Electronic Journal of Differential Equations, Vol. 2021 (2021), No. 98, pp. 1–16.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu

NEWTON-KANTOROVITCH METHOD FOR DECOUPLED FORWARD-BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

DAI TAGUCHI, TAKAHIRO TSUCHIYA

Abstract. We formulate a Newton-Kantorovitch method for solving decoupled forward-backward stochastic differential equations involving smooth and degenerate coefficients with uniformly bounded derivatives. We show that it converges globally and its rate of convergence is exponential.

1. Introduction

In this article, we study a method for approximating non-Markovian and decoupled forward-backward stochastic differential equations (FBSDEs) of the following form for arbitrary $T > 0$,

$$\begin{aligned} X(t) &= X(0) + \int_0^t b(s, X(s))\,\mathrm{d}s + \int_0^t \sigma(s, X(s))\,\mathrm{d}W(s), \\ Y(t) &= \varphi(X(T)) + \int_t^T f(s, X(s), Y(s), Z(s))\,\mathrm{d}s - \int_t^T Z(s)\,\mathrm{d}W(s), \end{aligned} \tag{1.1}$$

where the solution triplets $(X, Y, Z) \equiv \big(X(t), Y(t), Z(t)\big)_{t\in[0,T]}$ take values in $\mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times k}$, W is a k-dimensional Wiener process, and b, f, σ, and φ are measurable functions that could in general be random and defined on a probability space.

We propose a scheme for approximating such FBSDEs with random coefficients based on applying the Newton-Kantorovitch theory. Unlike a four-step scheme that relies on the Markov structure (see e.g., Ma, Protter and Yong [13], and Delarue [6]), this method allows us to find a non-Markovian approximation. In addition, as our approach is a type of contraction mapping [3, 20], the approximation works for arbitrary large durations with smooth random coefficients, and without monotonicity conditions; for details, see Hu and Peng [9], Peng and Wu [21], and Yong [25]. We also refer readers to [27, 15] for more details about the uniqueness and the existence of FBSDE solutions.

Most numerical algorithms for FBSDEs are based on time-space discretization schemes [7] for quasi-linear parabolic partial differential equations for coupled FBSDEs with monotonicity condition [5]. In contrast, to the best of our knowledge,

2010 *Mathematics Subject Classification.* 49M15, 65C30, 41A25, 60H35, 41A25, 60H10.
Key words and phrases. Stochastic differential equation; Newton-type methods; rate of convergence.

Submitted January 19, 2021. Published December 20, 2021.

very little has been published on state space discretization based approaches. Vidossich proved that the Chaplygin and Newton methods are equivalent for ordinary differential equations [22], and Kawabata and Yamada studied a state space discretization based on Newton-Kantorovitch approach for stochastic differential equations (SDEs) [12]. Ouknine [17] showed that the coefficients can be relaxed by a linear growth condition. Amano obtained a probabilistic second-order error estimate [2]. Wrzosek [23] extended to stochastic functional differential equations.

The aim of this paper is to formulate Newton-Kantorovitch approximation of the decoupled multi-dimensional FBSDEs with random coefficients and to show that it converges globally (1.2) as follows.

Theorem 1.1. *Suppose that b, σ, f, and φ are all C^1, their derivatives are uniformly bounded (s, ω)-a.e., and b, σ, and f are square-integrable with respect to their time variables. Then, there exists a positive constant $C > 0$ such that*

$$\|(X - X_{n+1}, Y - Y_{n+1}, Z - Z_{n+1})\| \leq C 2^{-n} \quad n \in \mathbb{N} \cup \{0\}. \tag{1.2}$$

As pointed out in [4, 8], solutions (X, Y, Z) of decoupled FBSDEs can be viewed as fixed points of a map. Here, we consider an alternative mapping

$$F_\varphi : \Omega \to \mathcal{M}(\mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2,\ \mathbb{S}^2_d \times \mathbb{L}^2_T),$$

where $\mathcal{M}(A, B)$ stands for a set of maps from A to B. It is inspired by a one-dimensional analog of Kawabata and Yamada [12] and Niwa [26], as follows. For given φ and $u = (x, y, z) \in \mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2$, we define

$$F_\varphi(u)(t) = \begin{pmatrix} x(t) - x(0) - \int_0^t b(s, x(s))\,\mathrm{d}s - \int_0^t \sigma(s, x(s))\,\mathrm{d}W(s) \\ y(t) - \varphi(x(T)) - \int_t^T f(s, u(s))\,\mathrm{d}s + \int_t^T z(s)\,\mathrm{d}W(s) \end{pmatrix}^\top. \tag{1.3}$$

Then, for any $(X_0, Y_0, Z_0) \in \mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2$ with $X_0(0) = X(0)$ and assuming that the driver f is smooth, the Newton-Kantorovitch approximation process is given by

$$\begin{gathered} F_\varphi(X_n, Y_n, Z_n) + F'_\varphi(X_n, Y_n, Z_n)(X_{n+1} - X_n, Y_{n+1} - Y_n, Z_{n+1} - Z_n) = 0, \\ X_{n+1}(0) = X(0), \end{gathered} \tag{1.4}$$

for $n \in \mathbb{N} \cup \{0\}$, where F'_φ stands for a Gâteaux derivative,

$$F'_\varphi : \Omega \to \mathcal{L}\big(\mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2, \mathbb{S}^2_d \times \mathbb{L}^2_T\big)$$

where $\mathcal{L}(A, B)$ stands for a set of linear maps from A to B. This sequence is well-posed iff a unique system of linear backward stochastic differential equations (BSDEs) formally by Theorem 3.6. Note that, if the given duration T is arbitrarily, then, even for linear FBSDEs with constant coefficients, the necessary and sufficient conditions become more complicated when the diffusion coefficients also depend on z, as was pointed out by Ma, Wu, Zhang and Zhang [14]. In this paper, we focus on decoupled FBSDEs and obtain the convergence result, Theorem 4.5.

The rest of this paper is structured as follows. Section 2 introduces the notations and assumptions used. Section 3 is devoted to formulating the Newton-Kantorovitch approximation process. Finally, Section 4 proves the main theorem for the decoupled FBSDEs.

2. Preliminaries

For $x, y \in \mathbb{R}^d$, $|x|$ denotes the Euclidian norm and $\langle x, y\rangle$ denotes the inner product. Matrixes size of $m \times k$ will be represented as an element $y \in \mathbb{R}^{m\times k}$, whose Euclidean norms are also given by $|y| = \sqrt{yy^\top}$ and for which $\langle y, z\rangle = \operatorname{trace}(yz^\top)$ where $A^\top$ is the transpose matrix of A. In this paper, we only consider the derivatives with respect to the space variable.

We also use the following notation, based on that in [19]. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $(\mathcal{F}_t)_{t\geq 0}$ be a Brownian filtration. $\mathcal{P} = \mathcal{P}((\mathcal{F}_t)_{t\geq 0})$ is the σ-algebra of the progressively measurable subset $A \subset \Omega\times[0,\infty)$ such that, for all $t \geq 0$, $A\cap([0,t]\times\Omega) \in \mathcal{B}[0,t]\otimes\mathcal{F}_t$. Let $T > 0$ be a fixed, and final, deterministic time. For $m, k \in \mathbb{N}$, we define

$$
\begin{aligned}
&\mathbb{L}^2 = \{\xi : \Omega \to \mathbb{R}^m \text{ is } \mathcal{F}_T\text{-measurable and } \|\xi\|_{\mathbb{L}^2} = \{\mathbb{E}[|\xi|^2]\}^{1/2} < \infty\},\\
&\mathbb{L}^2_T = \{Y : \Omega \times [0,T] \to \mathbb{R}^m \text{ continuous} : Y(t) \in m\mathcal{F}_T, \forall t \in [0,T],\ \|Y\|_{\mathbb{L}^2_T} < \infty\},\\
&\mathbb{S}^2_m = \{Y : \Omega \times [0,T] \to \mathbb{R}^m \text{ continuous, adapted} : \|Y\|_{\mathbb{S}^2_m} < \infty\},\\
&\mathbb{H}^2 = \{Z : \Omega \times [0,T] \to \mathbb{R}^{m\times k} \text{ adapted} : \|Z\|_{\mathbb{H}^2} < \infty\},
\end{aligned}
$$

where the norms $\|\cdot\|_{\mathbb{L}^2_T}$, $\|\cdot\|_{\mathbb{S}^2_m}$, and $\|\cdot\|_{\mathbb{H}^2}$ are defined by

$$
\|Y\|_{\mathbb{L}^2_T} = \|Y\|_{\mathbb{S}^2_m} = \{\mathbb{E}[\sup_{0\leq s\leq T}|Y(s)|^2]\}^{1/2}, \quad \|Z\|_{\mathbb{H}^2} = \{\mathbb{E}[\int_0^T |Z(s)|^2\,\mathrm{d}s]\}^{1/2}.
$$

For the sake of simplicity, we also write the operator norm of the operator A as $\|A\|$. The Banach spaces $\mathbb{S}^2_m \times \mathbb{H}^2$ and $\mathbb{S}^2_d$ are defined by

$$
\|(Y,Z)\|^2 = \|Y\|^2_{\mathbb{S}^2_m} + \|Z\|^2_{\mathbb{H}^2}, \quad \|X\|^2 = \|X\|^2_{\mathbb{S}^2_d}.
$$

For $\alpha \in \mathbb{R}$, we introduce the weighted norm

$$
\|(Y,Z)\|^2_\alpha = \mathbb{E}[\sup_{0\leq s\leq T} e^{\alpha s}|Y(s)|^2] + \mathbb{E}[\int_0^T e^{\alpha s}|Z(s)|^2\,\mathrm{d}s].
$$

For $p, q \in \mathbb{N}$, $C^k(\mathbb{R}^p, \mathbb{R}^q)$ is the set of functions of class C^k from $\mathbb{R}^p$ to $\mathbb{R}^q$, and $C^k_b(\mathbb{R}^p, \mathbb{R}^q)$ is the subset of these functions whose partial derivatives of order at most of the k values are bounded. When the domain and range dimensions are clear based on context and when there is no risk of confusion, we often eliminate the spaces to simplify the notation.

For a smooth g such that $g(\cdot,\cdot,\cdot) \in C^1(\mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^r, \mathbb{R}^q)$, it is convenient to obtain a concrete representation of $g' : \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^r \to \mathbb{R}^q$. The Fréchet derivative at $u \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^q$ is the matrix representation of $g'(u)$ [16, page 60] and it can be obtained using the Jacobian matrix:

$$
g(u + \Delta u) - g(u) = g'(u)\Delta u + \mathrm{R}_g(u)\Delta u, \quad \Delta u \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^r \tag{2.1}
$$

where the Lagrange remainder is

$$
\mathrm{R}_g(u)\Delta u = \Big(\int_0^1 \{g'(u + \theta\Delta u) - g'(u)\}\mathrm{d}\theta\Big)\Delta u.
$$

Notice that, for all $u = (x,y,z)^\top, \Delta u = (\Delta x, \Delta y, \Delta z)^\top \in \mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^r$,

$$
g'(u)\Delta u = g'_x(u)\Delta x + g'_y(u)\Delta y + g'_z(u)\Delta z, \tag{2.2}
$$

where we obtain $g'_x(u) : \mathbb{R}^p \to \mathbb{R}^p \times \mathbb{R}^p$ such that

$$g'_x(u)\Delta x = (\frac{\partial}{\partial x_j} g^{(i)}(x,y,z)\Delta x^{(i)})_{i\le p,\ j\le p},$$

where $\Delta x^{(i)}$ stands for the ith component for $i \le p$. We define g'_y and g'_z, similarly. Finally, for all $s \in [0,T]$, we define

$$\|g'\|_\infty = \sup_{(s,x,y,z)\in[0,T]\times\mathbb{R}^p\times\mathbb{R}^q\times\mathbb{R}^r} |g'(s,x,y,z)|,$$

for $g(s,\cdot,\cdot,\cdot) \in C_b^1(\mathbb{R}^p \times \mathbb{R}^q \times \mathbb{R}^r, \mathbb{R}^q)$.

3. Newton-Kantorovitch scheme for FBSDEs

In this section, we formulate the Newton-Kantorovitch scheme of the following system of FBSDEs:

$$\begin{aligned} X(t) &= X(0) + \int_0^t b(s,X(s))\,\mathrm{d}s + \int_0^t \sigma(s,X(s))\,\mathrm{d}W(s), \\ Y(t) &= \varphi(X(T)) + \int_t^T f(s,X(s),Y(s),Z(s))\,\mathrm{d}s - \int_t^T Z(s)\,\mathrm{d}W(s), \end{aligned} \tag{3.1}$$

where the solutions $(X,Y,Z) = (X(t),Y(t),Z(t))_{t\in[0,T]}$ take values in $\mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m\times k}$, W is the k-dimensional Wiener process, and the (progressively) measurable functions b, f, σ and φ are defined on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge 0}, \mathbb{P})$.

In this paper, we also assume the following.

- $X(0) : \Omega \to \mathbb{R}^d$ is an $\mathcal{F}_0$-measurable and square-integrable random vector.
- $b : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^d$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable.
- $\sigma : [0,T] \times \Omega \times \mathbb{R}^d \to \mathbb{R}^{d\times k}$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable.
- There exists a constant $C > 0$ such that, for any $x \in \mathbb{R}^d$,
$$\begin{aligned} |b(s,x)| &\le |b(s,0)| + C|x|, \quad (s,\omega)\text{-a.e.}, \\ |\sigma(s,x)| &\le |\sigma(s,0)| + C|x|, \quad (s,\omega)\text{-a.e.} \end{aligned}$$
- $\varphi : \Omega \times \mathbb{R}^d \to \mathbb{R}^m$ is $\mathcal{F}_T \otimes \mathcal{B}(\mathbb{R}^d)$-measurable.
- $f : [0,T]\times\Omega\times\mathbb{R}^d\times\mathbb{R}^m\times\mathbb{R}^{m\times k} \to \mathbb{R}^m$ is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^m)\otimes\mathcal{B}(\mathbb{R}^{m\times k})$-measurable.
- There exists a constant $C > 0$ such that for any $(x,y,z) \in \mathbb{R}^d\times\mathbb{R}^m\times\mathbb{R}^{m\times k}$,
$$\begin{aligned} |f(s,x,y,z)| &\le |f(s,0,0,0)| + C(|x|+|y|+|z|), \quad (s,\omega)\text{-a.e.} \\ |\varphi(x)| &\le |\varphi(0)| + C|x|, \quad \omega\text{-a.e.} \end{aligned}$$

In addition, in the BSDE,

$$Y(t) = \xi + \int_t^T f(s,Y(s),Z(s))\,\mathrm{d}s - \int_t^T Z(s)\,\mathrm{d}W(s),$$

we replace the above assumption on φ with the condition

- $\xi : \Omega \to \mathbb{R}^m$ is an $\mathcal{F}_T$-measurable and square integrable random vector; $\xi \in \mathbb{L}^2$.

We also introduce the following assumptions.

Assumption 3.1. $b(\cdot,0), \sigma(\cdot,0), f(\cdot,0,0,0) \in \mathbb{H}^2$, i.e.,

$$\mathbb{E}\Big[\int_0^T \{|b(s,0)|^2 + |\sigma(s,0)| + |f(s,0,0,0)|^2\}\,\mathrm{d}s\Big] < \infty$$

and $\varphi(0) \in \mathbb{L}^2$.

Assumption 3.2. $b(s,\cdot) \in C_b^1(\mathbb{R}^d,\mathbb{R}^d)$, $\sigma(s,\cdot) \in C_b^1(\mathbb{R}^d,\mathbb{R}^{d\times k})$, $f(s,\cdot,\cdot,\cdot) \in C_b^1(\mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^m \times k, \mathbb{R}^m), (s,\omega)$-a.e. and $\varphi \in C_b^1(\mathbb{R}^d,\mathbb{R}^m)$, ω-a.e.

For a given φ, we consider the operator F_φ (defined by (1.3)),

$$F_\varphi : \Omega \to \mathcal{M}(\mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2, \mathbb{S}_d^2 \times \mathbb{L}_T^2)$$

where $\mathcal{M}(A,B)$ stands for a collection of maps from A to B, namely,

$$F_\varphi(u)(t) = \begin{pmatrix} x(t) - x(0) - \int_0^t b(s,x(s))\,\mathrm{d}s - \int_0^t \sigma(s,x(s))\,\mathrm{d}W(s) \\ y(t) - \varphi(x(T)) - \int_t^T f(s,u(s))\,\mathrm{d}s + \int_t^T z(s)\,\mathrm{d}W(s) \end{pmatrix}^\top,$$

for $u = (x,y,z) \in \mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2$. As an immediate consequence of the result given in [12, Lemma 3.1], we obtain the following corresponding result for FBSDEs.

Lemma 3.3. *If Assumption 3.1 holds, then the operator F_φ defined by (1.3) maps the space $\mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2$ into $\mathbb{S}_d^2 \times \mathbb{L}_T^2$ and $t \mapsto F_\varphi(u)(t)$ is a continuous modification for $u \in \mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2$.*

Proof. For any $u = (x,y,z) \in \mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2$, it follows, from $\|x\|_{\mathbb{S}_d^2} < \infty$ and $\|z\|_{\mathbb{H}^2}^2 < \infty$, that the Itô integrals $t \mapsto \int_0^t \sigma(s,x(s))\,\mathrm{d}W(s)$ and $t \mapsto \int_t^T z(s)\,\mathrm{d}W(s)$, respectively, are continuous modifications. By the Jensen inequality, we obtain

$$\mathbb{E}\big[\int_0^t |b(s,x(s))|^2\,\mathrm{d}s\big] \le 2\mathbb{E}\big[\int_0^t |b(s,0)|^2\,\mathrm{d}s\big] + 2C^2T\|x\|_{\mathbb{S}_d^2}^2 < \infty$$

and

$$\begin{aligned}
&\mathbb{E}\big[\int_t^T |f(s,x(s),y(s),z(s))|^2\,\mathrm{d}s\big] \\
&\le 4\mathbb{E}\big[\int_t^T |f(s,0,0,0)|^2\,\mathrm{d}s\big] + 4C^2\big(T\{\|x\|_{\mathbb{S}_d^2}^2 + \|y\|_{\mathbb{S}_m^2}^2\} + \|z\|_{\mathbb{H}^2}^2\big) < \infty.
\end{aligned}$$

By Doob's inequality and Itô's isometry property, there exists a $c > 0$ such that

$$\begin{aligned}
&\mathbb{E}[\sup_{0\le t\le T} |\int_0^t \sigma(s,x(s))\,\mathrm{d}W(s)|^2] \\
&\le c\mathbb{E}\big[|\int_0^T \sigma(s,x(s))\,\mathrm{d}W(s)|^2\big] \\
&\le 2c\mathbb{E}\big[\int_0^T |\sigma(s,0)|^2\,\mathrm{d}s\big] + 2cC^2T\|x\|_{\mathbb{S}_d}^2 < \infty
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\Big[\sup_{0\le t\le T} \Big|\int_t^T z(s)\,\mathrm{d}W(s)\Big|^2\Big] &\le c\mathbb{E}\Big[\Big|\int_0^T z(s)\,\mathrm{d}W(s)\Big|^2\Big] \\
&= c\mathbb{E}\big[\int_0^T |z(s)|^2\,\mathrm{d}s\big] < \infty.
\end{aligned}$$

Using the Jensen inequality, we further obtain that

$$\frac{1}{9}\mathbb{E}\Big[\sup_{0\le t\le T} |F_\varphi(u)(t)|^2\Big]$$

$$
\begin{aligned}
&\le \|x\|_{\mathbb{S}_d^2}^2 + \mathbb{E}[|x(0)|^2] + \mathbb{E}\Big[\int_0^t |b(s,x(s))|^2\,\mathrm{d}s\Big] \\
&\quad + \mathbb{E}\Big[\sup_{0\le t\le T}\Big|\int_0^t \sigma(s,x(s))\,\mathrm{d}W(s)\Big|^2\Big] + \|y\|_{\mathbb{S}_m^2}^2 + \mathbb{E}|[\varphi(0)|^2 + C|x(T)|^2] \\
&\quad + \mathbb{E}\Big[\Big|\int_t^T f(s,x(s),y(s),z(s))\,\mathrm{d}s\Big|^2\Big] + \mathbb{E}\Big[\sup_{0\le t\le T}\Big|\int_t^T z(s)\,\mathrm{d}W(s)\Big|^2\Big],
\end{aligned}
$$

which is bounded by the above estimates. This completes the proof. □

Denote by $\mathcal{L}(A,B)$ a collection of linear maps from A to B. The following lemma shows that the operator $F_\varphi : \Omega \to \mathcal{L}(\mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2, \mathbb{S}_d^2 \times \mathbb{L}_T^2)$ has a Gâteaux derivative.

Lemma 3.4. *If Assumptions 3.1 and 3.2 hold, then for all $u = (x,y,z) \in \mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2$, the Gâteaux derivative $F'_\varphi(u) : \mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2 \to \mathbb{S}_d^2 \times \mathbb{L}_T^2$ of F_φ at $u \in \mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2$ in the direction $\overline{u} = (\overline{x},\overline{y},\overline{z}) \in \mathbb{S}_m^2 \times \mathbb{H}^2$ exists and is given for any $t \in [0,T]$, by*

$$
\begin{aligned}
&F'_\varphi(u)\overline{u}(t) \\
&= \begin{pmatrix} \overline{x}(t) - \overline{x}(0) - \int_0^t b'_x(s,x(s))\overline{x}(s)\,\mathrm{d}s - \int_0^t \sigma'_x(s,x(s))\overline{x}(s)\,\mathrm{d}W(s) \\ \overline{y}(t) - \varphi'(x(T))\overline{x}(T) - \int_t^T f'(s,u(s))\overline{u}(s)\,\mathrm{d}s + \int_t^T \overline{z}(s)\,\mathrm{d}W(s) \end{pmatrix}^\top .
\end{aligned} \tag{3.2}
$$

Proof. We denote the right-hand side of (3.2) by $A(u)\overline{u}(t)$ for $\overline{u} = (\overline{x},\overline{y},\overline{z}) \in \mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2$. We note that because

$$
\begin{aligned}
&\mathbb{E}\Big[\int_t^T \big\{|b'_x(s,x(s))\overline{x}(s)|^2 + |\sigma'_x(s,x(s))\overline{x}(s)|^2 + |f'(s,u(s))\overline{u}(s)|^2\big\}\,\mathrm{d}s\Big] \\
&\le (1\vee T)\{\|b'\|_\infty^2 + \|\sigma'\|_\infty^2 + \|f'\|_\infty^2\}\|\overline{u}\|^2,
\end{aligned}
$$

we obtain that $A(u)\overline{u} \in \mathbb{L}_T^2$ by the same argument of Lemma 3.3. It follows from (2.1) that, for all $s \in [0,T]$,

$$
\begin{aligned}
b(s,x(s)+\delta\overline{x}(s)) - b(s,x(s)) &= \delta b'_x(s,x(s))\overline{x}(s) + \mathrm{R}_b(x(s))(\delta\overline{x})(s), \\
\sigma(s,x(s)+\delta\overline{x}(s)) - \sigma(s,x(s)) &= \delta \sigma'_x(s,x(s))\overline{x}(s) + \mathrm{R}_\sigma(x(s))(\delta\overline{x})(s), \\
f(s,u(s)+\delta\overline{u}(s)) - f(s,u(s)) &= \delta f'(s,u(s))\overline{u}(s) + \mathrm{R}_f(u(s))(\delta\overline{u})(s), \\
\varphi(x(T)+\delta\overline{x}(T)) - \varphi(x(T)) &= \delta\varphi'(x(T))\overline{x}(T) + \mathrm{R}_\varphi(x(T))(\delta\overline{x})(T).
\end{aligned}
$$

Hence, for all $\delta > 0$ and $t \in [0,T]$, we have

$$
\begin{aligned}
&\frac{F_\varphi(u+\delta\overline{u})(t) - F_\varphi(u)(t)}{\delta} \\
&= A(u)\overline{u}(t) + \frac{1}{\delta}\begin{pmatrix} -\int_0^t \mathrm{R}_b(x(s))(\delta\overline{x})(s)\,\mathrm{d}s - \int_0^t \mathrm{R}_\sigma(x(s))(\delta\overline{x})(s)\,\mathrm{d}W(s) \\ -\mathrm{R}_\varphi(x(T))(\delta\overline{x})(T) - \int_t^T \mathrm{R}_f(u(s))(\delta\overline{u})(s)\,\mathrm{d}s. \end{pmatrix}^\top .
\end{aligned}
$$

Since $b'_x(s,\cdot)$, $\sigma'_x(s,\cdot)$, $f'(s,\cdot)$ and φ' are bounded and continuous (s,ω)-a.e., by using the dominated convergence theorem, we obtain

$$
\lim_{\delta\to 0}\Big\|\frac{F_\varphi(u+\delta\overline{u}) - F_\varphi(u)}{\delta} - A(u)\overline{u}\Big\|_{\mathbb{S}_d^2\times\mathbb{L}_T^2} = 0,
$$

thus completing the proof. □

The following lemma is a key for defining the Newton-Kantorovitch approximation process.

Lemma 3.5. *Suppose that Assumptions 3.1 and 3.2 hold and let* $u = (x, y, z) \in \mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2$. *Then, there exists a unique* $\widetilde{u} = (\widetilde{x}, \widetilde{y}, \widetilde{z}) \in \mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2$ *such that* $F_\varphi(u) + F'_\varphi(u)(\widetilde{u} - u) = 0$ *with initial condition* $\widetilde{x}(0) = x(0)$.

Proof. Let $u = (x, y, z)$ be in $\mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2$. We show that there exists a unique $\widetilde{u} = (\widetilde{x}, \widetilde{y}, \widetilde{z}) \in \mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2$ such that $F_\varphi(u) + F'_\varphi(u)(\widetilde{u} - u) = 0$ with $\widetilde{x}(0) = x(0)$, i.e.,

$$
\begin{aligned}
\widetilde{x}(t) &= x(0) + \int_0^t \{b(s, x(s)) + b'_x(s, x(s))(\widetilde{x}(s) - x(s))\}\,\mathrm{d}s \\
&\quad + \int_0^t \{\sigma(s, x(s)) + \sigma'_x(s, x(s))(\widetilde{x}(s) - x(s))\}\,\mathrm{d}W(s), \\
\widetilde{y}(t) &= \varphi(x(T)) + \varphi'(x(T))(\widetilde{x}(T) - x(T)) \\
&\quad + \int_t^T \{f(s, u(s)) + f'(s, u(s))(\widetilde{u}(s) - u(s))\}\,\mathrm{d}s - \int_t^T \widetilde{z}(s)\,\mathrm{d}W(s),
\end{aligned}
\tag{3.3}
$$

Because the above equation is a linear decoupled FBSDE with uniformly bounded coefficients, there exists a unique $\widetilde{u} \in \mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2$ as required. □

From Lemma 3.5, we can conclude that for any initial condition $(X_0, Y_0, Z_0) \in \mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2$, we can define the Newton-Kantorovitch approximation process $(X_{n+1}, Y_{n+1}, Z_{n+1}) \in \mathbb{S}^2_d \times \mathbb{S}^2_m \times \mathbb{H}^2$ as solving the equation

$$
\begin{gathered}
F_\varphi(X_n, Y_n, Z_n) + F'_\varphi(X_n, Y_n, Z_n)(X_{n+1} - X_n, Y_{n+1} - Y_n, Z_{n+1} - Z_n) = 0, \\
X_{n+1}(0) = X_0(0),
\end{gathered}
\tag{3.4}
$$

for $n \in \mathbb{N} \cup \{0\}$, which is equivalent to

$$
(X_{n+1}, Y_{n+1}, Z_{n+1}) = (X_n, Y_n, Z_n) - F'_\varphi(X_n, Y_n, Z_n)^{-1} F_\varphi(X_n, Y_n, Z_n).
$$

The following theorem shows that (3.4) has a unique solution that satisfies a linear decoupled FBSDE with uniformly bounded derivatives of coefficients.

Theorem 3.6. *If Assumptions 3.1 and 3.2 hold, then* $(X_{n+1}, Y_{n+1}, Z_{n+1})$ *satisfies the following linear decoupled FBSDE for* $0 \le t \le T$*:*

$$
\begin{aligned}
X_{n+1}(t) &= X_{n+1}(0) + \int_0^t b_n(s, X_{n+1}(s))\,\mathrm{d}s + \int_0^t \sigma_n(s, X_{n+1}(s))\,\mathrm{d}W(s), \\
Y_{n+1}(t) &= \varphi_n(X_{n+1}(T)) + \int_t^T f_n(s, X_{n+1}(s), Y_{n+1}(s), Z_{n+1}(s))\,\mathrm{d}s \\
&\quad - \int_t^T Z_{n+1}(s)\,\mathrm{d}W(s),
\end{aligned}
\tag{3.5}
$$

where we define for $0 \le s \le T$ *and* $(x, y, z) \in \mathbb{R}^d \times \mathbb{R}^m \times \mathbb{R}^{m \times k}$,

$$
\begin{aligned}
b_n(s, x) &= b(s, X_n(s)) + b'_x(s, X_n(s))(x - X_n(s)), \\
\sigma_n(s, x) &= \sigma(s, X_n(s)) + \sigma'_x(s, X_n(s))(x - X_n(s)), \\
f_n(s, x, y, z) &= f(s, X_n(s), Y_n(s), Z_n(s)), \\
&\quad + f'(X_n(s), Y_n(s), Z_n(s))(x - X_n(s), y - Y_n(s), z - Z_n(s)), \\
\varphi_n(x) &= \varphi(X_n(T)) + \varphi'_x(X_n(T))(x - X_n(T)).
\end{aligned}
$$

In particular, if $X_0(0) = X(0)$, *then* $X_{n+1}(0) = X(0)$ *for all* $n \in \mathbb{N} \cup \{0\}$.

Proof. The existence and uniqueness of $(X_{n+1}(t), Y_{n+1}(t), Z_{n+1}(t))_{t\in[0,T]}$ follows from Lemma 3.5, where, for all $n \in \mathbb{N}\cup\{0\}$. The equation (3.3) can be re-expressed as the desired FBSDE for $0 \le t \le T$. Finally, by the initial condition $X_{n+1}(0) = X_n(0)$, if $X_0(0) = X(0)$, then $X_{n+1}(0) = X(0)$, for all $n \in \mathbb{N} \cup \{0\}$. □

4. Convergence of the Newton-Kantorovitch approximation process for FBSDEs

In this section, we investigate the convergence of the Newton-Kantorovitch approximation processes defined by (3.4).

4.1. **Convergence of the Newton-Kantorovitch approximation process for SDEs.** In this subsection, we consider the Newton-Kantorovitch scheme for the forward process X and extend earlier studies of Newton-Kantorovitch methods for SDEs,

$$X(t) = X(0) + \int_0^t b(s, X(s))\,\mathrm{d}s + \int_0^t \sigma(s, X(s))\,\mathrm{d}W(s). \tag{4.1}$$

Theorem 4.1. *Let X be a solution of* (4.1). *If Assumptions 3.1 and 3.2 hold, then, for any $X_0 \in \mathbb{S}_d^2$ where $X_0(0) = X(0)$, we obtain that, for all $n \in \mathbb{N}\cup\{0\}$ and $\epsilon \in (0,1)$,*

$$\|X - X_{n+1}\|^2 \le \frac{C_0^{n+1}}{(n+1)!}\|X - X_0\|^2 \le \epsilon^{n+1} e^{C_0 T/\epsilon}\|X - X_0\|^2, \tag{4.2}$$

where the constant C_0 is given by

$$C_0 = 8c_{b,\sigma} T \exp(4c_{b,\sigma}T), \quad c_{b,\sigma} = \|b'\|_\infty + 18\|\sigma'\|_\infty + \|\sigma'\|_\infty^2.$$

Remark 4.2. An estimate was initially given in [12] for one-dimensional SDEs with uniformly bounded coefficients and an alternative estimation was proposed by [1].

Proof of Theorem 4.1. The proof is based on [1]. By a fundamental result in [10], X exists and is unique. For $0 \le s \le T$ and $n \in \mathbb{N}$, define

$$\overline{X}_n(s) = X(s) - X_n(s),$$

we obtain for all $s \in [0,T]$ and $n \in \mathbb{N}$, $\mathbb{P}$-a.s. By the mean value theorem, for $s \in [0,T]$ and $n \in \mathbb{N}$, we obtain

$$\begin{aligned}
&b(s, X(s)) - b_n(s, X_{n+1}(s)) \\
&= \{b(s, X(s)) - b(s, X_n(s)) - b'_x(s, X_n(s))\overline{X}_n(s)\} + b'_x(s, X_n(s))\overline{X}_{n+1}(s) \\
&= \mathrm{R}_b(X_n(s))\overline{X}_n(s) + b'_x(s, X_n(s))\overline{X}_{n+1}(s)
\end{aligned}$$

and

$$\begin{aligned}
&\sigma(s, X(s)) - \sigma_n(s, X_{n+1}(s)) \\
&= \{\sigma(s, X(s)) - \sigma(s, X_n(s)) - \sigma'_x(s, X_n(s))\overline{X}_n(s)\} + \sigma'_x(s, X_n(s))\overline{X}_{n+1}(s) \\
&= \mathrm{R}_\sigma(X_n(s))\overline{X}_n(s) + \sigma'_x(s, X_n(s))\overline{X}_{n+1}(s).
\end{aligned}$$

Recall (2.1) and we have, for $s \in [0,T]$, $n \in \mathbb{N}$ and $h \in \mathbb{S}_d^2$,

$$\mathrm{R}_b(X_n(s))h(s) = \Big\{\int_0^1 b'_x(s, (X_n(s)) + \theta h(s))\mathrm{d}\theta - b'_x(s, X_n(s))\Big\}h(s),$$

$$\mathrm{R}_\sigma(X_n(s))h(s) = \Big\{ \int_0^1 \sigma_x'(s,(X_n(s))+\theta h(s))\mathrm{d}\theta - \sigma_x'(s,X_n(s))\Big\}h(s).$$

This allows us to obtain

$$\begin{aligned}\overline{X}_{n+1}(t) &= \int_0^t \big\{b_x'(s,X_n(s))\overline{X}_{n+1}(s) + \mathrm{R}_b(X_n(s))\overline{X}_n(s)\big\}\,\mathrm{d}s\\ &\quad + \int_0^t \{\sigma_x'(s,X_n(s))\overline{X}_{n+1}(s) + \mathrm{R}_\sigma(X_n(s))\overline{X}_n(s)\}\,\mathrm{d}W(s),\end{aligned}$$

which, by applying Itô's formula yields

$$\begin{aligned}|\overline{X}_{n+1}(t)|^2 &= 2\int_0^t \langle \overline{X}_{n+1}(s), b_x'(s,X_n(s))\overline{X}_{n+1}(s) + \mathrm{R}_b(X_n(s))\overline{X}_n(s)\rangle\,\mathrm{d}s\\ &\quad + 2\int_0^t \langle \overline{X}_{n+1}(s), \sigma_x'(s,X_n(s))\overline{X}_{n+1}(s) + \mathrm{R}_\sigma(X_n(s))\overline{X}_n(s)\,\mathrm{d}W(s)\rangle\\ &\quad + \int_0^t |\sigma_x'(s,X_n(s))\overline{X}_{n+1}(s) + \mathrm{R}_\sigma(X_n(s))\overline{X}_n(s)|^2\,\mathrm{d}s.\end{aligned}$$

Note that we obtain

$$\begin{aligned}&2\langle \overline{X}_{n+1}(s), b_x'(s,X_n(s))\overline{X}_{n+1}(s) + \mathrm{R}_b(X_n(s))\overline{X}_n(s)\rangle\\ &\le 2\|b'\|_\infty|\overline{X}_{n+1}(t)|^2 + 4\|b'\|_\infty|\overline{X}_n(t)|^2\end{aligned}$$

and

$$\begin{aligned}&|\sigma_x'(s,X_n(s))\overline{X}_{n+1}(s) + \mathrm{R}_\sigma(X_n(s))\overline{X}_n(s)|^2\\ &\le 2\|\sigma'\|_\infty^2|\overline{X}_{n+1}(t)|^2 + 4\|\sigma'\|_\infty^2|\overline{X}_n(t)|^2.\end{aligned}$$

The Burkholder-Davis-Gundy inequality implies that there exists a c_0 such that

$$\begin{aligned}&\mathbb{E}\Big[\big|2\int_0^t \langle \overline{X}_{n+1}(t), \sigma_x'(s,X_n(s))\overline{X}_{n+1}(s)\,\mathrm{d}W(s)\rangle\big|\Big]\\ &\le \mathbb{E}\Big[2\Big(\frac{1}{4}\sup_{0\le s\le t}|\overline{X}_{n+1}(s)|^2\Big)^{1/2}\Big(4c_0^2\|\sigma'\|_\infty\int_0^t |X_{n+1}(s)|^2\,\mathrm{d}s\Big)^{1/2}\Big]\\ &\le \frac{1}{4}\mathbb{E}\big[\sup_{0\le s\le t}|\overline{X}_{n+1}(s)|^2\big] + 4c_0^2\|\sigma'\|_\infty\mathbb{E}\big[\int_0^t |X_{n+1}(s)|^2\,\mathrm{d}s\big],\end{aligned}$$

where we obtain the last inequality by applying the inequality $ab \le (a^2/2)+(b^2/2)$ for all $a,b\in\mathbb{R}$. Similarly, we have

$$\begin{aligned}&\mathbb{E}\Big[\big|2\int_0^t \langle \overline{X}_{n+1}(t), \mathrm{R}_\sigma(X_n(s))\overline{X}_n(s)\,\mathrm{d}W(s)\rangle\big|\Big]\\ &\le \mathbb{E}\Big[2\Big(\frac{1}{4}\sup_{0\le s\le t}|\overline{X}_{n+1}(s)|^2\Big)^{1/2}\Big(8c_0^2\|\sigma'\|_\infty\int_0^t |X_n(s)|^2\,\mathrm{d}s\Big)^{1/2}\Big]\\ &\le \frac{1}{4}\mathbb{E}\big[\sup_{0\le s\le t}|\overline{X}_{n+1}(s)|^2\big] + 8c_0^2\|\sigma'\|_\infty\mathbb{E}\big[\int_0^t |X_n(s)|^2\,\mathrm{d}s\big],\end{aligned}$$

where we note that an explicit upper bounded of c_0 can be obtained by 3; see [11, Theorem 3.28]. By setting $c_{b,\sigma} = \|b'\|_\infty + 2c_0^2\|\sigma'\|_\infty + \|\sigma'\|_\infty^2$, we obtain, for any $t'\in[0,T]$,

$$\mathbb{E}\big[\sup_{0\le t\le t'}|\overline{X}_{n+1}(t)|^2\big]$$

$$\leq 4c_{b,\sigma}\int_0^{t'}\mathbb{E}[\sup_{0\leq t\leq s}|\overline{X}_{n+1}(t)|^2]\,\mathrm{d}s+8c_{b,\sigma}\int_0^{t'}\mathbb{E}[\sup_{0\leq t\leq s}|\overline{X}_n(t)|^2]\,\mathrm{d}s.$$

Gronwall's inequality further implies that

$$\mathbb{E}\big[\sup_{0\leq t\leq t'}|\overline{X}_{n+1}(t)|^2\big]\leq C_0\int_0^{t'}\mathbb{E}\big[\sup_{0\leq t\leq s}|\overline{X}_n(t)|^2\big]\,\mathrm{d}s, \tag{4.3}$$

where $C_0=8c_{b,\sigma}\exp(4c_{b,\sigma}T)$. Iterating (4.3) yields

$$\begin{aligned}\mathbb{E}\big[\sup_{0\leq t\leq T}|\overline{X}_{n+1}(t)|^2\big]&\leq C_0^2\int_0^T\mathrm{d}s_1\int_0^{s_1}\mathrm{d}s_2\mathbb{E}\big[\sup_{0\leq t\leq s_2}|\overline{X}_{n-1}(t)|^2\big]\\&\leq C_0^{n+1}\mathbb{E}\big[\sup_{0\leq t\leq T}|\overline{X}_0(t)|^2\big]\int_0^T\mathrm{d}s_1\int_0^{s_1}\mathrm{d}s_2\cdots\int_0^{s_n}\mathrm{d}s_{n+1}\\&=\frac{(TC_0)^{n+1}}{(n+1)!}\mathbb{E}\big[\sup_{0\leq t\leq T}|\overline{X}_0(t)|^2\big].\end{aligned}$$

In addition, we obtain

$$\epsilon^{-(n+1)}\frac{(TC_0)^{n+1}}{(n+1)!}\leq\sum_{k=0}^{\infty}\frac{(TC_0/\epsilon)^k}{k!}=e^{C_0T/\epsilon},$$

thus completing the proof. □

We can also use the same argument to show that the approximation process is a Cauchy sequence in $\mathbb{S}_d^2$.

4.2. **Toward decoupled forward-backward stochastic differential equations.** The inequality in Theorem 4.3 below indicates that approximating the terminal condition is the key to estimating the error between the solution and the Newton-Kantorovitch approximation process. In this section, we consider the Newton-Kantorovitch approximation with respect to the terminal condition. First, we show that it converges with respect to the weighted norm $\|\cdot\|_\alpha$.

Theorem 4.3. *Let (X,Y,Z) be a solution of the FBSDE* (1.1). *If Assumptions 3.1 and 3.2 hold and $(X_0,Y_0,Z_0)\in\mathbb{S}_d^2\times\mathbb{S}_m^2\times\mathbb{H}^2$ such that $X_0(0)=X(0)$, then, for all $n\in\mathbb{N}\cup\{0\}$, $\epsilon\in(0,1)$ and $T>0$, we obtain*

$$\|(Y-Y_{n+1},Z-Z_{n+1})\|_\alpha^2\leq\epsilon^{n+1}\{C_1\|X-X_0\|_{\mathbb{S}_d^2}^2+\|(Y-Y_0,Z-Z_0)\|_\alpha^2\}, \tag{4.4}$$

where $C_1\equiv C_1(\epsilon)=\epsilon^{-1}\{1+(1+4(2+TC_0)\|\varphi'\|_\infty^2)(1+4c_0^2)\}e^{\alpha T+C_0/\epsilon}$, $c_0=3$ and C_0 is given by Theorem 4.1 and $\alpha=2\|f'\|_\infty+4\|f'\|_\infty^2+12\|f'\|_\infty(1+4c_0^2)(1\vee T)\epsilon^{-1}$.

Proof. By the fundamental result in [18], the solution (X,Y,Z) exists and is unique. For $s\in[0,T]$ and $n\in\mathbb{N}$, we define

$$h_n(s)=(\overline{X}_n(s),\overline{Y}_n(s),\overline{Z}_n(s))=(X(s)-X_n(s),Y(s)-Y_n(s),Z(s)-Z_n(s)),$$

and apply the mean value theorem, yielding

$$\begin{aligned}&f(s,X(s),Y(s),Z(s))-f_n(s,X_{n+1}(s),Y_{n+1}(s),Z_{n+1}(s))\\&=f'(s,X_n(s),Y_n(s),Z_n(s))h_{n+1}(s)+\{f(s,X(s),Y(s),Z(s))\\&\quad-f(s,X_n(s),Y_n(s),Z_n(s))-f'(s,X_n(s),Y_n(s),Z_n(s))h_n(s)\}\\&=f'(s,X_n(s),Y_n(s),Z_n(s))h_{n+1}(s)+\mathrm{R}_f(X_n(s),Y_n(s),Z_n(s))h_n(s).\end{aligned}$$

Recall (2.1) and we have, for $s \in [0,T]$, $n \in \mathbb{N}$ and $h \in \mathbb{S}_d^2 \times \mathbb{S}_m^2 \times \mathbb{H}^2$,

$$\begin{aligned}
&\mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h(s)\\
&= \Big\{ \int_0^1 f'(s, (X_n(s), Y_n(s), Z_n(s)) + \theta h(s))\mathrm{d}\theta - f'(s, X_n(s), Y_n(s), Z_n(s)) \Big\} h(s).
\end{aligned}$$

This allows us to show that $(\overline{Y}_{n+1}, \overline{Z}_{n+1})$ satisfies the following linear BSDE

$$\begin{aligned}
&\overline{Y}_{n+1}(t) - \overline{Y}_{n+1}(T) + \int_t^T \overline{Z}_{n+1}(s)\,\mathrm{d}W(s)\\
&= \int_t^T \{ f'(s, X_n(s), Y_n(s), Z_n(s))h_{n+1}(s) + \mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h_n(s)\}\,\mathrm{d}s.
\end{aligned}$$

Applying Itô's formula to $e^{\alpha t}|\overline{Y}_{n+1}(t)|^2$ for all $\alpha \in \mathbb{R}$, we obtain

$$\begin{aligned}
&e^{\alpha t}|\overline{Y}_{n+1}(t)|^2 - e^{\alpha T}|\overline{Y}_{n+1}(T)|^2 + \int_t^T e^{\alpha s}|\overline{Z}_{n+1}(s)|^2\,\mathrm{d}s\\
&= \int_t^T e^{\alpha s}(-\alpha)|\overline{Y}_{n+1}(s)|^2\,\mathrm{d}s - 2\int_t^T e^{\alpha s}\langle \overline{Y}_{n+1}(s), \overline{Z}_{n+1}(s)\,\mathrm{d}W(s)\rangle\\
&\quad + 2\int_t^T e^{\alpha s}\langle \overline{Y}_{n+1}(s), f'(s, X_n(s), Y_n(s), Z_n(s))h_{n+1}(s)\\
&\quad + \mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h_n(s)\rangle\,\mathrm{d}s.
\end{aligned} \tag{4.5}$$

From the Cauchy-Schwarz inequality and the inequality $2ab \le \delta^{-1}|a|^2 + \delta|b|^2$ for all $\delta > 0$, we have

$$\begin{gathered}
2|\langle \overline{Y}_{n+1}(s), f'_x(s, X_n(s), Y_n(s), Z_n(s))\overline{X}_{n+1}(s)\rangle|\\
\le 2\|f'\|_\infty^2|\overline{Y}_{n+1}(s)|^2 + \frac{1}{2}|\overline{X}_{n+1}(s)|^2,\\
2|\langle \overline{Y}_{n+1}(s), f'_y(s, X_n(s), Y_n(s), Z_n(s))\overline{Y}_{n+1}(s)\rangle| \le 2\|f'\|_\infty|\overline{Y}_{n+1}(s)|^2,\\
2|\langle \overline{Y}_{n+1}(s), f'_z(s, X_n(s), Y_n(s), Z_n(s))\overline{Z}_{n+1}(s)\rangle|\\
\le 2\|f'\|_\infty^2|\overline{Y}_{n+1}(s)|^2 + \frac{1}{2}|\overline{Z}_{n+1}(s)|^2,\\
2|\langle \overline{Y}_{n+1}(s), \mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h_n(s)\rangle|\\
\le \delta^{-1}|\overline{Y}_{n+1}(s)|^2 + \delta|\mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h_n(s)|^2,
\end{gathered}$$

and

$$|\mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h_n(s)| \le 2\|f'\|_\infty|h_n(s)|. \tag{4.6}$$

The right-hand side of (4.5) is therefore less than or equal to

$$\begin{aligned}
&\int_t^T e^{\alpha s}\cdot\{(-\alpha) + 2\|f'\|_\infty + 4\|f'\|_\infty^2 + \delta^{-1}\}|\overline{Y}_{n+1}(s)|^2\,\mathrm{d}s\\
&+ \int_t^T e^{\alpha s}\{\frac{1}{2}|\overline{Z}_{n+1}(s)|^2 + \frac{1}{2}|\overline{X}_{n+1}(s)|^2 + \delta|\mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h_n(s)|^2\}\,\mathrm{d}s\\
&- 2\int_t^T e^{\alpha s}\langle \overline{Y}_{n+1}(s), \overline{Z}_{n+1}(s)\,\mathrm{d}W(s)\rangle.
\end{aligned}$$

Setting $\alpha \equiv \alpha(\delta) = 2\|f'\|_\infty + 4\|f'\|_\infty^2 + \delta^{-1}$, we obtain

$$
\begin{aligned}
&e^{\alpha t}|\overline{Y}_{n+1}(t)|^2 - e^{\alpha T}|\overline{Y}_{n+1}(T)|^2 + \frac{1}{2}\int_t^T e^{\alpha s}\{|\overline{Z}_{n+1}(s)|^2 - |\overline{X}_{n+1}(s)|^2\}\,\mathrm{d}s \\
&\le \delta \int_t^T e^{\alpha s}|\mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h_n(s)|^2\,\mathrm{d}s \\
&\quad - 2\int_t^T e^{\alpha s}\langle \overline{Y}_{n+1}(s), \overline{Z}_{n+1}(s)\,\mathrm{d}W(s)\rangle.
\end{aligned} \tag{4.7}
$$

As $(Y, Z), (Y_n, Z_n) \in \mathbb{S}_m^2 \times \mathbb{H}^2$, the local martingale

$$
\Big\{ \int_0^t e^{\alpha s}\langle \overline{Y}_{n+1}(s), \overline{Z}_{n+1}(s)\,\mathrm{d}W(s)\rangle \Big\}_{t\in[0,T]}
$$

vanishes at 0. Thus, in particular, by setting $t = 0$, we obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^T e^{\alpha s}\{|\overline{Z}_{n+1}(s)|^2 - |\overline{X}_{n+1}(s)|^2\}\,\mathrm{d}s\Big] \\
&\le 2\delta \mathbb{E}\Big[\int_0^T e^{\alpha s}|\mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h_n(s)|^2\,\mathrm{d}s\Big] + 2\mathbb{E}[e^{\alpha T}|\overline{Y}_{n+1}(T)|^2].
\end{aligned} \tag{4.8}
$$

Note that

$$
\int_t^T e^{\alpha s}\langle \overline{Y}_{n+1}(s), \overline{Z}_{n+1}(s)\,\mathrm{d}W(s)\rangle = \int_0^T 1_{[t,T]}(s) e^{\alpha s}\langle \overline{Y}_{n+1}(s), \overline{Z}_{n+1}(s)\,\mathrm{d}W(s)\rangle.
$$

An application of the Burkholder-Davis-Gundy inequality indicates that there is a universal c_0 such that

$$
\begin{aligned}
&2\mathbb{E}\Big[\sup_{0\le t\le T}\Big|\int_t^T e^{\alpha s}\langle \overline{Y}_{n+1}(s), \overline{Z}_{n+1}(s)\,\mathrm{d}W(s)\rangle\Big|\Big] \\
&\le 2c_0\mathbb{E}\Big[\Big(\int_0^T 1_{[t,T]}(s)e^{2\alpha s}|\overline{Y}_{n+1}(s)|^2|\overline{Z}_{n+1}(s)|^2\,\mathrm{d}s\Big)^{1/2}\Big] \\
&\le \mathbb{E}\Big[\sup_{0\le t\le T} e^{\alpha t/2}|\overline{Y}_{n+1}(t)|\Big(4c_0^2\int_0^T e^{\alpha s}|\overline{Z}_{n+1}(s)|^2\,\mathrm{d}s\Big)^{1/2}\Big],
\end{aligned}
$$

where we note that an explicit upper bounded of c_0 can be obtained by 3; refer to [11, Theorem 3.28]. Hence, by considering the supremum of (4.7) and using the inequality $ab \le (a^2/2) + (b^2/2)$ for all $a, b \in \mathbb{R}$, we obtain

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\le t\le T} e^{\alpha t}|\overline{Y}_{n+1}(t)|^2 - e^{\alpha T}|\overline{Y}_{n+1}(T)|^2 \\
&+ \frac{1}{2}\int_0^T e^{\alpha s}\{|\overline{Z}_{n+1}(s)|^2 - |\overline{X}_{n+1}(s)|^2\}\,\mathrm{d}s\Big] \\
&\le \delta\int_0^T e^{\alpha s}\mathbb{E}[|\mathrm{R}_f(X_n(s), Y_n(s), Z_n(s))h_n(s)|^2]\,\mathrm{d}s \\
&\quad + \frac{1}{2}\mathbb{E}\big[\sup_{0\le t\le T} e^{\alpha t}|\overline{Y}_{n+1}(t)|^2\big] + 2c_0^2\mathbb{E}\Big[\int_0^T e^{\alpha s}|\overline{Z}_{n+1}(s)|^2\,\mathrm{d}s\Big],
\end{aligned}
$$

which implies that

$$
\mathbb{E}\Big[\sup_{0\le t\le T} e^{\alpha t}|\overline{Y}_{n+1}(t)|^2 - 2e^{\alpha T}|\overline{Y}_{n+1}(T)|^2
$$

$$
\begin{aligned}
&+\int_0^T e^{\alpha s}\{|\overline{Z}_{n+1}(s)|^2-|\overline{X}_{n+1}(s)|^2\}\,\mathrm{d}s\Big]\\
&\le 2\delta\int_0^T e^{\alpha s}\mathbb{E}[|\mathrm{R}_f(X_n(s),Y_n(s),Z_n(s))h_n(s)|^2]\,\mathrm{d}s\\
&\quad+4c_0^2\mathbb{E}\Big[\int_0^T e^{\alpha s}|\overline{Z}_{n+1}(s)|^2\,\mathrm{d}s\Big].
\end{aligned}
$$

Then by taking the inequality (4.8) into consideration, we observe that

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\le t\le T}e^{\alpha t}|\overline{Y}_{n+1}(t)|^2+\int_0^T e^{\alpha s}\{|\overline{Z}_{n+1}(s)|^2-|\overline{X}_{n+1}(s)|^2\}\,\mathrm{d}s\Big]\\
&\le 2\delta(1+4c_0^2)\int_0^T e^{\alpha s}\mathbb{E}[|\mathrm{R}_f(X_n(s),Y_n(s),Z_n(s))h_n(s)|^2]\,\mathrm{d}s\\
&\quad+2(1+4c_0^2)e^{\alpha T}\mathbb{E}[|\overline{Y}_{n+1}(T)|^2]+4c_0^2\mathbb{E}\Big[\int_0^T e^{\alpha s}|\overline{X}_{n+1}(s)|^2\,\mathrm{d}s\Big].
\end{aligned}
$$

If we also consider the inequality (4.6), we obtain

$$
\begin{aligned}
&\int_0^T e^{\alpha s}\mathbb{E}[|\mathrm{R}_f(X_n(s),Y_n(s),Z_n(s))h_n(s)|^2]\,\mathrm{d}s\\
&\le 6\|f'\|_\infty(1\vee T)\Big\{\mathbb{E}\Big[\sup_{0\le t\le T}e^{\alpha t}|\overline{X}_n(t)|^2+\sup_{0\le t\le T}e^{\alpha t}|\overline{Y}_n(t)|^2\\
&\quad+\int_0^T e^{\alpha s}|\overline{Z}_n(s)|^2\,\mathrm{d}s\Big]\Big\}.
\end{aligned}
$$

Selecting δ such that $12\|f'\|_\infty(1+4c_0^2)(1\vee T)\delta=\epsilon$, we obtain

$$
\alpha\equiv\alpha(\delta)=2\|f'\|_\infty+4\|f'\|_\infty^2+12\|f'\|_\infty(1+4c_0^2)(1\vee T)\epsilon^{-1}.
$$

This leads to

$$
\begin{aligned}
&\|(\overline{Y}_{n+1},\overline{Z}_{n+1})\|_\alpha^2\\
&\le\epsilon\|(\overline{X}_n,\overline{Y}_n,\overline{Z}_n)\|_\alpha^2+2(1+4c_0^2)e^{\alpha T}\|\overline{Y}_{n+1}(T)\|_{\mathbb{L}^2}^2+(1+4c_0^2)\|\overline{X}_{n+1}\|_\alpha^2.
\end{aligned}
$$

Hence, using (4.3),

$$
\begin{aligned}
\mathbb{E}[|\overline{Y}_{n+1}(T)|^2]&\le\|\varphi'\|_\infty^2\{4\mathbb{E}[|\overline{X}_n(T)|^2]+2\mathbb{E}[|\overline{X}_{n+1}(T)|^2]\}\\
&\le 2(2+TC_0)\|\varphi'\|_\infty^2\mathbb{E}[|\overline{X}_n(T)|^2].
\end{aligned}
$$

Thus, we obtain

$$
2(1+4c_0^2)e^{\alpha T}\|\overline{Y}_{n+1}(T)\|_{\mathbb{L}^2}^2\le 4(2+TC_0)\|\varphi'\|_\infty^2(1+4c_0^2)e^{\alpha T}\|\overline{X}_n\|_{\mathbb{S}_d^2}^2.
$$

Applying inequality (4.2) for $c_1=e^{\alpha T}+(1+4(2+TC_0)\|\varphi'\|_\infty^2)(1+4c_0^2)e^{\alpha T}$, yields

$$
\|(\overline{Y}_{n+1},\overline{Z}_{n+1})\|_\alpha^2\le\epsilon\|(\overline{Y}_n,\overline{Z}_n)\|_\alpha^2+c_1\|X-X_0\|^2\frac{C_0^n}{n!}.
$$

For any positive sequence $\{a_n\}$, $\{b_n\}$ and $\epsilon>0$ such that $a_{n+1}\le b_n+\epsilon a_n$ for all $n\in\mathbb{N}\cup\{0\}$, we obtain

$$
\begin{aligned}
a_{n+1}&\le b_n+\epsilon(\epsilon a_{n-1}+b_{n-1})\\
&=\epsilon^2(\epsilon^{-2}b_n+\epsilon^{-1}b_{n-1}+a_{n-1})\le\cdots
\end{aligned}
$$

$$\le \epsilon^{n+1}\Big\{\sum_{k=0}^{n}\epsilon^{k-(n+1)}b_{n-k}+a_0\Big\}.$$

Replacing $b_n = c_1\|X-X_0\|^2\frac{C_0^n}{n!}$ for all $n\in\mathbb{N}$, we obtain

$$\sum_{k=0}^{n}\epsilon^{k-(n+1)}b_{n-k}\le \epsilon^{-1}c_1\|X-X_0\|^2\sum_{k=0}^{n}\frac{(C_0/\epsilon)^k}{k!}\le \epsilon^{-1}c_1\|X-X_0\|^2e^{C_0/\epsilon},$$

for $n\in\mathbb{N}\cup\{0\}$. Thus, we obtain

$$\|(\overline{Y}_{n+1},\overline{Z}_{n+1})\|_\alpha^2\le \epsilon^{n+1}(C_1\|X-X_0\|^2+\|(\overline{Y}_0,\overline{Z}_0)\|_\alpha^2),$$

where

$$C_1=\epsilon^{-1}c_1e^{C_0/\epsilon}=\epsilon^{-1}\{1+(1+4(2+TC_0)\|\varphi'\|_\infty^2)(1+4c_0^2)\}e^{\alpha T+C_0/\epsilon}.$$

This completes the proof. □

Remark 4.4. We can also prove the so called "semilocal theorem", which states that the approximation process is a Cauchy sequence in $\mathbb{S}_m^2\times\mathbb{H}^2$ by the same argument. For the definition of "semilocal", refer to [24].

We can now prove our main result based on the following theorem.

Theorem 4.5. *Let (X,Y,Z) be a solution of the FBSDE* (1.1). *If Assumptions 3.1 and 3.2 hold and $(X_0,Y_0,Z_0)\in\mathbb{S}_d^2\times\mathbb{S}_m^2\times\mathbb{H}^2$ with $X_0(0)=X(0)$, then, there exists a $C_3>0$ such that, for all $n\in\mathbb{N}\cup\{0\}$, we obtain*

$$\begin{aligned}&\|(X-X_{n+1},Y-Y_{n+1},Z-Z_{n+1})\|^2\\&\le \epsilon^{n+1}C_3\|(X-X_0,Y-Y_0,Z-Z_0)\|^2,\end{aligned}\tag{4.9}$$

where the constant $C_3=C_3(\epsilon)$ is bounded by the coefficients and independent of n.

Proof. By inequalities (4.2) and (4.4), we obtain

$$\|X-X_{n+1}\|^2\le \epsilon^{n+1}e^{C_0T/\epsilon}\|X-X_0\|^2,$$

and

$$\begin{aligned}\|(Y-Y_{n+1},Z-Z_{n+1})\|^2&\le\|(Y-Y_{n+1},Z-Z_{n+1})\|_\alpha^2\\&\le \epsilon^{n+1}\big\{C_1\|X-X_0\|_{\mathbb{S}_d^2}^2+\|(Y-Y_0,Z-Z_0)\|_\alpha^2\big\},\end{aligned}$$

respectively, where

$$\begin{gathered}C_0=8c_{b,\sigma}T\exp(4c_{b,\sigma}T),\quad c_{b,\sigma}=\|b'\|_\infty+18\|\sigma'\|_\infty+\|\sigma'\|_\infty^2,\\ C_1\equiv C_1(\epsilon)=\epsilon^{-1}\{1+(1+4(2+TC_0)\|\varphi'\|_\infty^2)(1+4c_0^2)\}e^{\alpha T+C_0/\epsilon},\\ c_0=3,\quad \alpha=2\|f'\|_\infty+4\|f'\|_\infty^2+12\|f'\|_\infty(1+4c_0^2)(1\vee T)\epsilon^{-1}.\end{gathered}$$

Defining

$$C_3=\{(C_1+e^{C_0T/\epsilon})\vee e^{\alpha T}\},$$

we complete the proof. □

This finally allows us to prove our main theorem.

Proof of Theorem 1.1. By setting $\epsilon=1/2$, the desired result can be obtained from Theorem 4.5. □

Acknowledgements. We would like to express our sincere appreciation to Professor Toshio Yamada for insightful comments. This work was supported by JSPS KAKENHI Grant Numbers 16J00894 and 17H06833.

References

[1] K. Amano; A note on Newton's method for stochastic differential equations and its error estimate. *Proc. Japan Acad., Ser. A*, 85(3) (2009), 19–21.
[2] K. Amano; Newton's method for stochastic differential equations and its probabilistic second-order error estimate. *Electron. J. Differ. Equ.*, 2012 (2012) no. 3, 1–8.
[3] F. Antonelli; Backward-Forward Stochastic Differential Equations. *Ann. Appl. Probab.*, 3(3) (1193), 777–793.
[4] K. Bahlali, M. Eddahbi, Y. Ouknine; Quadratic BSDE with L^2 -terminal data: Krylov's estimate, Itô-Krylov's formula and existence results. *Annals of Probability*, 45(4) (2017), 2377–2397.
[5] C. Bender and J. Zhang. Time discretization and Markovian iteration for coupled FBSDEs. *Ann. Appl. Probab.*, 18(1) (2008), 143–177.
[6] F. Delarue; On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stochastic Processes and their Applications*, 99(2) (2002), 209 – 286.
[7] F. Delarue, S. Menozzi. A forward–backward stochastic algorithm for quasi-linear PDEs. *Ann. Appl. Probab.*, 16(1) (2006), 140–184.
[8] N. El Karoui, S. Peng, M. C. Quenez; Backward stochastic differential equations in finance. *Math. Finance*, 7(1) (1997), 1–71.
[9] Y. Hu, S. Peng; Solution of forward-backward stochastic differential equations. *Probability Theory and Related Fields*, 103 (2) (1995), 273–283.
[10] K. Itô; Differential equations determining a Markov process. *Journal of the Pan-Japanese Mathematical Colloquium*, 1077 (1942).
[11] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus (Graduate Texts in Mathematics)*. Springer, 2nd edition, aug 1991.
[12] S. Kawabata, T. Yamada; On Newton's method for stochastic differential equations. In *Séminaire de Probabilités, XXV*, volume 1485 of *Lecture Notes in Math.*, pages 121–137. Springer, Berlin, 1991.
[13] J. Ma, P. Protter, J. Yong; Solving forward-backward stochastic differential equations explicitly — a four step scheme. *Probability Theory and Related Fields*, 98(3) (1994), 339–359.
[14] J. Ma, Z. Wu, D. Zhang, J. Zhang; On well-posedness of forward-backward SDEs – a unified approach. *Ann. Appl. Probab.*, 25(4) (2015), 2168–2214.
[15] J. Ma, H. Yin, J. Zhang; On non-Markovian forward–backward SDEs and backward stochastic PDEs. *Stochastic Processes and their Applications*, 122(12) (2012), 3980– 4004.
[16] J. M. Ortega, W. C. Rheinboldt; *Iterative solution of nonlinear equations in several variables*, volume 30 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1970 original.
[17] Y. Ouknine; Approximation de Newton pour les équations différentielles stochastiques. *Stochastics and Stochastic Reports*, 45 (3-4) (1993), 237–247.
[18] E. Pardoux, S. Peng; Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1) (1990), 55–61.
[19] E. Pardoux, A. Răşcanu. *Stochastic differential equations, backward SDEs, partial differential equations*. Cham: Springer, 2014.
[20] E. Pardoux, S. Tang; Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probability Theory and Related Fields*, 114 (2) (1999), 123–150.
[21] S. Peng, Z. Wu; Fully Coupled Forward-Backward Stochastic Differential Equations and Applications to Optimal Control. *SIAM Journal on Control and Optimization*, 37(3) (1999), 825–843.
[22] G. Vidossich; Chaplygin's method is Newton's method. *Journal of Mathematical Analysis and Applications*, 66 (1) (1978), 188–206.
[23] M. Wrzosek; Newton's method for stochastic functional differential equations. *Electron. J. Differential Equations*, 2012 (2012) No. 130, 1–10.

[24] T. Yamamoto; Historical developments in convergence analysis for Newton's and Newton-like methods. *J. Comput. Appl. Math.*, 124(1-2):1–23, 2000. Numerical analysis 2000, Vol. IV, Optimization and nonlinear equations.

[25] J. Yong; Finding adapted solutions of forward–backward stochastic differential equations: method of continuation. *Probability Theory and Related Fields*, 107(4) (1997), 537–572.

[26] Y. Niwa; Newton's method for backward stochastic differential equations. Master's thesis, Ritsumeikan university (Japanese), 2003.

[27] J. Zhang; *Backward stochastic differential equations. From linear to fully nonlinear theory.* New York, NY: Springer, 2017.

Dai Taguchi
Research Institute for Interdisciplinary Science, Okayama University 3-1-1 Tsushima-naka, Kita-ku Okayama 700-8530, Japan
Email address: `dai.taguchi.dai@gmail.com`

Takahiro Tsuchiya
School of Computer Science and Engineering, The University of Aizu, Japan
Email address: `suci@probab.com`