem we have $Y_t^l = \tilde{u}(t, X_t^l), t \in$

$[(n-1)\delta, n\delta]$ for some deterministic function $\tilde{u}$. Then obviously $F(\tilde{u}) = \tilde{u}$ on $[(n-1)\delta, n\delta]$. Since, by Antonelli (1993), we have

$$|Y_t^l|^2 \le C[1 + |X_t^l|^2],$$

for some constant $C > 0$ which may depend on δ and the coefficients, $\tilde{u}$ has linear growth. Then by Theorem 6.1, (iii), we derive $\tilde{u} = u$. Therefore, $Y_t^l = u(t, X_t^l)$ for $t \in [(n-1)\delta, n\delta]$. Now we consider

$$\begin{cases} X_t^l = X_{(n-2)\delta}^l + \int_{(n-2)\delta}^t b(s, X_s^l, Y_s^l)ds + \int_{(n-2)\delta}^t \sigma(s, X_s^l, Y_s^l)dW_s; \\ Y_t^l = u((n-1)\delta, X_{(n-1)\delta)}^l) + \int_t^{(n-1)\delta} f(s, X_s^l, Y_s^l, Z_s^l)ds - \int_t^{(n-1)\delta} Z_s^l dW_s; \end{cases}$$

for $t \in [(n-2)\delta, (n-1)\delta]$. Thanks to Theorem 6.1, (ii), we may choose the same δ to ensure the uniqueness of solutions to the above FBSDE. Then by the same arguments we have $Y_t^l = u(t, X_t^l)$ for $t \in [(n-2)\delta, (n-1)\delta]$. Repeating the arguments backwardly we get $Y_t^l = u(t, X_t^l)$ for $t \in [0, T]$. Then on $[0, \delta]$ we have

$$\begin{cases} X_t^l = x + \int_0^t b(s, X_s^l, u(s, X_s^l))ds + \int_0^t \sigma(s, X_s^l, u(s, X_s^l)dW_s; \\ Y_t^l = u(\delta, X_\delta^l) + \int_t^\delta f(s, X_s^l, Y_s^l, Z_s^l)ds - \int_t^\delta Z_s^l dW_s. \end{cases}$$

By the uniqueness of solutions to this decoupled FBSDE, we get $(X_t^1, Y_t^1, Z_t^1) = (X_t^2, Y_t^2, Z_t^2)$ for $t \in [0, \delta]$. In particular, $X_\delta^1 = X_\delta^2$. Now repeating the arguments forwardly we obtain $(X_t^1, Y_t^1, Z_t^1) = (X_t^2, Y_t^2, Z_t^2)$ for $t \in [0, T]$.

We now prove the existence. For the same δ as above, we consider

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s)ds + \int_0^t \sigma(s, X_s, Y_s)dW_s; \\ Y_t = u(\delta, X_\delta) + \int_t^\delta f(s, X_s, Y_s, Z_s)ds - \int_t^\delta Z_s dW_s. \end{cases}$$

By Antonelli (1993) (X, Y, Z) exists on $[0, \delta]$. By the arguments in the above proof of uniqueness, we know that $Y_t = u(t, X_t), t \in [0, \delta]$. Now we can construct forwardly a solution (X, Y, Z) on $[0, T]$. Obviously, (X, Y, Z) satisfies both (1.1) and (6.2) over $[0, T]$. ■

B Proof of Theorem 6.6

Throughout the proof we will apply Corollary 6.3 several times without further notice. By (6.2) we have

$$\begin{aligned} X_{t_{i+1}} &= X_{t_i} + b(t_i, X_{t_i}, u(t_i, X_{t_i}))h + \int_{t_i}^{t_{i+1}} [b(t, X_t, Y_t) - b(t_i, X_{t_i}, Y_{t_i})]dt \\ &\quad + \sigma(t_i, X_{t_i}, u(t_i, X_{t_i}))\Delta W_{i+1} + \int_{t_i}^{t_{i+1}} [\sigma(t, X_t, Y_t) - \sigma(t_i, X_{t_i}, Y_{t_i})]dW_t. \end{aligned}$$

Applying Lemma 3.2 on X and $X^{n,m}$ we get

$$\begin{aligned} & E\{|X_{t_{i+1}} - X_{i+1}^{n,m}|^2\} \\ & \le E\Big\{[1 + A_1 h + (1+\lambda_2)A_2 h L_1]|X_{t_i} - X_i^{n,m}|^2 \\ & \quad +(1+\lambda_2^{-1})A_2 h |u(t_i, X_i^{n,m}) - u_i^{n,m}(X_i^{n,m})|^2 + 2(1+\lambda_1^{-1}) \times \\ & \quad \int_{t_i}^{t_{i+1}} [|b(t, X_t, Y_t) - b(t_i, X_{t_i}, Y_{t_i})|^2 + |\sigma(t, X_t, Y_t) - \sigma(t_i, X_{t_i}, Y_{t_i})|^2]dt\Big\} \\ & \le E\Big\{[1 + A_1 h + (1+\lambda_2)A_2 h L_1]|X_{t_i} - X_i^{n,m}|^2 \\ & \quad +C(1+\lambda_2^{-1})A_2 h(mc_2^m + h) \\ & \quad \times[1 + |X_i^{n,m} - X_{t_i}|^2 + |X_{t_i}|^2] + C(1+\lambda_1^{-1})(1+|x|)h^2\Big\} \\ & \le (1+Ch)E\{|X_{t_i} - X_i^{n,m}|^2\} + Ch[1+|x|^2][mc_2^m + h]. \end{aligned}$$

Since $X_{t_0} - X_0^{n,m} = 0$, we obtain

$$\sup_i E\{|X_{t_i} - X_i^{n,m}|^2\} \le C[1+|x|^2][mc_2^m + h].$$

Moreover,

$$\begin{aligned} E\Big\{\sup_{t_i \le t \le t_{i+1}} |X_t - X_i^{n,m}|^2\Big\} &\le 2E\Big\{\sup_{t_i \le t \le t_{i+1}} |X_t - X_{t_i}|^2 + |X_{t_i} - X_i^{n,m}|^2\Big\} \\ &\le C[1+|x|^2][mc_2^m + h]. \end{aligned}$$

hence, the estimate for X is proved.

Similarly, recall (6.8) and note that

$$\begin{aligned} Y_{t_i} &= Y_{t_{i+1}} + f(t_i, X_{t_i}, Y_{t_{i+1}}, \hat{Z}_{t_i})h - \int_{t_i}^{t_{i+1}} Z_t dW_t \\ &\quad + \int_{t_i}^{t_{i+1}} [f(t, X_t, Y_t, Z_t) - f(t_i, X_{t_i}, Y_{t_{i+1}}, \hat{Z}_{t_i})]dt. \end{aligned}$$

Applying Lemma 3.3 and following the arguments in (6.9), we have

$$
\begin{aligned}
&E\Big\{|Y_{t_i} - Y_i^{n,m}|^2 + (1-A_3)h|\hat{Z}_{t_i} - \hat{Z}_i^{n,m}|^2\Big\} \\
&\le E\Big\{[1+A_4h]|Y_{t_{i+1}} - Y_{i+1}^{n,m}|^2 + A_5h|X_{t_i} - X_i^{n,m}|^2 \\
&\quad +(\lambda_1^{-1} + h + 2\lambda_3^{-1}h)\int_{t_i}^{t_{i+1}} |f(t,X_t,Y_t,Z_t) - f(t_i,X_{t_i},Y_{t_{i+1}},\hat{Z}_{t_i})|^2 dt\Big\} \\
&\le E\Big\{[1+A_4h]|Y_{t_{i+1}} - Y_{i+1}^{n,m}|^2 + Ch\int_{t_i}^{t_{i+1}} |Z_t|^2 dt + C\int_{t_i}^{t_{i+1}} |Z_t - \tilde{Z}_{t_i}|^2 dt\Big\} \\
&\quad +Ch[1+|x|^2][mc_2^m + h]. \qquad \text{(B.6)}
\end{aligned}
$$

Note that

$$
\begin{aligned}
&E\{|Y_{t_n} - Y_n^{n,m}|^2\} = E\{|g(X_{t_n}) - g(X_n^{n,m})|^2\} \\
&\le CE\{|X_{t_n} - X_n^{n,m}|^2\} \le C[1+|x|^2][mc_2^m + h].
\end{aligned}
$$

Choose λ appropriately such that $A_3 < 1$ for small h. Then

$$
\begin{aligned}
&\sup_{0\le i\le n} E\{|Y_{t_i} - Y_i^{n,m}|^2\} \\
&\le CE\Big\{h\int_0^T |Z_t|^2 dt + \sum_{i=0}^{n-1}\int_{t_i}^{t_{i+1}} |Z_t - \tilde{Z}_{t_i}|^2 dt\Big\} + C[1+|x|^2][mc_2^m + h] \\
&\le C[1+|x|^2][mc_2^m + h].
\end{aligned}
$$

The estimate for Y easily follows.

Moreover, (B.6) implies

$$
\begin{aligned}
&(1-A_3)hE\Big\{|\hat{Z}_{t_i} - \hat{Z}_i^{n,m}|^2\Big\}(1+A_4)^i \\
&\le E\Big\{(1+A_4h)^{i+1}|Y_{t_{i+1}} - Y_{i+1}^{n,m}|^2 - (1+A_4)^i|Y_{t_i} - Y_i^{n,m}|^2\Big\} \\
&\quad +CE\Big\{h\int_{t_i}^{t_{i+1}} |Z_t|^2 dt + \int_{t_i}^{t_{i+1}} |Z_t - \tilde{Z}_{t_i}|^2 dt\Big\} + Ch[1+|x|^2][mc_2^m + h].
\end{aligned}
$$

Summing over i, we get (with the same estimates as for Y),

$$
\sum_{i=0}^{n-1} E\Big\{|\hat{Z}_{t_i} - \hat{Z}_i^{n,m}|^2\Big\}h \le Ch[1+|x|^2][mc_2^m + h].
$$

We can finally write

$$
\sum_{i=1}^{n} E\Big\{\int_{t_{i-1}}^{t_i} |Z_t - \hat{Z}_{i-1}^{n,m}|^2 dt\Big\} \le 2\sum_{i=1}^{n} E\Big\{\int_{t_{i-1}}^{t_i} |Z_t - \hat{Z}_{t_{i-1}}|^2 dt + |\hat{Z}_{t_{i-1}} - \hat{Z}_{i-1}^{n,m}|^2 h\Big\}
$$

The first term can be treated along the lines of (6.9), and so the estimate for Z follows. ■

References

Antonelli, A. (1993) Backward-Forward Stochastic Differential Equations. *Ann. Appl. Probab.*, **3**, 777-793.

Bender, C., Denk, R. (2005) A Forward Scheme for Backward SDEs. Preprint.

Bouchard, B., Touzi, N. (2004) Discrete-Time Approximation and Monte-Carlo Simulation of Backward Stochastic Differential Equations. *Stochastic Process. Appl.*, **111**, 175-206.

Carrière, J. F. (1996) Valuation of the Early-Exercise Price for Options using Simulations and Nonparametric Regression. *Insurance: Math. Econom.*, **19**, 19-30.

Cvitanić, J., Zhang, J. (2005) The Steepest Descent Method for FBSDEs. *Electronic Journal of Probability*, **10**, 1468-1495.

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

Delarue, F., Menozzi, S. (2006) A Forward Backward Stochastic Algorithm for Quasi-Linear PDEs. *Ann. Appl. Probab.*, **16**, 140-184.

Douglas, J., Ma, J., Protter, P. (1996) Numerical Methods for Forwad Backward Stochastic Differential Equations. *Ann. Appl. Probab.*, **6**, 940-968.

Gobet, E., Lemor, J.-P., Warin, X. (2005) A Regression-Based Monte-Carlo Method to Solve Backward Stochastic Differential Equations. *Ann. Appl. Probab.*, **15**, 2172-2202.

Longstaff, F. A., Schwartz, R. S. (2001) Valuing American Options by Simulation: A Simple Least-Square Approach. *Review of Financial Studies*, **14**, 113-147.

Ma, J., Protter, P., Yong, J. (1994) Solving Forward-Backward Stochastic Differential Equations Explicitly – a Four Step Scheme. *Prob. Th. Rel. Fields*, **98**, 339-359.

Ma, J., Shen, J., Zhao, Y. (2006) Numerical Method for Forward Backward SDEs: Four Step Scheme Revisited. Preprint.

Milstein, G. N., Tretyakov, M. V. (2006) Numerical Algorithms for Forward-Backward Stochastic Differential Equations. *SIAM J. Sci. Comput.*, **28**, 561-582.

Pardoux, E., Peng, S. (1992) Backward Stochastic Differential Equations and Quasilinear Parabolic Partial Differential Equations. *Lecture Notes in Control and Inform. Sci, Vol. 176*, Springer, New York, 200-217.

Pardoux, E., Tang, S. (1999) Forward-Backward Stochastic Differential Equations and Quasilinear Parabolic PDEs. *Probab. Theory Relat. Fields*, **114**, 123-150.

Riviere, O. (2005) *Equations différentielles stochastiques progressives rétrogrades couplées: equations aux dérivées partielles et discrétisation.* PhD Thesis, Paris.

Zhang, J. (2001) *Some Fine Properties of Backward Stochastic Differential Equations.* PhD Thesis, Purdue University.

Zhang, J. (2004) A Numerical Scheme for BSDEs. *Ann. Appl. Prob.*, **14**, 459-488.

Zhang, J. (2006) The Wellposedness of FBSDEs. *Discrete and Continuous Dynamic Systems (B)*, **6**, 927-940.