SIAM J. NUMER. ANAL.
Vol. 58, No. 4, pp. 2351–2375

A UNIFIED PROBABILISTIC DISCRETIZATION SCHEME FOR FBSDEs: STABILITY, CONSISTENCY, AND CONVERGENCE ANALYSIS*

JIE YANG†, WEIDONG ZHAO‡, AND TAO ZHOU§

Abstract. In this work, we propose a general discretization framework for numerical solution of forward backward stochastic differential equations (FBSDEs). The framework covers several existing temporal discretization probabilistic schemes in the literature. The consistency, stability, and convergence analysis for the proposed scheme are presented. In particular, we prove a stochastic mean square version of Lax equivalence theorem—showing that a consistent discretization scheme for FBSDEs is convergent if and only if it is stable. Applications of the analysis results to existing numerical schemes are also discussed.

Key words. forward backward stochastic differential equations, numerical schemes, stability, consistency, convergence analysis

AMS subject classifications. 60H35, 65C20, 93E15, 60H10

DOI. 10.1137/19M1260177

1. Introduction.

The study of forward backward stochastic differential equations (FBSDEs) is motivated by their wide application in many important fields, such as mathematical finance, stochastic optimal control, risk measure, and game theory (see, e.g., [23, 29, 34, 41] and references therein).

As analytical solutions of FBSDEs are seldom available, numerical methods have become popular tools for solving FBSDEs. In recent years, great efforts have been made for designing efficient numerical schemes for FBSDEs. Up to now two main types of schemes have been considered: the first type is based on numerical solution of a parabolic PDE which is related to the BSDE [19, 33, 36], while the second type of algorithms works backward through time and tries to tackle the BSDEs directly [1, 5, 7, 17, 27, 35, 42, 44, 48]. This work falls into the second type of almost all existing schemes, where a spatial-temporal discretization is usually needed. From the temporal discretization viewpoint, popular strategies include Euler-type methods [24, 26, 47], generalized θ-schemes [43, 51], Runge–Kutta schemes [12], and multistep schemes [11, 22, 49, 53, 54], to name a few. From the spatial discretization point of view, approximation techniques can be roughly classified into grid-type methods

*Received by the editors May 6, 2019; accepted for publication (in revised form) May 15, 2020; published electronically August 17, 2020.
https://doi.org/10.1137/19M1260177
Funding: The work of the authors was supported by National Key Basic Research Program grants 2018YFB0704304 and 2018YFA0703903, Science Challenge Project grant TZ2018001, National Natural Science Foundation of China grants 11822111, 11688101, 11731006, 11571206, 11831010 and 11801320, the Youth Innovation Promotion Association of CAS, and National Natural Science Foundation of Shandong Province grant ZR2018BA005.
†School of Mathematics and Statistics, Shandong University, Weihai, Shandong 264209, China (jieyang@sdu.edu.cn).
‡School of Mathematics and Institute of Finance, Shandong University, Jinan, Shandong 250100, China (wdzhao@sdu.edu.cn).
§LSEC and NCMIS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (tzhou@lsec.cc.ac.cn).

[11, 21, 49], regression approaches [24, 26], cubature methods [15], machine learning approach [3, 20], etc.

It is well known that for numerical schemes of ODEs/PDEs, the most important features are stability, consistency, and the convergence property. While this has been well studied in past decades for ODEs [10], the corresponding study for FBSDEs is not well developed due to the complex solution structure of FBSDEs. In particular, the direct extension of the stability and convergence analysis of numerical ODEs to FBSDEs seems to be highly nontrivial [13, 50, 52].

In this work, we shall first propose a very general discretization framework for numerical solution of FBSDEs, by combining an integral discretization method and a derivative discretization method. The framework covers many interesting temporal discretization probabilistic schemes in the literature, such as Euler-type schemes [4, 6, 14, 24, 25, 26, 32, 47], θ-schemes [43, 48, 52], multistep schemes [11, 22, 45, 49, 54], etc. Another main contribution is to build a theoretical framework for analyzing the proposed discretization scheme; more precisely, we aim at analyzing the stability, consistency, and convergence property of the proposed discretization scheme. We also obtain a stochastic mean square version of the Lax equivalence theorem, which shows that a consistent discretization scheme for FBSDEs is convergent if and only if it is stable. We would like to mention that the stability analysis in this work is restricted to a simplified version of the general numerical scheme (Scheme 3.1), which, however, still covers many existing probabilistic schemes as mentioned above. Thus, the analysis here is applicable to many popular schemes in the literature, and one may also expect to generate new efficient numerical schemes based on the general discretization framework.

The rest of the paper is organized as follows. In section 2, we introduce some preliminaries that will be used in this work. In section 3, we present a general discretization framework for FBSDEs. The stability, consistency, and convergence properties of the proposed scheme are analyzed in section 4. Applications of the analysis results to some specific schemes are discussed in section 5. Finally, we give some concluding remarks in section 6.

2. Preliminaries.

Let $(\Omega, \mathcal{F}, \mathbb{F}, P)$ be a complete filtered probability space with $\mathbb{F} = \{\mathcal{F}_t\}_{0\leq t\leq T}$ being the natural filtration generated by a standard m-dimensional Brownian motion $W_t = (W_t^1, \dots, W_t^m)^\top$, and $0 \leq t \leq T$ with T being a fixed positive number. We list some notation that will be used throughout this paper:

- $|\cdot|$: the Euclidean norm in the space $\mathbb{R}^d$, $\mathbb{R}^p$ or $\mathbb{R}^{p\times m}$.
- $x^\top$: the transpose of a vector x.
- $\mathcal{F}_s^{t,x}(t \leq s \leq T)$: the σ-algebra generated by the diffusion process $\{X_r, t \leq r \leq s, X_t = x\}$.
- $\mathbb{E}_s^{t,x}[\cdot]$: the conditional expectation under $\mathcal{F}_s^{t,x}$, i.e., $\mathbb{E}_s^{t,x}[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_s^{t,x}]$. For $s = t$, we denote $\mathbb{E}_t^x[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_t^{t,x}]$.
- $L^2([0,T])$: the Hilbert space of deterministic square integrable functions.
- $L^2(\Omega\times[0,T])$: the space of square integrable, progressively measurable functions ϕ_s on $\Omega\times[0,T]$ satisfying $\mathbb{E}[\int_0^T \phi_s^2 ds] < \infty$.
- C_p^k: the space of k times continuously differentiable functions with at most polynomial growth for their partial derivatives of order up to k.
- C_p^∞: the space of smooth functions with partial derivatives of polynomial growth.

We consider the following FBSDE that is defined on $(\Omega, \mathcal{F}, \mathbb{F}, P)$:

$$
\begin{cases} X_t = X_0 + \int_0^t b(r, X_r)dr + \int_0^t \sigma(r, X_r)dW_r & \text{(FSDE)}, \\ Y_t = \xi + \int_t^T f(r, X_r, Y_r, Z_r)dr - \int_t^T Z_r dW_r & \text{(BSDE)}. \end{cases} \tag{2.1}
$$

Here $t \in [0, T]$, $b(t, x) : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma(t, x) : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are respectively referred to as the drift and diffusion coefficients of the forward SDE, and $f(t, x, y, z) : [0, T] \times \mathbb{R}^d \times \mathbb{R}^p \times \mathbb{R}^{p\times m} \to \mathbb{R}^p$ and $\xi = \varphi(X_T)$ are the generator and the terminal condition of the BSDE, respectively. The two integrals in (2.1) with respect to an m-dimensional Brownian motion W_r are of Itô type. A triple (X_t, Y_t, Z_t) is called an L^2-adapted solution of the FBSDEs in (2.1) if it is $\mathcal{F}_t$-adapted and square integrable and satisfies (2.1).

In what follows, we assume that X_0 is $\mathcal{F}_0$-measurable, and we assume that ξ is $\mathcal{F}_T$-measurable with $\mathbb{E}[|\xi|^2] < \infty$. Moreover, we also make the following assumptions.

Assumption 1. We assume that the functions b, σ, f, and ξ satisfy the following conditions:

- The functions b and σ are L^2-measurable in $[0, T] \times \mathbb{R}^d$ that satisfy the following linear growth condition and the Lipschitz condition:
$$
\begin{aligned} &|b(t, x)| + |\sigma(t, x)| \le \tilde{K}(1 + |x|), \\ &|b(t, x) - b(t, x')| + |\sigma(t, x) - \sigma(t, x')| \le \tilde{L}|x - x'|. \end{aligned}
$$
Here $\tilde{K}$ and $\tilde{L}$ are positive constants.
- For a fixed $t \in [0, T]$, we have
$$
f(t, X_t, Y_t, Z_t) \in \mathcal{F}_t, \quad \mathbb{E}\left[|f(t, 0, 0, 0)|^2\right] < \infty.
$$
- The function f is uniformly Lipschitz with a positive constant L, i.e., for $x, x' \in \mathbb{R}^d$, $y, y' \in \mathbb{R}^p$, and $z, z' \in \mathbb{R}^{p\times m}$, we have
$$
|f(t, x, y, z) - f(t, x', y', z')| \le L(|x - x'| + |y - y'| + |z - z'|).
$$

Notice that the FBSDE (2.1) is well posed under the above assumption [28, 29, 40].

The Itô isometry. The following Itô isometry formula [30] will play an important role when designing discretization schemes.

LEMMA 2.1. *For adapted square measurable processes g_t and h_t, it holds that*

$$
\mathbb{E}\left[\left(\int_0^T g_t dW_t\right)\left(\int_0^T h_t dW_t\right)\right] = \mathbb{E}\left[\int_0^T g_t h_t dt\right]. \tag{2.2}
$$

The Malliavin derivative. Let $\mathcal{S}$ be the set of random variable F of the form $F = \psi(W(h_1), \ldots, W(h_l))$, where $\psi \in C_p^\infty$ and $W(h_i) = \int_0^T h_i(t)dW_t$ with $h_1, \ldots, h_l \in L^2([0, T])$. For $F \in \mathcal{S}$, the Malliavin derivative of F is defined as the process $D_t F$ valued in $L^2(\Omega \times [0, T])$:

$$
D_t F = \sum_i^l \frac{\partial f}{\partial x_i}(W(h_1), \ldots, W(h_l))h_i(t), \quad t \in [0, T].
$$

The associated norm of $\mathcal{S}$ is defined as

$$
\|F\|_{1,2} = \left(\mathbb{E}\left[|F|^2\right] + \mathbb{E}\left[\int_0^T |D_t F|^2 dt\right]\right)^{1/2}.
$$

Then we denote by $\mathbb{D}^{1,2}$ the Banach space which is the completion of $\mathcal{S}$ with respect to $\| \|_{1,2}$.

For the Malliavin derivative, the following lemma holds.

LEMMA 2.2 (see [38]). *If $F = (F_1, \dots, F_d)$ ($F_i \in \mathbb{D}^{1,2}$) is a random vector and $u \in L^2(\Omega \times [0,T])$, then we have*

$$D_t \left(\int_0^T u_s dW_s \right) = u_t + \int_t^T D_t u_s dW_s.$$

Moreover, we have the integration-by-parts formula

$$\mathbb{E}\left[\int_0^T u_t\, D_t F dt \right] = \mathbb{E}\left[F \int_0^T u_t dW_t \right]. \tag{2.3}$$

3. A unified temporal discretization scheme for FBSDEs. In this section, we shall present a general temporal discretization framework for FBSDEs. To this end, we introduce the following time partition π_h for $[0,T]$:

$$\pi_h = \left\{ t_n, n = 0, 1, \dots, N \,|\, 0 = t_0 < t_1 < \cdots < t_N = T \right\}.$$

Here N is a positive integer. We impose the following regularity constraint for the partition π_h:

$$\frac{\max\limits_{0 \le n \le N-1} \Delta t_n}{\min\limits_{0 \le n \le N-1} \Delta t_n} \le c_0, \tag{3.1}$$

where c_0 is a positive constant. We set

$$\Delta t_n = t_{n+1} - t_n, \quad \Delta t = \max_{0 \le n \le N-1} \Delta t_n.$$

For the Brownian motion W_t, we denote by $\Delta W_{n,i}$ the increment $W_{t_{n+i}} - W_{t_n}$ and by $\Delta W_{t_n,s}$ the increment $W_s - W_{t_n}$ for $s \ge t_n$.

We will use (X^n, Y^n, Z^n) to represent the numerical solution of FBSDEs (2.1) for $t = t_n$. For notational simplicity, we shall also represent $f(t, X_t, Y_t, Z_t)$ by f_t and represent $f(t_n, X^n, Y^n, Z^n)$ by f_n.

Let (X_t, Y_t, Z_t) be the exact solution of FBSDE (2.1). Then for $t \in [t_n, T]$ it holds that

$$Y_{t_n} = Y_t + \int_{t_n}^t f_s ds - \int_{t_n}^t Z_s dW_s. \tag{3.2}$$

By taking the conditional expectation $\mathbb{E}_{t_n}^x[\cdot]$ on (3.2), we obtain

$$Y_{t_n} = \mathbb{E}_{t_n}^x[Y_t] + \int_{t_n}^t \mathbb{E}_{t_n}^x[f_s] ds. \tag{3.3}$$

Next we introduce the following Gaussian process $G_{t_n,t}$:

$$G_{t_n,t} = \left(\int_{t_n}^t \phi_s dW_s \right)^\top, \quad 0 \le t_n < t \le T, \quad \phi_s \in L^2(\Omega \times [0,T]).$$

Note that for any $\phi_s \in L^2(\Omega \times [0,T])$, it is clear that $G_{t_n,t}$ is a martingale process that satisfies $\mathbb{E}^x_{t_n}[G_{t_n,t}] = 0$ and $\mathbb{E}^x_{t_n}[G^2_{t_n,t}] = \mathbb{E}^x_{t_n}[\int_{t_n}^t \phi_s^2 ds]$. Throughout this paper we assume that

$$c_\phi(t-t_n) \le \mathbb{E}[|G_{t_n,t}|^2] \le (t-t_n) \max_{t_n\le s\le t} \mathbb{E}[\phi_s^2] \le C_\phi(t-t_n), \tag{3.4}$$

where c_ϕ and C_ϕ are two positive constants. A simple choice for ϕ_s is $\phi_s \equiv 1$ and general choices have also been discussed [11, 12, 13, 54], e.g., ϕ_s is a polynomial on $[0,T]$. For all these cases, the above assumption (3.4) holds true.

Now, by multiplying (3.2) by $G_{t_n,t}$ and taking the conditional expectation $\mathbb{E}^x_{t_n}[\cdot]$ on the derived equation and using the Itô isometry (2.2), we have

$$0 = \mathbb{E}^x_{t_n}[Y_t G_{t_n,t}] + \int_{t_n}^t \mathbb{E}^x_{t_n}[f_s G_{t_n,s}]ds - \int_{t_n}^t \mathbb{E}^x_{t_n}[Z_s\phi_s]ds. \tag{3.5}$$

The two derived equations, (3.3) and (3.5), are the basis for designing time-discretization schemes for FBSDEs. We shall present two different ways to design numerical schemes, namely, the differentiation discretization method and the integration discretization method. Then, we shall propose a more general framework by combining the two methods.

3.1. The differentiation discretization method. We first present the differentiation discretization method. Notice that under mild conditions (such as the Lipschitz conditions) on the data b, σ φ, and f, it can be verified that the integrands in (3.3) and (3.5) are continuous with respect to s. Thus, by taking the derivative with respect to t in (3.3) and (3.5) we obtain

$$\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t]}{\mathrm{d}t} = -\mathbb{E}^x_{t_n}[f_t], \tag{3.6a}$$

$$\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t G_{t_n,t}]}{\mathrm{d}t} = -\mathbb{E}^x_{t_n}[f_t G_{t_n,t}] + \mathbb{E}^x_{t_n}[Z_t\phi_t]. \tag{3.6b}$$

By taking $t \to t_{n+0}$ in the above equations we get

$$\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t]}{\mathrm{d}t}\right|_{t\to t_{n+0}} = -f_{t_n}, \tag{3.7a}$$

$$\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t G_{t_n,t}]}{\mathrm{d}t}\right|_{t\to t_{n+0}} = Z_{t_n}\phi_{t_n}. \tag{3.7b}$$

We can now approximate the derivatives in (3.7) by using k-step schemes:

$$\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t]}{\mathrm{d}t}\right|_{t\to t_{n+0}} = \sum_{i=0}^k \alpha_i^{n,k}\mathbb{E}^x_{t_n}\left[Y_{t_{n+i}}\right] + R^k_{y,n}, \tag{3.8a}$$

$$\left.\frac{\mathrm{d}\mathbb{E}^x_{t_n}[Y_t G_{t_n,t}]}{\mathrm{d}t}\right|_{t\to t_{n+0}} = \sum_{i=0}^k \tilde{\alpha}_i^{n,k}\mathbb{E}^x_{t_n}\left[Y_{t_{n+i}}G^{n,i}\right] + R^k_{z,n}, \tag{3.8b}$$

where the parameters $\{\alpha_i^{n,k}\}_{0\le i\le k}$ and $\{\tilde{\alpha}_i^{n,k}\}_{0\le i\le k}$ are real numbers, $R^k_{y,n}$ and $R^k_{z,n}$ are the associated truncation errors, and $G^{n,i} := G_{t_n,t_{n+i}}$. Notice that by the classical derivative approximation theory, the truncation errors $R^k_{y,n}$ and $R^k_{z,n}$ can be well controlled if the parameters are properly chosen [49].

Now by inserting (3.8) into (3.7), we obtain

$$
\begin{cases} \alpha_0^{n,k} Y_{t_n} = -\sum_{i=1}^{k} \alpha_i^{n,k} \mathbb{E}_{t_n}^x \left[Y_{t_{n+i}} \right] - f_{t_n} - R_{y,n}^k, \\ Z_{t_n} = \sum_{i=0}^{k} \tilde{\tilde{\alpha}}_i^{n,k} \mathbb{E}_{t_n}^x \left[Y_{t_{n+i}} G^{n,i} \right] + \dfrac{1}{\phi_{t_n}} R_{z,n}^k, \end{cases} \tag{3.9}
$$

where $\tilde{\tilde{\alpha}}_i^{n,k} = \tilde{\alpha}_i^{n,k}/\phi_{t_n}$. One can then obtain the k-step temporal discretization schemes by removing the error terms $R_{y,n}^k$ and $\frac{1}{\phi_{t_n}} R_{z,n}^k$. This strategy has been proposed in [49], and it was shown numerically that the approximation to the quantity of interests (Y_{t_n}, Z_{t_n}) admits a kth order convergence, even if the Euler method is used to solve the forward SDE.

3.2. The integration discretization method. We now present the integration discretization method. Let $t = t_{n+k}$ in (3.3) and (3.5); then the integrals therein can be represented as

$$
\int_{t_n}^{t_{n+k}} \mathbb{E}_{t_n}^x [f_s] ds = \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_{k_y,i}^{n,k} \mathbb{E}_{t_n}^x \left[f_{t_{n+i}} \right] + R_y^{n,k}, \tag{3.10a}
$$

$$
\int_{t_n}^{t_{n+k}} \mathbb{E}_{t_n}^x [f_s G_{t_n,s}] ds = \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_{k_f,i}^{n,k} \mathbb{E}_{t_n}^x \left[f_{t_{n+i}} G^{n,i} \right] + R_{zf}^{n,k}, \tag{3.10b}
$$

$$
\int_{t_n}^{t_{n+k}} \mathbb{E}_{t_n}^x [Z_s \phi_s] ds = \Delta t_{n,k} \sum_{i=0}^{k_z} \lambda_{k_z,i}^{n,k} \mathbb{E}_{t_n}^x [Z_{t_{n+i}} \phi_{t_{n+i}}] + R_z^{n,k}, \tag{3.10c}
$$

where $\Delta t_{n,k} = t_{n+k} - t_n$, $\{\beta_{k_y,i}^{n,k}\}_{0 \le i \le k_y}$, $\{\lambda_{k_f,i}^{k}\}_{1 \le i \le k_f}$, and $\{\lambda_{k_z,i}^{n,k}\}_{0 \le i \le k_z}$ are real numbers, and $R_y^{n,k}$, $R_{zf}^{n,k}$, and $R_z^{n,k}$ are the associated truncation error terms. The constant k_y presents the number of grid points $(t_{n+i}, \mathbb{E}_{t_n}^x[f_{t_{n+i}}])$, $i = 0, 1, \ldots, k_y$, that is used to approximate the integral $\int_{t_n}^{t_{n+k}} \mathbb{E}_{t_n}^x [f_s] ds$, and the constants k_f and k_z have similar meanings as k_y. Notice that specific choices of those parameters will result in different approximation schemes such as the Euler scheme, the trapezoidal scheme, and multistep schemes (e.g., [42, 48, 53]).

Now, by inserting the above equations into (3.3) and (3.5), we get the following discretization equations for solving (Y_t, Z_t):

$$
\begin{cases} Y_{t_n} = \mathbb{E}_{t_n}^x \left[Y_{t_{n+k}} \right] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_{k_y,i}^{n,k} \mathbb{E}_{t_n}^x \left[f_{t_{n+i}} \right] + R_y^{n,k}, \\ \lambda_{k_z,0}^{n,k} \Delta t_{n,k} Z_{t_n} = \mathbb{E}_{t_n}^x \left[Y_{t_{n+k}} G^{n,k} \right] / \phi_{t_n} + \Delta t_{n,k} \sum_{i=1}^{k_f} \bar{\gamma}_{k_f,i}^{n,k} \mathbb{E}_{t_n}^x \left[f_{t_{n+i}} G^{n,i} \right] \\ \qquad\qquad - \Delta t_{n,k} \sum_{i=1}^{k_z} \bar{\lambda}_{k_z,i}^{n,k} \mathbb{E}_{t_n}^x [Z_{t_{n+i}} \phi_{t_{n+i}}] + \dfrac{1}{\phi_{t_n}} \left(R_{zf}^{n,k} - R_z^{n,k} \right), \end{cases} \tag{3.11}
$$

where $\bar{\gamma}_{k_f,i}^{n,k} = \gamma_{k_f,i}^{n,k}/\phi_{t_n}$ and $\bar{\lambda}_{k_z,i}^{n,k} = \lambda_{k_z,i}^{n,k}/\phi_{t_n}$.

To make the above equation for Z_{t_n} more general, we use the integration-by-parts formula in Lemma 2.2 to obtain

$$\begin{aligned}\mathbb{E}_{t_n}^x\left[Y_{t_{n+i}}G^{n,i}\right] &= \int_{t_n}^{t_{n+i}} \mathbb{E}_{t_n}^x\left[D_sY_{t_{n+i}}\phi_s ds\right]\\ &= \int_{t_n}^{t_{n+i}} \mathbb{E}_{t_n}^x\left[Z_{t_{n+i}}\sigma_{t_{n+i}}^{-1}D_sX_{t_{n+i}}\phi_s ds\right]\\ &= \Delta t_{n,i}\mathbb{E}_{t_n}^x\left[Z_{t_{n+i}}\phi_{t_{n+i}}\right] + \int_{t_n}^{t_{n+i}} \mathbb{E}_{t_n}^x\left[Z_{t_{n+i}}\left(\sigma_{t_{n+i}}^{-1}D_sX_{t_{n+i}}\phi_s - \phi_{t_{n+i}}\right)\right]ds,\end{aligned}$$

which yields

$$\begin{aligned}\mathbb{E}_{t_n}^x\left[Z_{t_{n+i}}\phi_{t_{n+i}}\right] = {}& \frac{1}{\Delta t_{n,i}}\mathbb{E}_{t_n}^x\left[Y_{t_{n+i}}G^{n,i}\right]\\ &+ \frac{1}{\Delta t_{n,i}}\int_{t_n}^{t_{n+i}} \mathbb{E}_{t_n}^x\left[Z_{t_{n+i}}\left(\phi_{t_{n+i}} - \sigma_{t_{n+i}}^{-1}D_sX_{t_{n+i}}\phi_s\right)\right]ds.\end{aligned} \tag{3.12}$$

Then, by inserting (3.10b), (3.10c), and (3.12) into (3.5) we obtain

$$\begin{aligned}0 = {}& \mathbb{E}_{t_n}^x\left[Y_{t_{n+k}}G^{n,k}\right] + \Delta t_{n,k}\sum_{i=1}^{k_f}\gamma_{k_f,i}^{n,k}\mathbb{E}_{t_n}^x\left[f_{t_{n+i}}G^{n,i}\right] - \Delta t_{n,k}\lambda_{k_z,0}^{n,k}Z_{t_n}\phi_{t_n}\\ &- \Delta t_{n,k}\sum_{i=1}^{k_z}\frac{\lambda_{k_z,i}^{n,k}}{\Delta t_{n,i}}\mathbb{E}_{t_n}^x\left[Z_{t_{n+i}}\int_{t_n}^{t_{n+i}}\left(\phi_{t_{n+i}} - \sigma_{t_{n+i}}^{-1}D_sX_{t_{n+i}}\phi_s\right)ds\right]\\ &- \Delta t_{n,k}\sum_{i=1}^{k_z}\frac{\lambda_{k_z,i}^{n,k}}{\Delta t_{n,i}}\mathbb{E}_{t_n}^x\left[Y_{t_{n+i}}G^{n,i}\right] + R_{zf}^{n,k} - R_z^{n,k}.\end{aligned}$$

Consequently, we get another more general equation for Z_{t_n}:

$$\begin{aligned}\lambda_{k_z,0}^{n,k}\Delta t_{n,k}Z_{t_n} = {}& \sum_{i=1}^{k}\tilde{\tilde{\lambda}}_{k_z,i}^{n,k}\mathbb{E}_{t_n}^x\left[Y_{t_{n+i}}G^{n,i}\right] + \Delta t_{n,k}\sum_{i=1}^{k_f}\bar{\gamma}_{k_f,i}^{n,k}\mathbb{E}_{t_n}^x\left[f_{t_{n+i}}G^{n,i}\right]\\ &- \Delta t_{n,k}\sum_{i=1}^{k_z}\bar{\lambda}_{k_z,i}^{n,k}\mathbb{E}_{t_n}^x\left[Z_{t_{n+i}}H_{t_n}^{n,i}\right] + \frac{1}{\phi_{t_n}}\left(R_{zf}^{n,k} - R_z^{n,k}\right),\end{aligned} \tag{3.13}$$

where

$$H_{t_n}^{n,i} = \frac{1}{\Delta t_{n,i}}\int_{t_n}^{t_{n+i}}\left(\phi_{t_{n+i}} - \sigma_{t_{n+i}}^{-1}D_sX_{t_{n+i}}\phi_s\right)ds,$$

and $\tilde{\tilde{\lambda}}_{k_z,i}^{n,k}$ is defined as

$$\tilde{\tilde{\lambda}}_{k_z,i}^{n,k} = \begin{cases} -\frac{\Delta t_{n,k}}{\Delta t_{n,i}}\frac{\lambda_{k_z,i}^{n,k}}{\phi_{t_n}}, & i = 1,\dots,k_z,\\ \frac{1}{\phi_{t_n}}, & i = k,\\ 0, & \text{others,}\end{cases}$$

if $k_z < k$, and otherwise

$$\tilde{\tilde{\lambda}}_{k_z,i}^{n,k} = \begin{cases} -\frac{\Delta t_{n,k}}{\Delta t_{n,i}}\frac{\lambda_{k_z,i}^{n,k}}{\phi_{t_n}}, & i \neq k,\\ \frac{1}{\phi_{t_n}}\left(1 - \frac{\Delta t_{n,k}}{\Delta t_{n,i}}\lambda_{k_z,i}^{n,k}\right), & i = k.\end{cases}$$

Notice that the above equation for Z_{t_n} is more general than the one in (3.11). In particular, (3.13) allows the use of a more general expansion of the multistep information (see the difference on the right-hand side).

Again, by removing the truncation error terms, one can obtain numerical schemes for solving the FBSDEs.

3.3. A unified discretization scheme for FBSDEs. We now propose a general time-discretization framework by combining the above two methods.

In view of the discretization equations (3.9), (3.11), and (3.13), we can propose the following general equations for the solution Y_{t_n} and Z_{t_n} of BSDEs:

$$\begin{cases} \alpha_0^{n,k} Y_{t_n} = \sum_{i=1}^{k} \alpha_i^{n,k} \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+i}}\right] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_{k_y,i}^{n,k} \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}\right] + \mathcal{R}_y^{n,k}, \\ \lambda_{k_z,0}^{n,k} \Delta t_{n,k} Z_{t_n} = \sum_{i=1}^{k} \tilde{\alpha}_i^{n,k} \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+i}} G^{n,i}\right] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_{k_f,i}^{n,k} \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}} G^{n,i}\right] \\ \qquad - \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_{k_z,i}^{n,k} \mathbb{E}_{t_n}^{X^n}\left[Z_{t_{n+i}} H^{n,i}\right] + \mathcal{R}_z^{n,k}, \end{cases} \tag{3.14}$$

where $\alpha_i^{n,k}$, $\tilde{\alpha}_i^{n,k}$, $\beta_{k_y,i}^{n,k}$, $\lambda_{k_z,i}^{n,k}$, $\gamma_{k_f,i}^{n,k}$ are deterministic artificial parameters, $G^{n,i}$ and $H^{n,i}$ are stochastic processes, and $\mathcal{R}_y^{n,k}$ and $\mathcal{R}_z^{n,k}$ are truncation errors.

Notice that the above discretization equations are combinations of the derivative discretization (3.9) and the integral discretization (3.11) and (3.13), with general parameters. Taking the equation of Z_{t_n}, for example, if one chooses $\tilde{\alpha}_i^{n,k}$ coinciding with $\tilde{\bar{\lambda}}_{k_z,i}^{n,k}$, $\gamma_{k_f,i}^{n,k}$ coinciding with $\bar{\gamma}_{k_f,i}^{n,k}$, and $\lambda_{k_z,i}^{n,k}$ coinciding with $\bar{\lambda}_{k_z,i}^{n,k}$, then we recover (3.13). If we choose $\lambda_{k_z,0}^{n,k} = 1$, $\tilde{\alpha}_i^{n,k}/\Delta t_{n,k} = \tilde{\tilde{\alpha}}_i^{n,k}$ and set other parameters to be zero, then the second equation in (3.14) recovers the second equation in (3.9). Moreover, we expect that new choices of the parameters will possibly generate new discretization schemes.

We are left to propose discretization schemes for the forward SDEs in (2.1). To this end, we introduce the following general scheme:

$$X^{n+i} = X^n + \Phi\left(t_n, X^n, \Delta t_n, \Delta W_{n,i}, \xi_n^{n,i}\right), \quad 0 \le n \le N-k,\ 1 \le i \le k, \tag{3.15}$$

where $X^0 = X_0$, and $\xi_n^{n,i}$ is a random variable related to the increment $\Delta W_{n,i}$. Here X^{n+i} can be viewed as a numerical approximation of the exact solution X_t to the forward SDE at $t = t_{n+i}$, which is calculated forwardly based on X^n. The value of X^{n+i} is involved in (3.16b) and (3.16c) and thus will be involved in the numerical analysis, e.g., (4.4) and (4.35). Notice that this general scheme contains the well-known Itô–Taylor type schemes, for instance [30],

1. the Euler scheme,

$$X^{n+1} = X^n + b(t_n, X^n)\Delta t_n + \sigma(t_n, X^n)\Delta W_{n,1};$$

2. the Milstein scheme,

$$\begin{aligned} X^{n+1} = {} & X^n + b(t_n, X^n)\Delta t_n + \sigma(t_n, X^n)\Delta W_{n,1} \\ & + \frac{1}{2}\sigma(t_n, X^n)\sigma_x(t_n, X^n)\left((\Delta W_{n,1})^2 - \Delta t_n\right); \end{aligned}$$

3. the weak order-2 Itô–Taylor scheme,

$$\begin{aligned} X^{n+1} = {} & X^n + b(t_n, X^n)\Delta t_n + \sigma(t_n, X^n)\Delta W_{n,1} + \frac{1}{2}L^0 b(t_n, X^n)(\Delta t_n)^2 \\ & + \frac{1}{2}\left(L^1 b(t_n, X^n) + L^0\sigma(t_n, X^n)\right)\Delta t_n \Delta W_{n,1} \\ & + \frac{1}{2}L^1\sigma(t_n, X^n)\left((\Delta W_{n,1})^2 - \Delta t_n\right), \end{aligned}$$

where the operators L^0 and L^1 are defined by

$$L^0 = \frac{\partial}{\partial t} + b(t,x)\frac{\partial}{\partial x} + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2}{\partial x^2}, \quad L^1 = \sigma(t,x)\frac{\partial}{\partial x}.$$

We are now ready to present a unified discretization scheme (by removing truncation error terms) for FBSDEs by combining the scheme (3.14) for BSDEs and the forward SDE discretization (3.15).

SCHEME 3.1. *Let* $K = \max\{k, k_y, k_f, k_z\}$*. Given* X^0 *and* $\{(Y^{N-i}, Z^{N-i})\}_{0\le i\le K-1}$*, we solve* $\{X^n\}_{1\le n\le N}$ *and* $\{(Y^n, Z^n)\}_{0\le n\le N-K}$ *by*

$$X^{n+i} = X^n + \Phi(t_n, X^n, \Delta t_n, \Delta W_{n,i}, \xi_n^{n,i}), \tag{3.16a}$$

$$\alpha_0^{n,k} Y^n = \sum_{i=1}^{k} \alpha_i^{n,k} \mathbb{E}_{t_n}^{X^n}\left[Y^{n+i}\right] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_{k_y,i}^{n,k} \mathbb{E}_{t_n}^{X^n}\left[f^{n+i}\right], \tag{3.16b}$$

$$\begin{aligned} \lambda_{k_z,0}^{n,k} \Delta t_{n,k} Z^n = &\sum_{i=1}^{k} \tilde{\alpha}_i^{n,k} \mathbb{E}_{t_n}^{X^n}\left[Y^{n+i} G^{n,i}\right] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_{k_f,i}^{n,k} \mathbb{E}_{t_n}^{X^n}\left[f^{n+i} G^{n,i}\right] \\ &- \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_{k_z,i}^{n,k} \mathbb{E}_{t_n}^{X^n}\left[Z^{n+i} H^{n,i}\right]. \end{aligned} \tag{3.16c}$$

Here $\{\alpha_i^{n,k}\}_{i=0}^{k}$, $\{\tilde{\alpha}_i^{n,k}\}_{i=1}^{k}$, $\{\beta_{k_y,i}^{n,k}\}_{i=0}^{k_y}$, $\{\gamma_{k_f,i}^{n,k}\}_{i=1}^{k_f}$, and $\{\lambda_{k_z,i}^{n,k}\}_{i=0}^{k_z}$ are deterministic real numbers (parameters) that may depend on the mesh, $G^{n,i}$ and $H^{n,i}$ are stochastic processes, and

$$f^{n+i} := f\left(t_{n+i}, X^{n+i}, Y^{n+i}, Z^{n+i}\right).$$

In what follows, for notational simplicity we shall omit the superscript ^{n,k} in those parameters unless possible confusions may rise. Then Scheme 3.1 becomes as follows.

SCHEME 3.2. *Let* $K = \max\{k, k_y, k_f, k_z\}$*. Given* X^0 *and* $\{(Y^{N-i}, Z^{N-i})\}_{0\le i\le K-1}$*, we solve* $\{X^n\}_{1\le n\le N}$ *and* $\{(Y^n, Z^n)\}_{0\le n\le N-K}$ *via*

$$X^{n+i} = X^n + \Phi\left(t_n, X^n, \Delta t_n, \Delta W_{n,i}, \xi_n^{n,i}\right), \tag{3.17a}$$

$$\alpha_0 Y^n = \sum_{i=1}^{k} \alpha_i \mathbb{E}_{t_n}^{X^n}\left[Y^{n+i}\right] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}_{t_n}^{X^n}\left[f^{n+i}\right], \tag{3.17b}$$

$$\begin{aligned} \lambda_0 \Delta t_{n,k} Z^n = &\sum_{i=1}^{k} \tilde{\alpha}_i \mathbb{E}_{t_n}^{X^n}\left[Y^{n+i} G^{n,i}\right] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}_{t_n}^{X^n}\left[f^{n+i} G^{n,i}\right] \\ &- \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}_{t_n}^{X^n}\left[Z^{n+i} H^{n,i}\right]. \end{aligned} \tag{3.17c}$$

Here the parameters $\{\alpha_i\}_{i=0}^{k}$, $\{\tilde{\alpha}_i\}_{i=1}^{k}$, $\{\beta_i\}_{i=0}^{k_y}$, $\{\gamma_i\}_{i=1}^{k_f}$, and $\{\lambda_i\}_{i=0}^{k_z}$ are real numbers with the constraint $\alpha_0 \neq 0$ and $\lambda_0 \neq 0$.

It is noticed that by properly choosing the parameters $\alpha_i, \tilde{\alpha}_i, \beta_i, \gamma_i$, and λ_i, the above unified scheme covers most of the existing discretization schemes; please see detailed discussions in section 5.

We remark that the $H^{n,i}$ in Scheme 3.1 is a generic stochastic process, and it is different from that in (3.13). We also remark that the above scheme is a temporal semidiscrete scheme. To get a fully discrete scheme, one should resort to spatial discretizations, and this can be done via, for instance, grid-type methods or the regression approach (see, e.g., [22, 26] and references therein).

4. Numerical analysis. In this section, we shall perform stability, consistency, and convergence analysis for Scheme 3.2. To this end, we first present some definitions.

Stability. In stability theory, one investigates small perturbations on terminal conditions and on the generator of BSDE, and then quantifies the stability of the numerical scheme due to these perturbations. To this end, we impose perturbations $\{(\zeta_y^{N-i}, \zeta_z^{N-i})\}_{0\le i\le k-1}$ on the terminal values $\{(Y^{N-i}, Z^{N-i})\}_{0\le i\le k-1}$. We also set a perturbation ζ_f on the generator $f(t,x,y,z)$:

$$f_\varsigma(t,x,y,z) = f(t,x,y,z) + \zeta_f.$$

The perturbed numerical solution pair $\{(Y_\varsigma^n, Z_\varsigma^n)\}_{0\le n\le N-k}$ under the perturbed conditions satisfies

$$\alpha_0 Y_\varsigma^n = \sum_{i=1}^{k} \alpha_i \mathbb{E}_{t_n}^{X^n}\left[Y_\varsigma^{n+i}\right] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}_{t_n}^{X^n}\left[f_\varsigma^{n+i}\right], \tag{4.1a}$$

$$\begin{aligned}\lambda_0 \Delta t_{n,k} Z_\varsigma^n = &\sum_{i=1}^{k} \tilde{\alpha}_i \mathbb{E}_{t_n}^{X^n}\left[Y_\varsigma^{n+i} G^{n,i}\right] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}_{t_n}^{X^n}\left[f_\varsigma^{n+i} G^{n,i}\right] \\ &- \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}_{t_n}^{X^n}\left[Z_\varsigma^{n+i} H^{n,i}\right],\end{aligned} \tag{4.1b}$$

where

$$f_\varsigma^n = f(t_n, X^n, Y_\varsigma^n, Z_\varsigma^n) + \zeta_f^n.$$

DEFINITION 4.1 (Stability). *Consider the numerical solution (Y^n, Z^n) of Scheme 3.2 and the above perturbed solution $(Y_\varsigma^n, Z_\varsigma^n)$. We say Scheme 3.2 is stable if for sufficiently small Δt, it holds that $(0 \le n \le N-k)$*

$$\begin{aligned}&\mathbb{E}\left[|Y_\varsigma^n - Y^n|^2\right] + \Delta t \sum_{j=n}^{N-k} \mathbb{E}\left[|Z_\varsigma^j - Z^j|^2\right] \\ &\le C\left(\sum_{j=0}^{k-1}\left(\mathbb{E}\left[\left|\zeta_y^{N-j}\right|^2\right] + \mathbb{E}\left[\left|\zeta_z^{N-j}\right|^2\right]\right) + \sum_{j=n}^{N-k} \Delta t \mathbb{E}\left[\left|\zeta_f^j\right|^2\right]\right),\end{aligned} \tag{4.2}$$

where C is a positive constant independent of Δt, Y_ζ^n, and Y^n.

Consistency. To introduce the consistency, we set

$$R_y^n = R_{yy}^n + R_{xy}^n, \quad R_z^n = R_{zz}^n + R_{xz}^n, \tag{4.3}$$

where the terms R_{yy}^n, R_{xy}^n, R_{zz}^n, and R_{xz}^n are, respectively, defined by

$$R_{yy}^n = \alpha_0 Y_{t_n}^{t_n, X^n} - \sum_{i=1}^{k} \alpha_i \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+i}}^{t_n,X^n}\right] - \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}^{t_n,X^n}\right],$$

$$R_{xy}^n = \sum_{i=1}^{k} \alpha_i \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+i}}^{t_n,X^n} - Y_{t_{n+i}}^{t_{n+i},X^{n+i}}\right] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}^{t_n,X^n} - f_{t_{n+i}}^{t_{n+i},X^{n+i}}\right],$$

(4.4)

$$
\begin{aligned}
R_{zz}^n = {} & \lambda_0 \Delta t_{n,k} Z_{t_n}^{t_n,X^n} - \sum_{i=1}^{k} \tilde{\alpha}_i \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+i}}^{t_n,X^n} G^{n,i}\right] - \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}^{t_n,X^n} G^{n,i}\right] \\
& - \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}_{t_n}^{X^n}\left[Z_{t_{n+i}}^{t_n,X^n} H^{n,i}\right], \\
R_{xz}^n = {} & \sum_{i=1}^{k} \tilde{\alpha}_i \mathbb{E}_{t_n}^{X^n}\left[\left(Y_{t_{n+i}}^{t_n,X^n} - Y_{t_{n+i}}^{t_{n+i},X^{n+i}}\right) G^{n,i}\right] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}_{t_n}^{X^n}\left[\left(f_{t_{n+i}}^{t_n,X^n}\right.\right. \\
& \left.\left. - f_{t_{n+i}}^{t_{n+i},X^{n+i}}\right) G^{n,i}\right] - \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}_{t_n}^{X^n}\left[\left(Z_{t_{n+i}}^{t_n,X^n} - Z_{t_{n+i}}^{t_{n+i},X^{n+i}}\right) H^{n,i}\right],
\end{aligned}
$$

where

$$
\begin{aligned}
f_{t_{n+i}}^{t_n,X^n} &= f\left(t_{n+1}, X^{n+i}, Y_{t_{n+i}}^{t_n,X^n}, Z_{t_{n+i}}^{t_n,X^n}\right), \\
f_{t_{n+i}}^{t_{n+i},X^{n+i}} &= f\left(t_{n+i}, X^{n+i}, Y_{t_{n+i}}^{t_{n+i},X^{n+i}}, Z_{t_{n+1}}^{t_{n+i},X^{n+i}}\right).
\end{aligned}
$$

Here the numerical solutions X^n are taken into account, and the superscript ${}^{t_n,X^n}$ means that the associated processes start at $t = t_n$ from the space point X^n.

Remark 1. The accuracy of Scheme 3.2 depends not only on the accuracy of the discretizations (3.17b) and (3.17c) for solving the BSDE but also on that of the discretization (3.17a) for solving the SDE. Here the terms R_y^n and R_z^n represent the total errors for solving solution Y_t and Z_t, respectively. R_{yy}^n and R_{zz}^n mean the errors from the integration or differentiation discretization; R_{xy}^n and R_{xz}^n mean the errors from the approximation accuracy of the discretization (3.17a).

DEFINITION 4.2 (Consistency). *Scheme* 3.2 *is said to be consistent in the mean square sense if for any* $\epsilon > 0$, *there exists* $\delta > 0$ *such that for* $0 < \Delta t < \delta$, *it holds that*

$$
\sum_{n=0}^{N-k} \frac{1}{\Delta t} \mathbb{E}\left[|R_y^n|^2\right] < \epsilon \quad \text{and} \quad \sum_{n=0}^{N-k} \frac{1}{\Delta t} \mathbb{E}\left[|R_z^n|^2\right] < \epsilon. \tag{4.5}
$$

Furthermore, it is said to be consistent with order p *if*

$$
\mathbb{E}\left[|R_y^n|^2\right] = \mathcal{O}\left((\Delta t)^{2(1+p)}\right) \quad \text{and} \quad \mathbb{E}\left[|R_z^n|^2\right] = \mathcal{O}\left((\Delta t)^{2(1+p)}\right). \tag{4.6}
$$

Convergence. Let (Y_{t_n}, Z_{t_n}) and (Y^n, Z^n) be the exact solution of FBSDEs (2.1) and the numerical solution of Scheme 3.2, respectively. The numerical errors $\{e_y^n\}_{n=0}^N$ and $\{e_z^n\}_{n=0}^N$ are defined by

$$
e_y^n = Y_{t_n}^{t_n,X^n} - Y^n, \quad e_z^n = Z_{t_n}^{t_n,X^n} - Z^n.
$$

DEFINITION 4.3 (Convergence). *Scheme* 3.2 *is said to be convergent in the mean square sense if for any* $\epsilon > 0$ *there exists* $\delta > 0$ *such that for* $0 < \Delta t < \delta$ *it holds that* $(0 \le n \le N - k)$

$$
\mathbb{E}\left[|e_y^n|^2\right] + \Delta t \sum_{j=n}^{N-k} \mathbb{E}\left[|e_z^j|^2\right] < \epsilon. \tag{4.7}
$$

4.1. Lax equivalence theorem. In classical numerical ODEs/PDEs theory, the consistency and stability of a numerical scheme in general imply the convergence, and this is known as the Lax equivalence convergence theorem. In this section, we shall present a stochastic version of the Lax equivalence convergence theorem.

THEOREM 1 (Lax equivalence Theorem). *Suppose that Scheme* 3.2 *is consistent. Then it is convergent in the mean square sense if and only if it is stable.*

Proof. First, we assume that Scheme 3.2 is stable and consistent. Let $e_y^n = Y_{t_n} - Y^n$ and $e_z^n = Z_{t_n} - Z^n$. The aim is to show that Scheme 3.2 is convergent, i.e., for sufficiently small Δt it holds that

$$\mathbb{E}\big[|e_y^n|^2\big] + \Delta t \sum_{j=n}^{N-k} \mathbb{E}\big[|e_z^j|^2\big] < \epsilon, \quad n = 0, 1, \ldots, N-k.$$

Observe that the exact solution (Y_{t_n}, Z_{t_n}) satisfies the following equations:

(4.8a)
$$\alpha_0 Y_{t_n}^{t_n,X^n} = \sum_{i=1}^{k} \alpha_i \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+i}}^{t_{n+i},X^{n+i}}\right] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}^{t_{n+i},X^{n+i}}\right] + R_y^n,$$

(4.8b)
$$\begin{aligned}\lambda_0 \Delta t_{n,k} Z_{t_n}^{t_n,X^n} &= \sum_{i=1}^{k} \tilde{\alpha}_i \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+i}}^{t_{n+i},X^{n+i}} G^{n,i}\right] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}^{t_{n+i},X^{n+i}} G^{n,i}\right] \\ &\quad - \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}_{t_n}^{X^n}\left[Z_{t_{n+i}}^{t_{n+i},X^{n+i}} H^{n,i}\right] + R_z^n,\end{aligned}$$

where R_y^n and R_z^n are defined by (4.3). Then e_y^n and e_z^n satisfy the following equations:

$$\alpha_0 e_y^n = \sum_{i=1}^{k} \alpha_i \mathbb{E}_{t_n}^{X^n}[e_y^{n+i}] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_r \mathbb{E}_{t_n}^{X^n}\left[e_f^{n+i}\right] + R_y^n, \tag{4.9a}$$

$$\begin{aligned}\lambda_0 \Delta t_{n,k} e_z^n &= \sum_{i=1}^{k} \tilde{\alpha}_i \mathbb{E}_{t_n}^{X^n}[e_y^{n+i} G^{n,i}] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}_{t_n}^{X^n}[e_f^{n+i} G^{n,i}] \\ &\quad - \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}_{t_n}^{X^n}[e_z^{n+i} H^{n,i}] + R_z^n.\end{aligned} \tag{4.9b}$$

Notice that R_y^n and R_z^n can be viewed as perturbations of Scheme 3.2. Thus the stability assumption leads to the following estimate:

$$\mathbb{E}\left[|e_y^n|^2\right] + \Delta t \sum_{i=n}^{N-k} \mathbb{E}\left[|e_z^i|^2\right] \le C\left(\sum_{i=0}^{k-1} \mathbb{E}\Big[\big|e_y^{N-i}\big|^2\Big] + \frac{1}{\Delta t}\sum_{i=n}^{N-k} \mathbb{E}\Big[\big|R_y^i\big|^2 + \big|R_z^i\big|^2\Big]\right).$$

By the consistency assumption, for sufficiently small Δt it holds that

$$\frac{1}{\Delta t}\sum_{i=n}^{N-k} \mathbb{E}\left[|R_y^i|^2 + |R_z^i|^2\right] < \frac{\epsilon}{2}. \tag{4.10}$$

We also suppose that the terminal conditions are given with the following prior estimate:

$$\mathbb{E}\left[\left|e_y^{N-i}\right|^2\right] + \mathbb{E}\left[\left|e_z^{N-i}\right|^2\right] < \frac{\epsilon}{2k}, \quad i = 0, 1, \ldots, k-1. \tag{4.11}$$

Then, we get by using together (4.10) and (4.11)

$$\mathbb{E}\left[\left|e_y^n\right|^2\right] + \Delta t \sum_{i=n}^{N-k} \mathbb{E}\left[\left|e_z^i\right|^2\right] \le C\left(\frac{\epsilon}{2} + \frac{\epsilon}{2}\right) < C\epsilon. \tag{4.12}$$

That is, the scheme is convergent.

Next, we assume that the scheme is convergent. We prove the stability by the contradiction argument. To this end, suppose that the scheme is unstable; then there exists $\varepsilon_0 > 0$ such that for any $\delta > 0$ and at least one $\Delta t \in (0, \delta]$ it holds that

$$\mathbb{E}\left[|e_y^n|^2\right] + \Delta t \sum_{i=n}^{N-k} \mathbb{E}\left[|e_z^i|^2\right] \ge \varepsilon_0, \quad n = 0, 1, \ldots, N-k. \tag{4.13}$$

This means that the solutions Y^n and Z^n of the scheme do not converge to the solutions Y_t and Z_t, which contradicts the convergence assumption. The proof is complete. ∎

We expect that the above Lax equivalence theorem will play an essential role in analyzing numerical schemes for FBSDEs. We will discuss some of its applications in section 5.

4.2. Stability analysis. In this section, we aim at analyzing the stability of the proposed scheme. In particular, the analysis is narrowed to the following simplified version of Scheme 3.2.

SCHEME 4.4. *Let* $K = \max\{k, k_y, k_f, k_z\}$. *Given* X^0 *and* $\{(Y^{N-i}, Z^{N-i})\}_{0\le i\le K-1}$, *we solve* $\{X^n\}_{1\le n\le N}$ *and* $\{(Y^n, Z^n)\}_{0\le n\le N-K}$ *by*

$$X^{n+i} = X^n + \Phi(t_n, X^n, \Delta t_n, \Delta W_{n,i}, \xi_n^{n,i}), \tag{4.14a}$$

$$Y^n = \sum_{i=1}^{k} \alpha_i \mathbb{E}_{t_n}^{X^n}\left[Y^{n+i}\right] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}_{t_n}^{X^n}\left[f^{n+i}\right], \tag{4.14b}$$

$$\begin{aligned}\lambda_0 \Delta t_{n,k} Z^n = {}& \tilde{\alpha}\mathbb{E}_{t_n}^{X^n}\left[Y^{n+k}G^{n,k}\right] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}_{t_n}^{X^n}\left[f^{n+i}G^{n,i}\right] \\ & - \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}_{t_n}^{X^n}\left[Z^{n+i}H^{n,i}\right].\end{aligned} \tag{4.14c}$$

The only difference between the above scheme and the original Scheme 3.2 is on the Z-equation, and this scheme is simplified by setting the first term in the right-hand side as $\tilde{\alpha}\mathbb{E}_{t_n}^{X^n}[Y^{n+k}G^{n,k}]$ instead of $\sum_{i=1}^{k} \tilde{\alpha}_i \mathbb{E}_{t_n}^{X^n}\left[Y^{n+i}G^{n,i}\right]$. Nevertheless, the above simplified scheme still covers most of existing schemes such as those mentioned in Table 5.1. To study the stability, consistency, and convergence of Scheme 4.4, we shall make some additional assumptions as listed below.

Assumption 2. We assume that the parameters in Scheme 4.4 satisfy

$$\sum_{i=1}^{k} |\alpha_i| \le (1 + C_*\Delta t), \quad 0 < \tilde{\alpha}^2 \le (1 + C_*'\Delta t), \quad \frac{\tilde{\alpha}^2}{|\alpha_k|} \le M \text{ if } \alpha_k \ne 0,$$

$\lambda_0 > |\lambda_1|$ and $\max_i \{|\beta_i|, |\gamma_i|, |\lambda_i|\} \leq \Upsilon$, where C_*, C'_*, M, and Υ are positive constants. Furthermore, we assume that the measurable process $H^{n,k}$ satisfies

$$H^{n,1} = 1 \text{ or } \mathbb{E}[H^{n,1}] = \mathcal{O}\left((\Delta t)^\alpha\right), \quad \mathbb{E}\left[|H^{n,k}|^2\right] = \mathcal{O}\left((\Delta t)^\alpha\right), \ k > 1,$$

where $\alpha > 0$.

It can be verified that for most existing probabilistic schemes (e.g., [4, 7, 14, 26, 32, 43, 47, 48, 54]), the involved parameters satisfy the above assumption. Now we are ready to give the main result of this section.

THEOREM 2. *Under Assumptions* 1–2 *and* (3.4)*, Scheme* 4.4 *is stable.*

Proof. Let (Y^n, Z^n) and $(Y^n_\varsigma, Z^n_\varsigma)$ be the numerical solution and the perturbed solution, respectively. We have

$$Y^n_\varsigma = \sum_{i=1}^{k} \alpha_i \mathbb{E}^{X^n}_{t_n}[Y^{n+i}_\varsigma] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}^{X^n}_{t_n}[f^{n+i}_\varsigma], \tag{4.15a}$$

$$\lambda_0 \Delta t_{n,k} Z^n_\varsigma = \tilde{\alpha} \mathbb{E}^{X^n}_{t_n}[Y^{n+k}_\varsigma G^{n,k}] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}^{X^n}_{t_n}[f^{n+i}_\varsigma G^{n,i}]$$
$$- \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}^{X^n}_{t_n}[Z^{n+i}_\varsigma H^{n,i}]. \tag{4.15b}$$

For notational simplicity, we set

$$e^n_{y,\varsigma} = Y^n_\varsigma - Y^n, \ \ e^n_{z,\varsigma} = Z^n_\varsigma - Z^n, \ \ e^n_{f,\varsigma} = f_\varsigma(t_n, X^n, Y^n_\varsigma, Z^n_\varsigma) - f(t_n, X^n, Y^n, Z^n).$$

Step 1. Estimate $e^n_{y,\varsigma} = Y^n_\varsigma - Y^n$. By subtracting (4.14b) from (4.15a) we obtain the perturbed error equation:

$$e^n_{y,\varsigma} = \sum_{i=1}^{k} \alpha_i \mathbb{E}^{X^n}_{t_n}\left[e^{n+i}_{y,\varsigma}\right] + \Delta t_{n,k} \sum_{r=0}^{k_y} \beta_r \mathbb{E}^{X^n}_{t_n}\left[e^{n+r}_{f,\varsigma}\right] + T^n_{y,\zeta}, \tag{4.16}$$

where the residual term $T^n_{y,\zeta}$ is defined by (recall the definition for ζ after (4.1))

$$T^n_{y,\zeta} = \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}^{X^n}_{t_n}[\zeta^{n+i}_f].$$

Then by the Lipschitz property of f we have

$$|e^n_{y,\varsigma}| \leq \left| \sum_{i=1}^{k} \alpha_i \mathbb{E}^{X^n}_{t_n}[e^{n+i}_{y,\varsigma}] \right| + \Delta t_{n,k} L \sum_{r=0}^{k_y} |\beta_r| \mathbb{E}^{X^n}_{t_n}\left[|e^{n+r}_{y,\varsigma}| + |e^{n+r}_{z,\varsigma}|\right] + |T^n_{y,\zeta}|. \tag{4.17}$$

By taking the square on both sides of (4.17) and by using together the inequality $(a+b)^2 \leq (1+\eta\Delta t)a^2 + (1+\frac{1}{\eta\Delta t})b^2$, the Hölder inequality, and Assumption 2 we deduce

(4.18)

$$\left|e^n_{y,\varsigma}\right|^2 \leq (1+\eta\Delta t) \left| \sum_{i=1}^{k} \alpha_i \mathbb{E}^{X^n}_{t_n}[e^{n+i}_{y,\varsigma}] \right|^2$$
$$+ \left(1 + \frac{1}{\eta\Delta t}\right) \left(\Delta t_{n,k} L \sum_{r=0}^{k_y} |\beta_r| \mathbb{E}^{X^n}_{t_n}\left[|e^{n+r}_{y,\varsigma}| + |e^{n+r}_{z,\varsigma}|\right] + |T^n_{y,\zeta}| \right)^2$$

$$
\begin{aligned}
&\leq (1+\eta\Delta t)\sum_{i=1}^{k}|\alpha_i|\sum_{i=1}^{k}|\alpha_i|\left|\mathbb{E}_{t_n}^{X^n}[e_{y,\varsigma}^{n+i}]\right|^2+(2k_y+3)\left(1+\frac{1}{\eta\Delta t}\right)\\
&\quad\times\left((\Delta t_{n,k})^2L^2\sum_{r=0}^{k_y}|\beta_r|^2\mathbb{E}_{t_n}^{X^n}\left[|e_{y,\varsigma}^{n+r}|^2+|e_{z,\varsigma}^{n+r}|^2\right]+|T_{y,\zeta}^n|^2\right)\\
&\leq (1+\eta\Delta t)(1+C_*\Delta t)\sum_{i=1}^{k}|\alpha_i|\left|\mathbb{E}_{t_n}^{X^n}[e_{y,\varsigma}^{n+i}]\right|^2\\
&\quad+(2k_y+3)\left(1+\frac{1}{\eta\Delta t}\right)\left(k^2L^2\Upsilon^2(\Delta t)^2\sum_{r=0}^{k_y}\mathbb{E}_{t_n}^{X^n}\left[|e_{y,\varsigma}^{n+r}|^2+|e_{z,\varsigma}^{n+r}|^2\right]+|T_{y,\zeta}^n|^2\right).
\end{aligned}
$$

Step 2. Estimate $e_{z,\varsigma}^n = Z_\varsigma^n - Z^n$. Similarly, by subtracting (4.14c) from (4.15b) we get

$$
\begin{aligned}
\lambda_0\Delta t_{n,k}e_{z,\varsigma}^n &= \tilde{\alpha}\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}G^{n,k}\right]+\Delta t_{n,k}\sum_{i=1}^{k_f}\gamma_i\mathbb{E}_{t_n}^{X^n}\left[e_{f,\varsigma}^{n+i}G^{n,i}\right]\\
&\quad-\Delta t_{n,k}\sum_{i=1}^{k_z}\lambda_i\mathbb{E}_{t_n}^{X^n}\left[e_{z,\varsigma}^{n+i}H^{n,i}\right]+T_{z,\zeta}^n,
\end{aligned}
\tag{4.19}
$$

where the residual term $T_{z,\zeta}^n$ is defined by

$$
T_{z,\zeta}^n = \Delta t_{n,k}\sum_{i=1}^{k_f}\gamma_i\mathbb{E}_{t_n}^{X^n}\left[\zeta_f^{n+i}G^{n,i}\right].
$$

By (4.19) and setting $\bar{\lambda}_i = \frac{\lambda_i}{\lambda_0}$, we can derive

$$
\begin{aligned}
|e_{z,\varsigma}^n| &\leq |\bar{\lambda}_1|\left|\mathbb{E}_{t_n}^{X^n}\left[e_{z,\varsigma}^{n+1}H^{n,1}\right]\right|+\sum_{i=2}^{k_z}|\bar{\lambda}_i|\left|\mathbb{E}_{t_n}^{X^n}[e_{z,\varsigma}^{n+i}H^{n,i}]\right|+\frac{1}{|\lambda_0\Delta t_{n,k}|}|\tilde{\alpha}|\left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}G^{n,k}\right]\right|\\
&\quad+\frac{1}{|\lambda_0|}\sum_{i=1}^{k_f}|\gamma_i|\left|\mathbb{E}_{t_n}^{X^n}\left[e_{f,\varsigma}^{n+i}G^{n,i}\right]\right|+\frac{1}{|\lambda_0|\Delta t_{n,k}}|T_{z,\zeta}^n|.
\end{aligned}
\tag{4.20}
$$

Then, by squaring both sides of (4.20) and using the inequality $(a+b)^2 \leq (1+\varepsilon)a^2 + (1+\frac{1}{\varepsilon})b^2$ we have

$$
\begin{aligned}
|e_{z,\varsigma}^n|^2 &\leq (1+\varepsilon)\left(|\bar{\lambda}_1|\left|\mathbb{E}_{t_n}^{X^n}\left[e_{z,\varsigma}^{n+1}H^{n,1}\right]\right|\right)^2+2(1+1/\varepsilon)\\
&\quad\times\left(\frac{\tilde{\alpha}^2}{(\lambda_0\Delta t_{n,k})^2}\left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}G^{n,k}\right]\right|^2+\left(\frac{1}{|\lambda_0|}\sum_{i=1}^{k_f}|\gamma_i|\left|\mathbb{E}_{t_n}^{X^n}\left[e_{f,\varsigma}^{n+i}G^{n,i}\right]\right|\right.\right.\\
&\quad\left.\left.+\sum_{i=2}^{k_z}|\bar{\lambda}_i|\left|\mathbb{E}_{t_n}^{X^n}\left[e_{z,\varsigma}^{n+i}H^{n,i}\right]\right|+\frac{1}{|\lambda_0\Delta t_{n,k}|}|T_{z,\zeta}^n|\right)^2\right)\\
&\leq (1+\varepsilon)\bar{\lambda}_1^2\left|\mathbb{E}_{t_n}^{X^n}\left[e_{z,\varsigma}^{n+1}H^{n,1}\right]\right|^2+2(1+1/\varepsilon)\frac{\tilde{\alpha}^2}{\lambda_0^2(\Delta t_{n,k})^2}\left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}G^{n,k}\right]\right|^2
\end{aligned}
\tag{4.21}
$$

$$+ 2\big(1+1/\varepsilon\big)(k_f+k_z)\left(\frac{1}{\lambda_0^2}\sum_{i=1}^{k_f}\gamma_i^2\mathbb{E}_{t_n}^{X^n}\left[\left|e_{f,\varsigma}^{n+i}\right|^2\right]\mathbb{E}_{t_n}^{X^n}\left[\left|G^{n,i}\right|^2\right]\right.$$
$$\left.+\sum_{i=2}^{k_z}\bar{\lambda}_i^2\mathbb{E}_{t_n}^{X^n}\left[|e_{z,\varsigma}^{n+i}|^2\right]\mathbb{E}_{t_n}^{X^n}\left[|H^{n,i}|^2\right]+\frac{1}{(\lambda_0\Delta t_{n,k})^2}\left|T_{z,\zeta}^n\right|^2\right).$$

By Assumptions 1–2, we have

$$\begin{aligned}
&\mathbb{E}_{t_n}^{X^n}\left[\left|G^{n,k}\right|^2\right]\le \Delta t_{n,k}\max_{t_n\le s\le t_{n_k}}\mathbb{E}[\phi_s^2]\le C_\phi\Delta t_{n,k},\\
&\mathbb{E}_{t_n}^{X^n}\left[\left|H^{n,1}\right|^2\right]\le 1,\quad \mathbb{E}_{t_n}^{X^n}\left[|H^{n,k}|^2\right]\le C(\Delta t_{n,k})^\alpha,\quad k=2,3,\dots,\\
&\mathbb{E}_{t_n}^{X^n}\left[\left|e_{f,\varsigma}^{n+i}\right|^2\right]\le L^2\mathbb{E}_{t_n}^{X^n}\left[\left(|e_{y,\varsigma}^{n+i}|+|e_{z,\varsigma}^{n+i}|\right)^2\right]\le 2L^2\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+i}\right|^2+\left|e_{z,\varsigma}^{n+i}\right|^2\right].
\end{aligned}\tag{4.22}$$

Also by Hölder's inequality we have

$$\begin{aligned}
\left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}G^{n,k}\right]\right|^2 &= \left|\mathbb{E}_{t_n}^{X^n}\left[\left(e_{y,\varsigma}^{n+k}-\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}\right]\right)G^{n,k}\right]\right|^2\\
&\le C_\phi\Delta t_{n,k}\left(\mathbb{E}_{t_n}^{X^n}\left[|e_{y,\varsigma}^{n+k}|^2\right]-\left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}\right]\right|^2\right).
\end{aligned}\tag{4.23}$$

Then by inserting (4.22) and (4.23) into (4.21), we get

$$\begin{aligned}
\left|e_{z,\varsigma}^n\right|^2 \le{}& (1+\varepsilon)\bar{\lambda}_1^2\mathbb{E}_{t_n}^{X^n}\left[\left|e_{z,\varsigma}^{n+1}\right|^2\right]\\
&+2C_\phi(1+1/\varepsilon)\frac{\tilde{\alpha}^2}{\lambda_0^2\Delta t_{n,k}}\left(\mathbb{E}_{t_n}^{X^n}\left[|e_{y,\varsigma}^{n+k}|^2\right]-\left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}\right]\right|^2\right)\\
&+2(1+1/\varepsilon)(k_f+k_z)\left(2C_\phi L^2/\lambda_0^2\sum_{i=1}^{k_f}\gamma_i^2\Delta t_{n,i}\mathbb{E}_{t_n}^{X^n}\left[|e_{y,\varsigma}^{n+i}|^2+|e_{z,\varsigma}^{n+i}|^2\right]\right.\\
&\left.+\sum_{i=2}^{k_z}\bar{\lambda}_i^2(\Delta t_{n,i})^\alpha\mathbb{E}_{t_n}^{X^n}\left[|e_{z,\varsigma}^{n+i}|^2\right]+\left|T_{z,\zeta}^n\right|^2/\lambda_0^2(\Delta t_{n,k})^2\right).
\end{aligned}\tag{4.24}$$

For $\alpha_k\neq 0$, it holds that $\frac{\tilde{\alpha}^2}{|\alpha_k|}\le M$, thus $\tilde{\alpha}^2=|\alpha_k|\frac{\tilde{\alpha}^2}{|\alpha_k|}\le M|\alpha_k|$. Let $C'=C_\phi/\lambda_0^2M$. Now multiplying both sides of (4.24) by $\frac{\Delta t}{2C'(1+\frac{1}{\varepsilon})}$ one has

$$\begin{aligned}
\frac{\Delta t}{2C'(1+\frac{1}{\varepsilon})}\left|e_{z,\varsigma}^n\right|^2 \le{}& \frac{\varepsilon}{2C'}\bar{\lambda}_1^2\Delta t\mathbb{E}_{t_n}^{X^n}\left[\left|e_{z,\varsigma}^{n+1}\right|^2\right]+|\alpha_k|\left(\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+k}\right|^2\right]-\left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}\right]\right|^2\right)\\
&+2C_\phi C_0'L^2\Upsilon^2(\Delta t)^2\sum_{i=1}^{k_f}\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+i}\right|^2+\left|e_{z,\varsigma}^{n+i}\right|^2\right]\\
&+C_0'(\Delta t)^{1+\alpha}\sum_{i=2}^{k_z}\lambda_i^2\mathbb{E}_{t_n}^{X^n}\left[\left|e_{z,\varsigma}^{n+i}\right|^2\right]+C_0'\frac{\left|T_{z,\zeta}^n\right|^2}{\Delta t},
\end{aligned}\tag{4.25}$$

where $C_0'=\frac{k_f+k_z}{C'\lambda_0^2}=M(k_f+k_z)/C_\phi$.

For the case $\alpha_k = 0$, we set $C' = C_\phi/\lambda_0^2$. Then, by multiplying both sides of (4.24) by $\frac{\Delta t}{2C'(1+\frac{1}{\varepsilon})}$ we obtain

$$\begin{aligned} \frac{\Delta t}{2C'(1+\frac{1}{\varepsilon})}\left|e_{z,\varsigma}^n\right|^2 \leq{}& \frac{\varepsilon}{2C'}\bar{\lambda}_1^2\Delta t\mathbb{E}_{t_n}^{X^n}\left[\left|e_{z,\varsigma}^{n+1}\right|^2\right] + \tilde{\alpha}^2\left(\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+k}\right|^2\right] - \left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}\right]\right|^2\right) \\ &+ 2C_\phi C_0' L^2\Upsilon^2(\Delta t)^2\sum_{i=1}^{k_f}\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+i}\right|^2 + \left|e_{z,\varsigma}^{n+i}\right|^2\right] \\ &+ C_0'(\Delta t)^{1+\alpha}\sum_{i=2}^{k_z}\lambda_i^2\mathbb{E}_{t_n}^{X^n}\left[\left|e_{z,\varsigma}^{n+i}\right|^2\right] + C_0'\frac{|T_{z,\zeta}^n|^2}{\Delta t}, \end{aligned} \tag{4.26}$$

where $C_0' = \frac{k_f+k_z}{C'\lambda_0^2} = (k_f+k_z)/C_\phi$.

Step 3. *Toward the stability claim.* We first consider the case where $\alpha_k \neq 0$. By summing up (4.18) and (4.25) we get

$$\begin{aligned} &|e_{y,\varsigma}^n|^2 + \frac{\varepsilon}{2C'(1+\varepsilon)}\Delta t|e_{z,\varsigma}^n|^2 \\ &\quad\leq (1+\eta\Delta t)(1+C_*\Delta t)\sum_{i=1}^{k}|\alpha_i|\left|\mathbb{E}_{t_n}^{X^n}[e_{y,\varsigma}^{n+i}]\right|^2 \\ &\qquad + |\alpha_k|\left(\mathbb{E}_{t_n}^{X^n}\left[|e_{y,\varsigma}^{n+k}|^2\right] - \left|\mathbb{E}_{t_n}^{X^n}[e_{y,\varsigma}^{n+k}]\right|^2\right) + \frac{\varepsilon}{2C'}\bar{\lambda}_1^2\Delta t\mathbb{E}_{t_n}^{X^n}\left[\left|e_{z,\varsigma}^{n+1}\right|^2\right] \\ &\qquad + (2k_y+3)\left(1+\frac{1}{\eta\Delta t}\right)\left(k^2L^2\Upsilon^2(\Delta t)^2\sum_{r=0}^{k_y}\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+r}\right|^2 + \left|e_{z,\varsigma}^{n+r}\right|^2\right] + \left|T_{y,\zeta}^n\right|^2\right) \\ &\qquad + 2C_\phi C_0' L^2\Upsilon^2(\Delta t)^2\sum_{i=1}^{k_f}\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+i}\right|^2 + \left|e_{z,\varsigma}^{n+i}\right|^2\right] \\ &\qquad + C_0'(\Delta t)^{1+\alpha}\sum_{i=2}^{k_z}\lambda_i^2\mathbb{E}_{t_n}^{X^n}\left[\left|e_{z,\varsigma}^{n+i}\right|^2\right] + C_0'\frac{\left|T_{z,\zeta}^n\right|^2}{\Delta t}. \end{aligned} \tag{4.27}$$

For the case of $\alpha_k = 0$, we can sum up (4.18) and (4.26) to get

$$\begin{aligned} &|e_{y,\varsigma}^n|^2 + \frac{\varepsilon}{2C'(1+\varepsilon)}\Delta t|e_{z,\varsigma}^n|^2 \\ &\quad\leq (1+\eta\Delta t)(1+C_*\Delta t)\sum_{i=1}^{k}|\alpha_i|\left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+i}\right]\right|^2 \\ &\qquad + \tilde{\alpha}^2\left(\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+k}\right|^2\right] - \left|\mathbb{E}_{t_n}^{X^n}\left[e_{y,\varsigma}^{n+k}\right]\right|^2\right) + \frac{\varepsilon}{2C'}\bar{\lambda}_1^2\Delta t\mathbb{E}_{t_n}^{X^n}\left[\left|e_{z,\varsigma}^{n+1}\right|^2\right] \\ &\qquad + (2k_y+3)\left(1+\frac{1}{\eta\Delta t}\right)\left(k^2L^2\Upsilon^2(\Delta t)^2\sum_{r=0}^{k_y}\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+r}\right|^2 + \left|e_{z,\varsigma}^{n+r}\right|^2\right] + \left|T_{y,\zeta}^n\right|^2\right) \\ &\qquad + 2C_\phi C_0' L^2\Upsilon^2(\Delta t)^2\sum_{i=1}^{k_f}\mathbb{E}_{t_n}^{X^n}\left[\left|e_{y,\varsigma}^{n+i}\right|^2 + \left|e_{z,\varsigma}^{n+i}\right|^2\right] \\ &\qquad + C_0'(\Delta t)^{1+\alpha}\sum_{i=2}^{k_z}\lambda_i^2\mathbb{E}_{t_n}^{X^n}\left[\left|e_{z,\varsigma}^{n+i}\right|^2\right] + C_0'\left|T_{z,\zeta}^n\right|^2/\Delta t. \end{aligned} \tag{4.28}$$

Notice that the above two cases $\alpha_k \neq 0$ and $\alpha_k = 0$ can be unified. Now recall that $K = \max\{k, k_y, k_f, k_z\}$. Then, by taking the expectation $\mathbb{E}[\cdot]$ on the both sides of (4.27) or (4.28), the derived equation can be further simplified into

$$
\begin{aligned}
&(1 - C_1'\Delta t)\,\mathbb{E}\left[\left|e_{y,\varsigma}^{n}\right|^2\right] + \left(\frac{\varepsilon}{2C'(1+\varepsilon)} - C_1'\right)\Delta t\mathbb{E}\left[\left|e_{z,\varsigma}^{n}\right|^2\right] \\
&\le (1 + C^*\Delta t)\sum_{i=1}^{k}|\alpha_i|\,\mathbb{E}\left[\left|e_{y,\varsigma}^{n+i}\right|^2\right] + \frac{\varepsilon}{2C'}\bar{\lambda}_1^2\Delta t\mathbb{E}\left[\left|e_{z,\varsigma}^{n+1}\right|^2\right] \\
&\quad + C_2'\Delta t\sum_{i=1}^{K}\mathbb{E}\left[\left|e_{y,\varsigma}^{n+i}\right|^2\right] + C_3'\Delta t\sum_{i=2}^{K}\mathbb{E}\left[\left|e_{z,\varsigma}^{n+i}\right|^2\right] + \frac{C_4'}{\Delta t}\mathbb{E}\left[\left|T_{y,\zeta}^{n}\right|^2 + \left|T_{z,\zeta}^{n}\right|^2\right],
\end{aligned}
\tag{4.29}
$$

where

$$
\begin{aligned}
&C^* = \begin{cases} \eta + C_* + \eta C_*\Delta t, & |\alpha_k| \neq 0, \\ \max\{\eta + C_* + \eta C_*\Delta t, C_*'\}, & |\alpha_k| = 0, \end{cases} \\
&C_1' = (2K+3)\big(\Delta t + 1/\eta\big)K^2L^2\Upsilon^2, \\
&C_2' = (2K+3)\big(\Delta t + 1/\eta\big)K^2L^2\Upsilon^2 + 2KC_\phi C_0'L^2\Upsilon^2\Delta t, \\
&C_3' = C_2' + KC_0'\Upsilon^2(\Delta t)^\alpha, \quad C_4' = \max\big\{(2K+3)(\Delta t + 1/\eta),\, C_0'\big\}.
\end{aligned}
$$

Under the condition $\lambda_0 > |\lambda_1|$ and the definition $C' = C_\phi/\lambda_0^2$, it can be verified that there exist some fixed positive numbers ε_0, η_0, Δt_0, and $\underline{C}$ such that if $0 < \Delta t \le \Delta t_0$ and $\eta \ge \eta_0$, one has

$$
\frac{\varepsilon_0}{2C'(1+\varepsilon_0)} - C_1' - \frac{\varepsilon_0}{2C'}\bar{\lambda}_1^2 - KC_3' > \underline{C} > 0.
$$

Also, there exists a positive constant C such that for $0 < \Delta t \le \Delta t_0$ and $\eta = \eta_0$ one has

$$
C^* \le C, \quad C_i' \le C, \quad 1 - C\Delta t > 0.
$$

Now we set $\varepsilon = \varepsilon_0$ and $\eta = \eta_0$ in (4.29); then for $0 < \Delta t \le \Delta t_0$ we have

$$
\begin{aligned}
&\sum_{j=n}^{N-K}\mathbb{E}\left[\left|e_{y,\varsigma}^{j}\right|^2\right] + \frac{1}{1 - C\Delta t}\left(\frac{\varepsilon_0}{2C'(1+\varepsilon_0)} - C_1'\right)\sum_{j=n}^{N-K}\Delta t\mathbb{E}\left[\left|e_{z,\varsigma}^{j}\right|^2\right] \\
&\quad\le \frac{1 + C\Delta t}{1 - C\Delta t}\sum_{j=n}^{N-K}\sum_{i=1}^{k}|\alpha_i|\,\mathbb{E}\left[\left|e_{y,\varsigma}^{i+j}\right|^2\right] + \frac{\varepsilon_0}{2C'}\bar{\lambda}_1^2\sum_{j=n}^{N-K}\Delta t\mathbb{E}\left[\left|e_{z,\varsigma}^{j+1}\right|^2\right] \\
&\qquad + \frac{C\Delta t}{1 - C\Delta t}\sum_{j=n}^{N-K}\sum_{i=1}^{K}\mathbb{E}\left[\left|e_{y,\varsigma}^{i+j}\right|^2\right] + \frac{C_3'\Delta t}{1 - C\Delta t}\sum_{j=n}^{N-K}\sum_{i=1}^{K}\mathbb{E}\left[\left|e_{z,\varsigma}^{i+j}\right|^2\right] \\
&\qquad + \frac{C}{(1 - C\Delta t)\Delta t}\sum_{j=n}^{N-K}\mathbb{E}\left[\left|T_{y,\zeta}^{j}\right|^2 + \left|T_{z,\zeta}^{j}\right|^2\right].
\end{aligned}
\tag{4.30}
$$

By using the condition $\sum_{i=1}^{k}|\alpha_i| \le (1 + C_*\Delta t)$, we obtain

$$
\sum_{j=n}^{N-K}\sum_{i=1}^{k}|\alpha_i|\,\mathbb{E}\left[\left|e_{y,\varsigma}^{i+j}\right|^2\right] \le (1 + C_*\Delta t)\sum_{j=n+1}^{N}\mathbb{E}\left[\left|e_{y,\varsigma}^{j}\right|^2\right].
$$

Then we deduce

$$\mathbb{E}\left[\left|e_{y,\varsigma}^n\right|^2\right] + \underline{C}\Delta t \sum_{j=n}^{N-K} \mathbb{E}\left[\left|e_{z,\varsigma}^j\right|^2\right] \le CK\Delta t \sum_{j=n+1}^{N} \left(\mathbb{E}\left[\left|e_{y,\varsigma}^j\right|^2\right]\right) + R^{n,K}, \tag{4.31}$$

where

$$R^{n,K} = (1+C\Delta t) \sum_{j=N-K+1}^{N} \left(\mathbb{E}\left[\left|e_{y,\varsigma}^j\right|^2\right] + \mathbb{E}\left[\left|e_{z,\varsigma}^j\right|^2\right]\right) + \frac{C}{\Delta t} \sum_{j=n}^{N-K} \mathbb{E}\left[\left|T_{y,\zeta}^j\right|^2 + \left|T_{z,\zeta}^j\right|^2\right].$$

By the discrete backward Gronwall inequality [53], for $n = N-K, \dots, 0$ we have

$$\mathbb{E}\left[\left|e_{y,\varsigma}^n\right|^2\right] \le e^{CK\hat{T}} \left(R^{n,K} + CK^2\Delta t \max_{N-K+1\le j\le N} \mathbb{E}\left[\left|e_{y,\varsigma}^j\right|^2\right]\right), \tag{4.32}$$

where $\hat{T} = N\Delta t$. Then by using (4.31) again, we have the following estimate of $e_{z,\varsigma}^n$:

$$\begin{aligned} \underline{C}\Delta t \sum_{j=n}^{N-K} \mathbb{E}\left[\left|e_{z,\varsigma}^j\right|^2\right] &\le CK\Delta t \sum_{j=n+1}^{N} \mathbb{E}\left[\left|e_{y,\varsigma}^j\right|^2\right] + R^{n,K} \\ &\le (CK^2\hat{T}e^{CK\hat{T}} + 1) \left(R^{n,K} + CK^2\Delta t \max_{0\le j\le K-1} \mathbb{E}\left[\left|e_{y,\varsigma}^{N-j}\right|^2\right]\right). \end{aligned} \tag{4.33}$$

By combining the estimates in (4.32) and (4.33), we obtain

$$\begin{aligned} &\mathbb{E}\left[\left|e_{y,\varsigma}^n\right|^2\right] + \underline{C}\Delta t \sum_{j=n}^{N-K} \mathbb{E}\left[\left|e_{z,\varsigma}^j\right|^2\right] \\ &\quad \le \overline{C} \left(\sum_{j=N-K+1}^{N} \left(\mathbb{E}\left[\left|e_{y,\varsigma}^j\right|^2\right] + \mathbb{E}\left[\left|e_{z,\varsigma}^j\right|^2\right]\right) + \frac{1}{\Delta t} \sum_{j=n}^{N-K} \mathbb{E}\left[\left|T_{y,\zeta}^j\right|^2 + \left|T_{z,\zeta}^j\right|^2\right] \right), \end{aligned} \tag{4.34}$$

where

$$\overline{C} = \max\left\{C,\, 1 + C\Delta t,\, CK^2\Delta t\right\} \left(CK^2\hat{T}e^{CK\hat{T}} + 1\right).$$

The proof is completed. ∎

We have completed the stability analysis, and the above results can be used to show the convergence of Scheme 4.4. To show this, notice that for the exact solution (Y_{t_n}, Z_{t_n}), we have

$$\begin{aligned} Y_{t_n}^{t_n,X^n} &= \sum_{i=1}^{k} \alpha_i \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+i}}^{t_{n+i},X^{n+i}}\right] + \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}^{t_{n+i},X^{n+i}}\right] + \tilde{R}_y^n, \\ \lambda_0 \Delta t_{n,k} Z_{t_n}^{t_n,X^n} &= \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+k}}^{t_{n+k},X^{n,k}} G^{n,k}\right] + \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}^{t_{n+i},X^{n+i}} G^{n,i}\right] \\ &\quad - \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}_{t_n}^{X^n}\left[Z_{t_{n+i}}^{t_{n+i},X^{n+i}} H^{n,i}\right] + \tilde{R}_z^n. \end{aligned} \tag{4.35}$$

Here $\tilde{R}_y^n$ and $\tilde{R}_z^n$ are defined by

$$\tilde{R}_y^n = Y_{t_n}^{t_n,X^n} - \sum_{i=1}^{k} \alpha_i \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+i}}^{t_{n+i},X^{n+i}}\right] - \Delta t_{n,k} \sum_{i=0}^{k_y} \beta_i \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}^{t_{n+i},X^{n+i}}\right],$$

$$\tilde{R}_z^n = \lambda_0 \Delta t_{n,k} Z_{t_n}^{t_n,X^n} - \mathbb{E}_{t_n}^{X^n}\left[Y_{t_{n+k}}^{t_{n+k},X^{n,k}} G^{n,k}\right] - \Delta t_{n,k} \sum_{i=1}^{k_f} \gamma_i \mathbb{E}_{t_n}^{X^n}\left[f_{t_{n+i}}^{t_{n+i},X^{n+i}} G^{n,i}\right] + \Delta t_{n,k} \sum_{i=1}^{k_z} \lambda_i \mathbb{E}_{t_n}^{X^n}\left[Z_{t_{n+i}}^{t_{n+i},X^{n+i}} H^{n,i}\right].$$

By replacing $T_{y,\varsigma}^n$ in (4.16) and $T_{z,\varsigma}^n$ in (4.19) with $\tilde{R}_y^n$ and $\tilde{R}_z^n$ in (4.35), respectively, and using similar arguments as those used in the proof of Theorem 2, we obtain the following corollary, which implies the error estimate of Scheme 4.4.

COROLLARY 1. *Let (X_t, Y_t, Z_t) and (X^n, Y^n, Z^n) be the exact solution of* (2.1) *and the approximated solution by Scheme* 4.4. *For sufficiently small Δt it holds that*

$$\begin{aligned} &\mathbb{E}\left[|Y_{t_n} - Y^n|^2\right] + \Delta t \sum_{i=n}^{N-k} \mathbb{E}\left[\left|Z_{t_i} - Z^i\right|^2\right] \\ &\le C\left(\sum_{i=0}^{k-1}\left(\mathbb{E}\left[\left|e_y^{N-i}\right|^2\right] + \mathbb{E}\left[\left|e_z^{N-i}\right|^2\right]\right) + \frac{1}{\Delta t}\sum_{i=n}^{N-k} \mathbb{E}\left[\left|\tilde{R}_y^i\right|^2 + \left|\tilde{R}_z^i\right|^2\right]\right). \end{aligned} \tag{4.36}$$

We close this section by remarking that our stability analysis can be applied to many existing schemes such as those in [11, 42, 48, 54]. However, for the multistep schemes in [49], the involved parameters $\{\alpha_i^{n,k}\}_{i=1}^k$ do not satisfy the condition:

$$\sum_{i=1}^{k} |\alpha_i^{n,k}| \le (1 + C_* \Delta t),$$

and for the Crank–Nicolson scheme in [31, 50], the parameters $\{\lambda_i\}_{i=0}^1$ do not satisfy the condition $\lambda_0 > |\lambda_1|$. Thus, our analysis results do not cover these two cases. The stability analysis for the multistep schemes in [49] is still open.

5. Discussion and applications.

In this part, we shall give some discussions on Schemes 3.1 and 4.4. As mentioned, many existing probabilistic numerical schemes can be viewed as special cases of the proposed general scheme (Scheme 3.1 or 3.2). To see this, we shall give some specific examples as follows.

First, we consider the Euler-type schemes that have been well studied in the literature [4, 6, 7, 14, 24, 25, 26, 32, 47]. We take the following Euler scheme as an example [7, 47]:

$$\begin{cases} Y^n = \mathbb{E}_{t_n}^{X^n}\left[Y^{n+1}\right] + \Delta t_n \mathbb{E}_{t_n}^{X^n}\left[f^{n+1}\right] \quad \text{or} \quad Y^n = \mathbb{E}_{t_n}^{X^n}\left[Y^{n+1}\right] + \Delta t_n f^n, \\ Z^n = \dfrac{1}{\Delta t_n} \mathbb{E}_{t_n}^{X^n}\left[Y^{n+1} \Delta W_{n,1}\right], \end{cases} \tag{5.1}$$

where $f^n = f(t_n, X^n, Y^n, Z^n)$. Actually, the above scheme is a special case of our scheme when the parameters are chosen as $\alpha_0 = \alpha_1 = \tilde{\alpha}_1 = \lambda_0 = 1$, $\beta_0 = 1$ (or $\beta_0 = 0,\ \beta_1 = 1$) in (3.17).

Next, let's consider the θ-scheme in [42, 48, 52]:

$$\begin{cases} Y^n = \mathbb{E}_{t_n}^{X^n}\left[Y^{n+1}\right] + \Delta t(1-\theta_1)\mathbb{E}_{t_n}^{X^n}\left[f^{n+1}\right], \\ \quad 0 = \mathbb{E}_{t_n}^{X^n}\left[Y^{n+1} \Delta W_{n,1}^\top\right] + \Delta t(1-\theta_2)\mathbb{E}_{t_n}^{X^n}\left[f^{n+1} \Delta W_{n,1}^\top\right] \\ \qquad - \Delta t_n\{(1-\theta_2)\mathbb{E}_{t_n}^{X^n}\left[Z^{n+1}\right] + \theta_2 Z^n\}. \end{cases} \tag{5.2}$$

This scheme can be obtained by setting the parameters in (3.17) as $\alpha_0 = \alpha_1 = \tilde{\alpha}_1 = 1, \beta_0 = \theta_1, \beta_1 = 1-\theta_1$, $\gamma_1 = 1-\theta_2$, $\lambda_0 = \theta_2$, $\lambda_1 = 1-\theta_2$, $H^{n,1} = 1$, and $G^{n,1} = \Delta W_{n,1}^{\top}$.

Further, we consider the following multistep scheme [11]:

$$\begin{cases} Y^n = \sum_{i=1}^{r} a_i \mathbb{E}_{t_n}^{X^n}\left[Y^{n+i}\right] + \Delta t \sum_{j=0}^{r} b_j \mathbb{E}_{t_n}^{X^n}\left[f^{n+j}\right], \\ Z^n = \sum_{i=1}^{r} \alpha_i \mathbb{E}_{t_n}^{X^n}\left[Y^{n+i} H_{n,i}\right] + \Delta t \sum_{i=1}^{r} \beta_i \mathbb{E}_{t_n}^{X^n}\left[f^{n+i} H_{n,i}\right], \end{cases} \tag{5.3}$$

This scheme is a special case when the parameters in (3.17) are chosen as $\alpha_0 = 1$, $\lambda_0 = 1$, $\alpha_i = a_i$, $\beta_j = b_j$, $\tilde{\alpha}_i = \alpha_i$, $\lambda_i = 0$, $\gamma_i = \beta_i$, and $G^{n,i} = H_{n,i}/\Delta t$, $i \geq 1$, $j \geq 0$.

We also list in Table 5.1 the connection between some popular schemes and our general schemes.

In the following, we give further remarks on Schemes 3.1 and 4.4.

- We remark that Scheme 3.1 is a semidiscrete scheme. To get a fully discrete scheme, one has to deal with the associated conditional expectations. Popular approaches include regression methods [6, 14, 26, 32], the Gaussian quadrature method [49, 52, 53], Malliavin calculus methods [7, 8], cubature methods [15, 16], and quantization methods [2, 18, 39], to name a few.
- We remark that the Runge–Kutta schemes in [12] are not covered by the general scheme. Moreover, since the involved parameters in the multistep scheme in [22, 49] and the Crank–Nicolson scheme in [31, 50] do not satisfy Assumption 2, the stability analysis here is not applicable to these schemes.

Table 5.1
Relationships between the unified discretization (3.17) *and popular schemes in the literature.*

Schemes	Parameters	Reference
Euler	$\alpha_0 = \alpha_1 = \tilde{\alpha}_1 = \lambda_0 = 1,\ \beta_0 = 1$ or $\alpha_0 = \alpha_1 = \tilde{\alpha}_1 = \lambda_0 = 1,\ \beta_0 = 0,\ \beta_1 = 1$	[4, 14, 24, 25, 26, 32, 47]
Crank–Nicolson	$\alpha_0 = \alpha_1 = \tilde{\alpha}_1 = 1,\ \beta_0 = \beta_1 = \frac{1}{2}$, $\lambda_0 = \lambda_1 = \frac{1}{2},\ \gamma_1 = \frac{1}{2}$	[31, 50]
θ-scheme I	$\alpha_0 = \alpha_1 = \tilde{\alpha}_1 = 1,\ \beta_0 = \theta_1,\ \beta_1 = 1-\theta_1$, $\gamma_1 = 1-\theta_2,\ \lambda_0 = \theta_2,\ \lambda_1 = 1-\theta_2$, $H^{n,1} = 1, \quad G^{n,1} = \Delta W_{n,1}^{\top}$ $\{\theta_i\}$ are defined in the associated references.	[42, 48, 52]
θ-scheme II	$\alpha_0 = \alpha_1 = 1,\ \ \beta_0 = \theta_1,\ \ \beta_1 = 1-\theta_1,\ \ \lambda_0 = \theta_3$, $\tilde{\alpha}_1 = \theta_3 - \theta_4,\ \ \lambda_1 = -\theta_4,\ \ \gamma_1 = 1-\theta_2$, $H^{n,1} = 1,\ \ G^{n,1} = \Delta W_{n,1}^{\top}$ $\{\theta_i\}$ are defined in the associated reference.	[51]
Multistep I	$\alpha_0 = 1,\ \ \lambda_0 = 1,\ \ \alpha_i = a_i,\ \ \beta_j = b_j,\ \ \tilde{\alpha}_i = \alpha_i$, $\lambda_i = 0, \quad \gamma_i = \beta_i,\ \ G^{n,i} = H_{n,i}/\Delta t, \quad i \geq 1,\ j \geq 0$ $\{a_i, b_j, \alpha_i, \beta_i\}$ are defined in the associated reference.	[11]
Multistep II	$\alpha_0 = \alpha_{k,0}, \quad \lambda_0 \Delta t_{n,k} = 1,\ \ \beta_0 \Delta t_{n,k} = -1$, $\alpha_i = -\alpha_i^{n,k}, \quad \tilde{\alpha}_i = \alpha_i^{n,k}, \quad \beta_i = \lambda_i = \gamma_i = 0$, $G^{n,i} = \Delta W_{n,i}^{\top}, \quad i \geq 1$ $\{\alpha_i^{n,k}, \alpha_{k,0}\}$ are defined in the associated references.	[22, 49]
Multistep III	$\alpha_0 = \alpha_k \text{ (or } \alpha_{K_y}) = 1, \quad \beta_i = b_{K_y,i}^k \text{ (or } \gamma_{K_y,i}^{K_y})$, $\lambda_i = b_{K_z,i}^l \text{ (or } \gamma_{K_z,i}^1), \quad H^{n,i} = 1, \quad \gamma_j = b_{K_z,j}^l \text{ (or } \gamma_{K_z,j}^1)$, $\tilde{\alpha}_l \text{ (or } \tilde{\alpha}_1) = 1,\ \ G^{n,j} = \Delta W_{n,j}^{\top}, \quad i \geq 0,\ j \geq 1$ $\{b_{K_y,i}^k, b_{K_z,i}^l, \gamma_{K_y,i}^{K_y}, \gamma_{K_z,i}^1\}$ are defined in the associated references.	[45, 53]

- In addition, one may expect to generate new schemes based on (3.17). In particular, one can resort to the machine learning methods as in [37] to train the involved parameters using artificial data, and this will result in potentially accelerated new schemes.

Finally, we present an illustrated example to show that the analysis in this work can be directly applied to existing schemes. In particular, we consider the multistep scheme proposed in [54].

SCHEME 5.1. *Given the terminal condition* $\{(Y^{N-i}, Z^{N-i})\}_{0\leq i\leq k-1}$, *we solve* Y^n *and* Z^n *by*

$$Y^n = \mathbb{E}_{t_n}^{X^n}\left[Y^{n+1}\right] + \Delta t_n \sum_{i=0}^{k_y} b_{k_y,i}^n \mathbb{E}_{t_n}^{X^n}\left[f^{n+i}\right],$$

$$Z^n = \frac{1}{\Delta t_n}\mathbb{E}_{t_n}^{X^n}\left[Y^{n+1}\Delta\tilde{W}_{n,1}^\top\right] + \sum_{i=1}^{k_f} b_{k_f,i}^n \mathbb{E}_{t_n}^{X^n}\left[f^{n+i}\Delta\tilde{W}_{n,i}^\top\right].$$

Here we have

$$b_{K,i}^n = \frac{1}{\Delta t_n}\int_{t_n}^{t_{n+1}} \prod_{j=0,j\neq i}^{K}\left(\frac{s-t_{n+j}}{t_{n+i}-t_{n+j}}\right) ds, \quad \Delta\tilde{W}_{n,i} = \int_{t_n}^{t_{n+i}} \phi_s dW_s,$$

where ϕ_s is a backward orthogonal polynomial satisfying

$$\int_{t_n}^{t_{n+1}} \phi_s ds = 1.$$

Notice that this discretization scheme falls into Scheme 4.4 by choosing

$$\alpha_1 = 1, \quad \tilde{\alpha} = 1, \quad \lambda_0 = 1, \quad \beta_i = b_{k_y,i}^n, \quad \gamma_i = b_{k_f,i}^n,$$

and $\lambda_i = 0$ for $1 \leq i \leq k_z$. Moreover, it is easy to verify that Assumption 2 is fulfilled. Consequently, we learn by applying Theorem 2 that the multistep scheme is stable, and we can then further get its convergence result (which coincides with Theorem 5.1 in [54]) by Corollary 1.

THEOREM 3. *We suppose that the terminal conditions are given with certain accuracy:*

$$\max_{0\leq i\leq k-1} \mathbb{E}\left[|e_y^{N-i}|\right] = \mathcal{O}\Big((\Delta t)^{k_y+1}\Big),$$
$$\max_{0\leq i\leq k-1} \mathbb{E}\left[|e_z^{N-i}|\right] = \mathcal{O}\Big((\Delta t)^{k_z+1} + (\Delta t)^{k_f-k_z+1}\Big).$$

We also suppose that the numerical solution X^{n+1} *is solved with certain accuracy:*

$$\max_{0\leq n\leq N} \mathbb{E}\big[|X^n|^r\big] \leq C_g(1+\mathbb{E}\big[|X_0|^r\big]),$$
$$\left|\mathbb{E}_{t_n}^{X^n}\left[g\left(X_{t_{n+1}}^{t_n,X^n}\right) - g\left(X^{n+1}\right)\right]\right| \leq C_g(1+|X^n|^{2r_1})(\Delta t)^{\beta+1},$$
$$\left|\mathbb{E}_{t_n}^{X^n}\left[\left(g\left(X_{t_{n+1}}^{t_n,X^n}\right) - g\left(X^{n+1}\right)\right)G^{n,1}\right]\right| \leq C_g(1+|X^n|^{2r_2})(\Delta t)^{\gamma+1}.$$

Then for sufficiently small Δt *we have*

$$\mathbb{E}\left[|e_y^n|^2\right] + \Delta t \sum_{i=n}^{N-k} \mathbb{E}\left[|e_z^i|^2\right]$$
$$\leq C\Big((\Delta t)^{2k_y+2} + (\Delta t)^{2k_y+2} + (\Delta t)^{2(k_y-k_z+1)} + (\Delta t)^{2\beta} + (\Delta t)^{2\gamma}\Big).$$

6. Conclusions. We propose a general discretization framework for numerical solution of FBSDEs. The framework covers many interesting temporal discretization probabilistic schemes in the literature. We analyze the consistency, stability, and convergence for the proposed scheme. As a byproduct, we obtain a stochastic mean square version of the Lax equivalence theorem—showing that a consistent discretization scheme for FBSDEs is convergent if and only if it is stable. Applications of the analysis results to existing numerical schemes are discussed. In our future studies, we shall extend these results to second order FBSDEs [46, 55] and mean-field BSDEs [43]. Meanwhile, we will consider the use of the machine learning technique [37] to generate some potentially accelerated new schemes based on the unified scheme in section 3.

Acknowledgment. The authors would like to thank the referees for their valuable comments, which much improved the quality of the paper.

REFERENCES

[1] V. Bally, *Approximation scheme for solutions of BSDE*, in Backward Stochastic Differential Equations, N. El Karoui and L. Mazliak, eds., Addison-Wesley-Longman, Boston, 1997, pp. 177–191.

[2] V. Bally, G. Pagés, and J. Printems, *A stochastic quantization method for non-linear problems*, Monte Carlo Methods Appl., 7 (2001), pp. 21–33.

[3] C. Beck, W. E, and A. Jentzen, *Machine learning approximation algorithms for high-dimensional fully nonlinear partial differential equations and second-order backward stochastic differential equations*, J. Nonlinear Sci., 29 (2019), pp. 1563–1619.

[4] C. Bender and R. Denk, *A forward scheme for backward SDEs*, Stochastic Process. Appl., 117 (2007), pp. 1793–1812.

[5] C. Bender and J. Zhang, *Time discretization and Markovian iteration for coupled FBSDEs*, Ann. Appl. Probab., 18 (2008), pp. 143–177.

[6] C. Bender and J. Steiner, *Least-squares Monte Carlo for Backward SDEs*, in Numerical Methods in Finance, Springer Proc. Math. 12, Springer, New York, 2012, pp. 257–289.

[7] B. Bouchard and N. Touzi, *Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations*, Stochastic Process. Appl., 111 (2004), pp. 175–206.

[8] B. Bouchard and X. Warin, *Monte-Carlo valorisation of American options: Facts and new algorithms to improve existing methods*, in Numerical Methods Springer Proc. Math. 12, Springer, New York, 2012, pp. 215–255.

[9] P. Briand, B. Delyon, and J. Mémin, *Donsker-type theorem for BSDE*, in Finance, Electron. Comm. Probab., 6 (2001), pp. 1–14.

[10] J. Butcher, *Numerical Methods for Ordinary Differential Equations*, John Wiley & Sons, New York, 2003.

[11] J. F. Chassagneux, *Linear multi-step schemes for BSDEs*, SIAM J. Numer. Anal., 52 (2014), pp. 2815–2836.

[12] J. F. Chassagneux and D. Crisan, *Runge-Kutta schemes for backward stochastic differential equations*, Ann. Appl. Probab., 24 (2014), pp. 679–720.

[13] J. F. Chassagneux and A. Richou, *Numerical stability analysis of the Euler scheme for BSDEs*, SIAM J. Numer. Anal., 53 (2015), pp. 1172–1193.

[14] K. W. Chau and C. W. Oosterlee, *Stochastic grid bundling method for backward stochastic differential equations*, Int. J. Comput. Math., 96 (2019), pp. 2272–2301.

[15] D. Crisan and K. Manolarakis, *Solving backward stochastic differential equations using the cubature method: Application to nonlinear pricing*, SIAM J. Financ. Math., 3 (2012), pp. 534–571.

[16] D. Crisan and K. Manolarakis, *Second order discretization of a backward SDEs and simulation with the cubature method*, Ann. Appl. Probab., 24 (2014), pp. 652–678.

[17] D. Chevance, *Numerical methods for backward stochastic differential equations*, in Numerical Methods in Finance, L. C. G. Rogers and D. Talay, eds., Cambridge University Press, Cambridge, UK, 1997, pp. 232–244.

[18] F. Delarue and S. Menozzi, *A foward-backward stochastic algorithm for quasi-linear PDEs*, Ann. Appl. Probab., 16 (2006), pp. 140–184.

[19] J. Douglas, J. Ma, and P. Protter, *Numerical methods for forward backward stochastic differential equations*, Ann. Appl. Probab., 6 (1996), pp. 940–968.

[20] W. E, J. Han, and A. Jentzen, *Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations*, Commun. Math. Stat., 5 (2017), pp. 349–380.

[21] Y. Fu, W. Zhao, and T. Zhou, *Efficient spectral sparse grid approximations for solving multi-dimensional forward backward SDEs*, Discrete Contin. Dyn. Syst. Ser. B, 22 (2017), pp. 3439–3458.

[22] Y. Fu, W. Zhao, and T. Zhou, *Multistep schemes for forward backward stochastic differential equations with jumps*, J. Sci. Comput., 69 (2016), pp. 651–672.

[23] E. Gianin, *Risk measures via g-expectations*, Insurance Math. Econom., 39 (2006), pp. 19–34.

[24] E. Gobet and C. Labart, *Error expansion for the discretization of backward stochastic differential equations*, Stochastic Process. Appl., 117 (2007), pp. 803–829.

[25] E. Gobet and C. Labart, *Solving BSDE with adaptive control variate*, SIAM J. Numer. Anal., 48 (2010), pp. 257–277.

[26] E. Gobet, J. P. Lemor, and X. Warin, *A regression-based Monte Carlo method to solve backward stochastic differential equations*, Ann. Appl. Probab., 15 (2005), pp. 2172–2202.

[27] P. Henry-Labordere, X. Tan, and N. Touzi, *A numerical algorithm for a class of BSDEs via the branching process*, Stochastic Process. Appl., 124 (2014), pp. 1112–1140.

[28] Y. Hu and S. Peng, *Solutions of forward-backward stochastic differential equations*, Probab. Theory Related Fields, 103 (1995), pp. 273–283.

[29] N. El Karoui, S. Peng, and M. C. Quenez, *Backward stochastic differential equations in finance*, Math. Finance, 7 (1997), pp. 1–71.

[30] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1992.

[31] Y. Li, J. Yang, and W. Zhao, *Convergence error estimates of the Crank-Nicolson scheme for solving decoupled FBSDEs*, Sci. China Math., 60 (2017), pp. 923–948.

[32] J. P. Lemor, E. Gobet, and X. Warin, *Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations*, Bernoulli, 12 (2006), pp. 889–916.

[33] J. Ma, P. Protter, and J.Yong, *Solving forward-backward stochastic differential equations explicitly—a four step scheme*, Probab. Theory Related Fields, 98 (1994), pp. 339–359,.

[34] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*, Lecture Notes in Math. 1702, Springer-Verlag, Berlin, 1999.

[35] J. Ma, P. Protter, J. San Martin, and S. Soledad, *Numerical method for backward stochastic differential equations*, Ann. Appl. Probab., 12 (2002), pp. 302–316.

[36] G. N. Milstein and M. V. Tretyakov, *Numerical algorithms for forward-backward stochastic differential equations*, SIAM J. Sci. Comput., 28 (2006), pp. 561–582.

[37] S. Mishra, *A Machine Learning Framework for Data Driven Acceleration of Computations of Differential Equations*, Research Report 2018-28, Seminar for Applied Mathematics, ETH Zürich, 2018.

[38] D. Nualart, *The Malliavin Calculus and Related Topics*, Springer-Verlag, Berlin, 1995.

[39] G. Pagés and A. Sagna, *Improved error bounds for quantization based numerical schemes for BSDE and nonlinear filtering*, Stochastic Process. Appl., 128 (2018), pp. 847–883.

[40] E. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett., 14 (1990), pp. 55–61.

[41] S. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim., 28 (1990), pp. 966–979.

[42] M. J. Ruijter and C. W. Oosterlee, *A Fourier cosine method for an efficient computation of solutions to BSDEs*, SIAM J. Sci. Comput., 37 (2015), pp. A859–A889.

[43] Y. Sun, W. Zhao, and T. Zhou, *Explicit θ-schemes for mean-field backward stochastic differential equations*, SIAM J. Numer. Anal., 56 (2018), pp. 2672–2697.

[44] T. Tang, W. Zhao, and T. Zhou, *Deferred correction methods for forward backward stochastic differential equations*, Numer. Math. Theory Methods Appl., 10 (2017), pp. 222–242.

[45] L. Teng, A. Lapitckii, and M. Gunther, *A multi-step scheme based on cubic spline for solving backward stochastic differential equations*, Appl. Numer. Math., 150 (2020), pp. 117–138.

[46] J. Yang, W. Zhao, and T. Zhou, *Explicit deferred correction methods for second-order forward backward stochastic differential equations*, J. Sci. Comput., 79 (2019), pp. 1409–1432.
[47] J. Zhang, *A numerical scheme for BSDEs*, Ann. Appl. Probab., 14 (2004), pp. 459–488.
[48] W. Zhao, L. Chen, and S. Peng, *A new kind of accurate numerical method for backward stochastic differential equations*, SIAM J. Sci. Comput., 28 (2006), pp. 1563–1581.
[49] W. Zhao, Y. Fu, and T. Zhou, *New kinds of high-order multistep schemes for coupled forward backward stochastic differential equations*, SIAM J. Sci. Comput., 36 (2014), pp. A1731–A1751.
[50] W. Zhao, Y. Li, and L. Ju, *Error estimates of the Crank-Nicolson scheme for solving backward stochastic differential equations*, Int. J. Numer. Anal. Model., 10 (2013), pp. 876–898.
[51] W. Zhao, Y. Li, and G. Zhang, *A generalized θ-scheme for solving backward stochastic differential equations*, Discrete Contin. Dyn. Syst. Ser. B, 17 (2012), pp. 1585–1603.
[52] W. Zhao, J. Wang, and S. Peng, *Error estimates of the θ-scheme for backward stochastic differential equations*, Discrete Contin. Dyn. Syst. Ser. B, 12 (2009), pp. 905–924.
[53] W. Zhao, G. Zhang, and L. Ju, *A stable multistep scheme for solving backward stochastic differential equations*, SIAM J. Numer. Anal., 48 (2010), pp. 1369–1394.
[54] W. Zhao, W. Zhang, and L. Ju, *A multistep scheme for decoupled forward-backward stochastic differential equations*, Numer. Math. Theory Methods Appl., 9 (2016), pp. 262–288.
[55] W. Zhao, T. Zhou, and T. Kong, *High order numerical schemes for second-order FBSDEs with applications to stochastic optimal control*, Commun. Comput. Phys., 21 (2017), pp. 808–834.