Advances in Applied Mathematics and Mechanics
Adv. Appl. Math. Mech., Vol. **12**, No. 6, pp. 1457-1480

DOI: 10.4208/aamm.OA-2019-0345
December 2020

Finite Element Methods for Nonlinear Backward Stochastic Partial Differential Equations and Their Error Estimates

Xu Yang[1] and Weidong Zhao[2,*]

[1] *School of Mathematics, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China*
[2] *School of Mathematics, Shandong University, Jinan, Shandong 250100, China*

Received 28 November 2019; Accepted (in revised version) 10 March 2020

Abstract. In this paper, we consider numerical approximation of a class of nonlinear backward stochastic partial differential equations (BSPDEs). By using finite element methods in the physical space domain and the Euler method in the time domain, we propose a spatial finite element semi-discrete scheme and a spatio-temporal full discrete scheme for solving the BSPDEs. Errors of the schemes are rigorously analyzed and theoretical error estimates with convergence rates are obtained.

AMS subject classifications: 60H15, 60H35, 65C30, 65M60

Key words: Backward stochastic partial differential equations, finite element method, error estimate.

---

1 Introduction

Consider the following backward stochastic partial differential equations (BSPDEs):

$$-du(t,x)-\mathcal{L}u(t,x)dt=f\big(t,\nabla u(t,x),u(t,x),v(t,x)\big)dt-v(t,x)dW(t), \tag{1.1}$$

for $t\in[0,T)$, $x\in D$. The associated boundary and terminal conditions are given by

$$\begin{cases} u(t,x)=v(t,x)=0, & x\in\partial D,\\ u(T,x)=u_T(x), & x\in D. \end{cases}$$

Here $T>0$ is a positive constant and $D\subset\mathbb{R}^d$, $(d=1,2,3)$ is a bounded convex domain with a sufficiently smooth boundary ∂D. The operator $\mathcal{L}$ is a second order elliptic operator

*Corresponding author.
Emails: xuyang96@cumt.edu.cn (X. Yang), wdzhao@sdu.edu.cn (W. Zhao)

defined by

$$\mathcal{L}u =: \sum_{j,k=1}^{d} \frac{\partial}{\partial x_j}\left(a_{jk}(x)\frac{\partial u}{\partial x_k}\right) - a_0(x)u.$$

Moreover, $W=\{W(t)\colon t\in[0,T]\}$ is a standard Wiener process defined on a completed probability space $(\Omega,\mathcal{F},\mathbb{F},\mathbb{P})$ with $\mathbb{F}=:\{\mathcal{F}_t\colon t\in[0,T]\}$ being the augmented natural filtration generated by W. Also, u_T is an $\mathcal{F}_T$-measurable random field. In fact, the BSPDEs (1.1) can be mathematically interpreted as the following integral form

$$\begin{aligned} &u(t,x)-\int_t^T \mathcal{L}u(s,x)ds \\ =&u_T(x)+\int_t^T f(s,\nabla u(s,x),u(s,x),v(s,x))ds-\int_t^T v(s,x)dW(s). \end{aligned} \tag{1.2}$$

Precise assumptions on the operator $\mathcal{L}$, the function f, and the terminal function u_T will be discussed in Section 2.

The above BSPDEs play an essential role in many real applications. In particular, it serves as the adjoint equations in stochastic optimal control problems governed by stochastic parabolic equations [3, 11, 34]. Also, the nonlinear BSPDEs appears as the value functions in the optimal control problems of non-Markovian SDEs [20]. Other applications of BSPDEs include the nonlinear filtering problems [18], and mathematical finance [9, 16]. Recently, great attention has been paid to develop theoretical analysis of BSPDEs. Hu and Peng [11, 12] were the first to investigate the well-posedness of adapted solutions to semilinear backward stochastic evolution equations. The existence and uniqueness, as well as the regularity, of the adapted solutions for a class of BSPDEs was discussed in [16, 24, 33]. For more recent developments, one can refer to [1, 7, 9, 22] and references therein.

As analytic solutions to BSPDEs are seldom available, numerical methods become popular approaches for solving BSPDEs. As a generalization of backward stochastic differential equations (BSDEs), the associated numerical methods for BSPDEs have not been well studied. In contract, the numerical methods for BSDEs have been well developed in recent years [2, 4, 10, 15, 21, 23, 27, 28, 30, 32]. Up to now, there exist only a very limited number of works [8, 26] devoted to this field. In [8], finite element methods for linear forward-backward stochastic heat equations were considered and a rigorous convergence analysis for the spacial semi-discrete scheme was presented for linear BSPDEs. A semidiscrete Galerkin scheme based on spectral method for BSPDEs was studied in [26].

The primary goal of this paper is to develop and analyze numerical schemes which are used to approximate the solutions of nonlinear BSPDEs (1.1). More precisely, we study the strong approximation errors caused by spatial semi-discretization and space-time full-discretization of (1.1). We first consider semidiscrete finite element method for (1.1) and by a combination of the finite element method together with a linear implicit Euler time-stepping scheme, we also investigate a spatio-temporal discretization of (1.1). For both cases, we get the error estimates with precise strong convergence rate shown in

Theorem 3.1 and Theorem 4.2. The convergence analysis of the proposed semidiscrete scheme is based on some classic error estimates from deterministic finite element approximation and some techniques from stochastic calculus in Hilbert space such as Itô's formula and the Burkholder–Davis–Gundy inequality. The proof of Theorem 4.2 for the full-discretization is more complicated and more technical. We first established a stability theorem for the proposed fully discrete approximation, based on which, together with some classical error estimates from the deterministic numerical analysis, we eventually prove the strong approximation errors with convergence rate given. To the best of our knowledge, this work is the first try to deal with numerical approximation BSPDEs of type (1.1), the results we obtained are quite novel in the literature. The scientific novelty and main contributions of this work could be summarized as follows.

- In difference to already existing results, the problem we considered here is more general, especially the case f including the gradient of u is covered, which makes the problem more complicated. As a consequence, the numerical analysis of the problem becomes more involved.
- We present the convergence analysis for the fully discrete scheme–showing that the proposed scheme admits a first order rate of convergence in both the temporary domain and the physical domain. Compared to the spatially semi-discrete case, the analysis for the spatio-temporally discrete scheme is much more involved and technical. In particular, the Itô formula, which plays a powerful role in the error analysis for the spatially semi-discrete scheme, is not applicable for the fully discrete case.

The rest of this paper is organized as follows. In Section 2, some preliminaries are collected and the well-posedness of the BSPDEs is discussed. Section 3 is devoted to the convergence analysis of the semi-discrete scheme. This is followed by the analysis of the fully discrete scheme in Section 4. Finally, we give some concluding remarks in Section 5.

2 Preliminaries

In this section, we present some preliminaries. Let $|\cdot|$ denote the usual Euclidean vector norm. For a Hilbert space X with inner product $(\cdot,\cdot)_X$ and the induced norm $\|\cdot\|_X$, we denote by $L^2(\Omega,\mathcal{F}_t;X)$ the space of all $\mathcal{F}_t$-measurable X-valued random variables η satisfying

$$\mathbb{E}\left[\|\eta\|_X^2\right]<\infty.$$

Then for $p\geq 1$, $L^p_{\mathcal{F}}((0,T);X)$ is the space of all $\mathbb{F}$-adapted X-valued processes $\varphi(t)$ on $[0,T]$ that satisfy

$$\mathbb{E}\left[\int_0^T \|\varphi(t)\|_X^p dt\right]<\infty.$$

Obviously, for the case $p=2$, $L^2_{\mathcal{F}}(0,T;X)$ is a Hilbert space which is a subspace of the Hilbert space $L^2([0,T]\times\Omega;X)$. We also denoted by $C_{\mathcal{F}}\big([0,T];L^2(\Omega,X)\big)$ the space of all

$\mathbb{F}$-adapted, mean square continuous, X-valued processes $\psi(t)$ satisfying

$$\sup_{0\le t\le T}\mathbb{E}\left[\|\psi(t)\|_X^2\right]<\infty.$$

Similarly, $L^2\big(\Omega;C([0,T];X)\big)$ is a subspace of $C_{\mathcal{F}}([0,T];L^2(\Omega,X))$ in which it holds

$$\mathbb{E}\left[\sup_{0\le t\le T}\|\psi(t)\|_X^2\right]<\infty.$$

We introduce the following assumptions for the BSPDEs (1.1).

Assumption 2.1. For the operator $\mathcal{L}$, we suppose that

- The functions $a_{jk}(x)$ $(j,k=1,\cdots,d)$ and $a_0(x)$ are continuous functions in $\bar{D}$ such that $a_0(x)\ge 0$,
- The matrix $\left[a_{jk}(x)\right]_{j,k}$ is symmetric and uniformly positive definite in D, i.e.,

$$a_{jk}(x)=a_{kj}(x),\quad \sum_{j,k=1}^d a_{jk}(x)\xi_j\xi_k\ge\kappa\sum_{j=1}^d\xi_j^2\quad\text{with}\ \ \kappa>0\ \ \text{for any}\ \ \xi\in\mathbb{R}^d.$$

Now we define the bilinear form associated with the operator $-\mathcal{L}$ as follows:

$$a(u,\phi)=:\int_D\sum_{j,k=1}^d a_{jk}(x)\frac{\partial u}{\partial x_j}(x)\frac{\partial\phi}{\partial x_k}(x)+a_0(x)u(x)\phi(x)dx. \tag{2.1}$$

Assumption 2.2. The function $f\colon [0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ satisfies the following Lipschitz condition, i.e., there is a constant $L>0$ such that

$$|f(t,q_1,u_1,v_1)-f(t,q_2,u_2,v_2)|\le L\Big(|q_1-q_2|+|u_1-u_2|+|v_1-v_2|\Big), \tag{2.2}$$

for any $(q_1,u_1,v_1),(q_2,u_2,v_2)\in\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}$.

Assumption 2.3. The function $f(t,q,u,v)$ with $(t,q,u,v)\in[0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}$ has continuous and uniformly bounded first and second partial derivatives with respect to t, q, u and v.

Let $A=-\mathcal{L}$, we define the Hilbert space $H=L^2(D)$ equipped with inner product $(\cdot,\cdot)$ and the induced norm $\|\cdot\|$. Next, we introduce the space related to the fractional powers of the linear operator A. For $\alpha\in\mathbb{R}$, define $\dot{H}^\alpha=:\mathcal{D}(A^{\alpha/2})$ with the norm $\|\cdot\|_\alpha=:\|A^{\alpha/2}\cdot\|$. By the definition of fractional powers of the of A, it is well-known that $\dot{H}^0=H$, $\dot{H}^1=H_0^1(D)$ and $\dot{H}^2=H^2(D)\cap H_0^1(D)=\mathcal{D}(A)$.

Now, we formulate problem (1.1) in an abstract form in Itô's sense taking values in the Hilbert space $(H,(\cdot,\cdot),\|\cdot\|)$

$$-du(t)+Au(t)dt=f\big(t,\nabla u(t),u(t),v(t)\big)dt-v(t)dW(t),\quad u(T)=u_T, \tag{2.3}$$

for $t\in\left[0,T\right]$. We are ready to define the weak solution of (2.3). For more details, one can refer to [7,12,33].

Definition 2.1 (Weak solution)**.** A pair of random fields

$$(u,v)\in\Big(C_{\mathcal{F}}\big([0,T];L^2(\Omega,H)\big)\cap L^2_{\mathcal{F}}((0,T);\dot{H}^1)\Big)\times L^2_{\mathcal{F}}((0,T);H)$$

is called the weak solution to (2.3) if it satisfies almost surely

$$\begin{aligned}&(u(t),\phi)+\int_t^T a\big(u(s),\phi\big)\,ds\\ =&(u_T,\phi)+\int_t^T \big(f\big(s,\nabla u(s),u(s),v(s)\big),\phi\big)\,ds\\ &-\int_t^T (v(s),\phi)\,dW(s),\quad \forall\phi\in\dot{H}^1,\quad \forall t\in[0,T].\end{aligned}\tag{2.4}$$

Now we give the following theorem that shows that wellposedness of problem (2.3) [1, 7, 22, 24, 26, 33].

Theorem 2.1. *Under Assumptions* 2.1-2.2*, and assume that* $u_T:\Omega\to H$ *is* $\mathcal{F}_T$*-measurable and satisfies* $\mathbb{E}\left[\|u_T\|_1^2\right]<\infty$*, i.e.,* $u_T\in L^2\left(\Omega,\mathcal{F}_T;\dot{H}^1\right)$*. Then, the BSPDE* (2.3) *admits a unique weak solution satisfying*

$$\sup_{0\le t\le T}\mathbb{E}\left[\|u(t)\|^2\right]+\mathbb{E}\left[\int_0^T\|u(t)\|_1^2+\|v(t)\|^2dt\right]<\infty.\tag{2.5}$$

Moreover, if Assumption 2.3 *also holds and* $u_T\in L^2(\Omega,\mathcal{F}_T;\dot{H}^2)$*, then we have*

$$(u,v)\in\Big(L^2\big(\Omega;C([0,T];\dot{H}^1)\big)\cap L^2_{\mathcal{F}}((0,T);\dot{H}^2)\Big)\times L^2_{\mathcal{F}}((0,T);\dot{H}^1),$$

that is

$$\mathbb{E}\left[\sup_{0\le t\le T}\|u(t)\|_1^2\right]+\mathbb{E}\left[\int_0^T\|u(t)\|_2^2+\|v(t)\|_1^2dt\right]<\infty.\tag{2.6}$$

We also provide the following lemma:

Lemma 2.1. *Under Assumptions* 2.1*, the following assertions hold:*

- *The bilinear form* (2.1) *is bounded on* $\dot{H}^1$*: there exist* $\alpha>0$ *such that*

$$|a(u,\phi)|\le\alpha\|u\|_1\|\phi\|_1,\quad \forall u,\phi\in\dot{H}^1.$$

- *The bilinear form* (2.1) *satisfies the coercivity condition: there exists* $\beta>0$ *such that*

$$a(u,u)\ge\beta\|u\|_1^2,\quad \forall u\in\dot{H}^1.$$

We now close this section by providing the following Itô's formula.

Lemma 2.2 ([6, 18])**.** *Consider the Itô process*

$$X(t)=X(0)+\int_0^t b(s)ds+\int_0^t \sigma(s)dW(s), \quad t\in[0,T],$$

where $X(0)$ is a $\mathcal{F}_0$-measurable H-valued random variable, $b\in L^1_{\mathcal{F}}((0,T);H)$ and $\sigma\in L^2_{\mathcal{F}}((0,T);H)$. Assume that $F:[0,T]\times H\to\mathbb{R}$ and its partial derivatives F_t, F_x and F_{xx} are uniformly continuous on bounded subsets of $[0,T]\times H$, then it holds almost surely

$$\begin{aligned}F(t,X(t))=&F(0,X(0))+\int_0^t \big(F_x(s,X(s)),\sigma(s)\big)dW(s)\\&+\int_0^t \Big[F_t(s,X(s))+(F_x(s,X(s)),b(s))+\frac{1}{2}(F_{xx}(s,X(s))\sigma(s),\sigma(s))\Big]ds. \end{aligned}\tag{2.7}$$

3 A semi-discrete scheme and its error estimates

To introduce the semi-discrete scheme, we let $\mathcal{T}_h$ be a triangulation mesh of D indexed by the maximal mesh size h. Furthermore, we assume that $\mathcal{T}_h$ is quasi-uniform. On the mesh $\mathcal{T}_h$, we introduce a finite element space S_h consisting of continuous piecewise linear polynomials such that $S_h\subset\dot{H}^1$.

To proceed, we first introduce some useful operators on S_h and H [13,25]. Let $P_h: H\to S_h$ be the L^2-projection operator defined by

$$(P_h u,\chi)=(u,\chi), \quad \forall\chi\in S_h \quad \text{for } u\in H. \tag{3.1}$$

Also, we denote by A_h the discrete elliptic operator $A_h: S_h\to S_h$:

$$(A_h\varphi,\chi)=a(\varphi,\chi), \quad \forall\varphi,\chi\in S_h\subset\dot{H}^1. \tag{3.2}$$

The so-called Ritz-projection operator $R_h: \dot{H}^1\to S_h$ is defined by

$$a(R_h\varphi,\chi)=a(\varphi,\chi), \quad \forall\varphi\in\dot{H}^1, \quad \chi\in S_h. \tag{3.3}$$

Now, we are ready to introduce the semi-discrete finite element scheme. By Theorem 2.6 and the definition of weak solution, we have

$$\begin{aligned}&\big(u(t),\phi\big)+\int_t^T a\big(u(s),\phi\big)\,ds\\=&(u_T,\phi)+\int_t^T \Big(f\big(s,\nabla u(s),u(s),v(s)\big),\phi\Big)\,ds\\&-\int_t^T \big(v(s),\phi\big)\,dW(s), \quad \forall\phi\in\dot{H}^1, \quad t\in[0,T]. \end{aligned}\tag{3.4}$$

Based on the above weak formulation (3.4), the semi-discrete finite element scheme for BSPDEs (2.3) is defined as following: find $u_h \in L^2\big(\Omega;C([0,T];S_h)\big)$ and $v_h \in L^2_{\mathcal{F}}((0,T);S_h)$ such that

$$\begin{aligned}
&\big(u_h(t),\phi_h\big)+\int_t^T a\big(u_h(s),\phi_h\big)\,ds\\
=&(u_T,\phi_h)+\int_t^T \Big(f\big(s,\nabla u_h(s),u_h(s),v_h(s)\big),\phi_h\Big)\,ds\\
&-\int_t^T \big(v_h(s),\phi_h\big)\,dW(s), \quad \forall \phi_h \in S_h, \quad t\in[0,T].
\end{aligned}\tag{3.5}$$

The above scheme (3.5) can also be rewritten as:

$$-du_h(t)+A_hu_h(t)\,dt=P_hf\big(t,\nabla u_h(t),u_h(t),v_h(t)\big)\,dt-v_h(t)\,dW(t) \tag{3.6}$$

with $u_h(T)=P_hu_T$. Obviously, (3.6) is a BSDE system. Due to the boundedness of the operator P_h and the Lipschitz condition on f, the BSDE system admits a unique solution $(u_h,v_h)\in L^2\big(\Omega;C([0,T];S_h)\times L^2_{\mathcal{F}}((0,T);S_h)$ (see e.g., [19]). To give the error estimates for the semi-discrete scheme, we first introduce the following lemma:

Lemma 3.1 ([5, 13, 25])**.** *The operators A_h, P_h and R_h enjoy the following properties:*

1. *The operator A_h is self-adjoint and positive definite on S_h.*

2. *For the operator R_h, under the regularity assumptions on the finite element space, the following standard error estimates hold:*

$$\|R_hu-u\|+h\|R_hu-u\|_1\le Ch^s\|u\|_s, \quad \forall u\in \dot{H}^s, \quad 1\le s\le 2. \tag{3.7}$$

3. *For the operator P_h, under the regularity assumptions on the finite element space, it holds that:*

$$\|P_hu-u\|\le Ch^s\|u\|_s, \qquad \forall u\in \dot{H}^s, \quad 0\le s\le 2, \tag{3.8a}$$

$$\|P_hu\|_1\le C\|u\|_1, \qquad \forall u\in \dot{H}^1. \tag{3.8b}$$

Notice that by (3.7) and (3.8b), we have

$$h\|A^{1/2}(P_hu-u)\|=h\|P_hu-u\|_1\le Ch^s\|u\|_s, \quad \forall u\in \dot{H}^s, \quad 1\le s\le 2. \tag{3.9}$$

To see this, $\forall \chi\in S_h$ we have

$$\begin{aligned}
h\|A^{1/2}(P_hu-u)\|=&h\|P_hu-u\|_1=h\|P_hu-\chi+\chi-u\|_1\\
\le&h\|P_h(u-\chi)\|_1+h\|\chi-u\|_1\le Ch\|\chi-u\|_1.
\end{aligned}\tag{3.10}$$

Then, the conclusion (3.9) follows by setting $\chi=R_hu$.

We are now ready to give the following theorem which gives the error estimate for the semi-discrete approximation (3.6).

Theorem 3.1. *Let (u,v) and (u_h,v_h) be the solutions of* (2.3) *and* (3.6)*, respectively. We set*

$$e_u(t)=u_h(t)-u(t),\quad e_v(t)=v_h(t)-v(t).$$

Suppose that Assumptions 2.1-2.3 *hold. If $u_T\in L^2(\Omega,\mathcal{F}_T;\dot{H}^2)$, then there exists a constant C that is independent of h such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|e_u(t)\|^2\right]+\mathbb{E}\left[\int_0^T\|e_u(t)\|_1^2dt\right]+\mathbb{E}\left[\int_0^T\|e_v(t)\|^2dt\right]\le Ch^2. \tag{3.11}$$

Proof. We set $\phi=\phi_h$ in (3.4) and subtract (3.4) from (3.5) to get

$$\begin{aligned}&\big(u_h(t)-u(t),\phi_h\big)+\int_t^T a\big(u_h(s)-u(s),\phi_h\big)ds\\ =&\int_t^T\Big(f\big(s,\nabla u_h(s),u_h(s),v_h(s)\big)-f\big(s,\nabla u(s),u(s),v(s)\big),\phi_h\Big)ds\\ &-\int_t^T\big(v_h(s)-v(s),\phi_h\big)dW(s),\quad\forall\phi_h\in S_h.\end{aligned} \tag{3.12}$$

We define

$$\begin{aligned}&e_v(t)=\theta_v(t)+\rho_v(t),\\ &e_u(t)=\theta_u(t)+\rho_u(t)=\tilde{\theta}_u(t)+\tilde{\rho}_u(t),\end{aligned}$$

where

$$\begin{aligned}&\theta_v(t)=v_h(t)-P_hv(t), &&\rho_v(t)=P_hv(t)-v(t),\\ &\theta_u(t)=u_h(t)-P_hu(t), &&\rho_u(t)=P_hu(t)-u(t),\\ &\tilde{\theta}_u(t)=u_h(t)-R_hu(t), &&\tilde{\rho}_u(t)=R_hu(t)-u(t).\end{aligned}$$

Notice that by the definitions of the project operators R_h and P_h we have

$$(\rho_u(t),\phi_h)=0,\quad a(\tilde{\rho}_u(t),\phi_h)=0\quad\text{and}\quad(\rho_v(t),\phi_h)=0,\quad\forall\phi_h\in S_h. \tag{3.13}$$

Then, by (3.12) and (3.13), we have

$$\begin{aligned}&\big(\theta_u(t),\phi_h\big)+\int_t^T a\big(\tilde{\theta}_u(s),\phi_h\big)ds\\ =&\int_t^T\Big(f\big(s,\nabla u_h(s),u_h(s),v_h(s)\big)-f\big(s,\nabla u(s),u(s),v(s)\big),\phi_h\Big)ds\\ &-\int_t^T\big(\theta_v(s),\phi_h\big)dW(s),\quad\forall\phi_h\in S_h,\end{aligned} \tag{3.14}$$

which yields

$$\begin{aligned}&-d\theta_u(t)+A_h\tilde{\theta}_u(t)dt\\=&P_h\Big(f\big(t,\nabla u_h(t),u_h(t),v_h(t)\big)-f\big(t,\nabla u(t),u(t),v(t)\big)\Big)dt-\theta_v(t)dW(t)\end{aligned}\tag{3.15}$$

with $\theta_u(T)=0$. Now, by applying the Itô formula (2.7) to $\|\theta_u(t)\|^2$ we get

$$\begin{aligned}&\|\theta_u(t)\|^2+\int_t^T\|\theta_v(s)\|^2ds\\=&-2\int_t^T\Big(A_h\tilde{\theta}_u(s),\theta_u(s)\Big)ds-2\int_t^T\big(\theta_v(s),\theta_u(s)\big)dW(s)\\&+2\int_t^T\bigg(P_h\Big(f(s,\nabla u_h(s),u_h(s),v_h(s))-f(s,\nabla u(s),u(s),v(s))\Big),\theta_u(s)\bigg)ds.\end{aligned}$$

By the definition of the operator A_h, we have

$$\begin{aligned}&\|\theta_u(t)\|^2+\int_t^T\|\theta_v(s)\|^2ds\\=&-2\int_t^Ta\big(\tilde{\theta}_u(s),\theta_u(s)\big)ds-2\int_t^T\big(\theta_v(s),\theta_u(s)\big)dW(s)\\&+2\int_t^T\bigg(P_h\Big(f(s,\nabla u_h(s),u_h(s),v_h(s))-f(s,\nabla u(s),u(s),v(s))\Big),\theta_u(s)\bigg)ds\\=&2\int_t^Ta\big(\tilde{\theta}_u(s),\rho_u(s)\big)ds-2\int_t^Ta\big(\tilde{\theta}_u(s),\theta_u(s)+\rho_u(s)\big)ds\\&+2\int_t^T\bigg(P_h\Big(f(s,\nabla u_h(s),u_h(s),v_h(s))-f(s,\nabla u(s),u(s),v(s))\Big),\theta_u(s)\bigg)ds\\&-2\int_t^T\big(\theta_v(s),\theta_u(s)\big)dW(s).\end{aligned}\tag{3.16}$$

Notice that $\theta_u(t)+\rho_u(t)=e_u(t)=\tilde{\theta}_u(t)+\tilde{\rho}_u(t)$. Consequently, by replacing $\theta_u(t)+\rho_u(t)$ with $\tilde{\theta}_u(t)+\tilde{\rho}_u(t)$ in (3.16) and taking expectation on both sides one gets

$$\begin{aligned}&\mathbb{E}\left[\|\theta_u(t)\|^2\right]+\int_t^T\mathbb{E}\left[\|\theta_v(s)\|^2\right]ds\\=&2\int_t^T\mathbb{E}\left[a(\tilde{\theta}_u(s),\rho_u(s))\right]ds-2\int_t^T\mathbb{E}\left[\|\tilde{\theta}_u(s)\|_1^2\right]ds-2\int_t^T\mathbb{E}\left[a(\tilde{\theta}_u(s),\tilde{\rho}_u(s))\right]ds\\&+2\int_t^T\mathbb{E}\bigg[\bigg(P_h\Big(f(s,\nabla u_h(s),u_h(s),v_h(s))-f(s,\nabla u(s),u(s),v(s))\Big),\theta_u(s)\bigg)\bigg]ds.\end{aligned}$$

Since

$$a\big(\tilde{\theta}_u(s),\rho_u(s)\big)-a\big(\tilde{\theta}_u(s),\tilde{\rho}_u(s)\big)=-a\big(\tilde{\theta}_u(s),P_h\tilde{\rho}_u(s)\big),$$

we obtain

$$\begin{aligned}
&\mathbb{E}\left[\|\theta_u(t)\|^2\right]+\int_t^T\mathbb{E}\left[\|\theta_v(s)\|^2\right]ds\\
=&-2\int_t^T\mathbb{E}\left[a(\tilde{\theta}_u(s),P_h\tilde{\rho}_u(s))\right]ds-2\int_t^T\mathbb{E}\left[\|\tilde{\theta}_u(s)\|_1^2\right]ds\\
&+2\int_t^T\mathbb{E}\left[\Big(P_h\left(f\left(s,\nabla u_h(s),u_h(s),v_h(s)\right)-f\left(s,\nabla u(s),u(s),v(s)\right)\right),\theta_u(s)\Big)\right]ds.
\end{aligned}$$

Then using together Lemma 2.1 and the Cauchy–Schwarz inequality we obtain

$$\begin{aligned}
&\mathbb{E}\left[\|\theta_u(t)\|^2\right]+\mathbb{E}\left[\int_t^T\|\theta_v(s)\|^2ds\right]\\
\leq&2\alpha\mathbb{E}\left[\int_t^T\|\tilde{\theta}_u(s)\|_1\|P_h\tilde{\rho}_u(s)\|_1ds\right]-2\mathbb{E}\left[\int_t^T\|\tilde{\theta}_u(s)\|_1^2ds\right]\\
&+2\mathbb{E}\left[\int_t^T\Big\|P_h\Big(f\left(s,\nabla u_h(s),u_h(s),v_h(s)\right)-f\left(s,\nabla u(s),u(s),v(s)\right)\Big)\Big\|\|\theta_u(s)\|ds\right].
\end{aligned}$$

Now by using together (3.8b) and the Lipschitz condition on f, we deduce

$$\begin{aligned}
&\mathbb{E}\Big[\|\theta_u(t)\|^2\Big]+\mathbb{E}\Big[\int_t^T\|\theta_v(s)\|^2ds\Big]\\
\leq&2\alpha\mathbb{E}\Big[\int_t^T\|\tilde{\theta}_u(s)\|_1\|\tilde{\rho}_u(s)\|_1ds\Big]-2\mathbb{E}\Big[\int_t^T\|\tilde{\theta}_u(s)\|_1^2ds\Big]\\
&+C\mathbb{E}\Big[\int_t^T\|\nabla u_h(s)-\nabla u(s)\|\|\theta_u(s)\|ds\Big]+C\mathbb{E}\Big[\int_t^T\|u_h(s)-u(s)\|\|\theta_u(s)\|ds\Big]\\
&+C\mathbb{E}\Big[\int_t^T\|v_h(s)-v(s)\|\|\theta_u(s)\|ds\Big]\\
=&2\alpha\mathbb{E}\Big[\int_t^T\|\tilde{\theta}_u(s)\|_1\|\tilde{\rho}_u(s)\|_1ds\Big]-2\mathbb{E}\Big[\int_t^T\|\tilde{\theta}_u(s)\|_1^2ds\Big]\\
&+C\mathbb{E}\Big[\int_t^T\|\nabla\theta_u(s)+\nabla\rho_u(s)\|\|\theta_u(s)\|ds\Big]+C\mathbb{E}\Big[\int_t^T\|\theta_u(s)+\rho_u(s)\|\|\theta_u(s)\|ds\Big]\\
&+C\mathbb{E}\Big[\int_t^T\|\theta_v(s)+\rho_v(s)\|\|\theta_u(s)\|ds\Big].
\end{aligned}$$

Applying the triangle inequality yields

$$\begin{aligned}
&\mathbb{E}\left[\|\theta_u(t)\|^2\right]+\mathbb{E}\left[\int_t^T\|\theta_v(s)\|^2ds\right]\\
\leq&2\alpha\mathbb{E}\left[\int_t^T\|\tilde{\theta}_u(s)\|_1\|\tilde{\rho}_u(s)\|_1ds\right]-2\mathbb{E}\left[\int_t^T\|\tilde{\theta}_u(s)\|_1^2ds\right]\\
&+C\mathbb{E}\left[\int_t^T\|\theta_u(s)\|^2ds\right]+C\mathbb{E}\left[\int_t^T\|\rho_u(t)\|\|\theta_u(s)\|ds\right]
\end{aligned}$$

$$
\begin{aligned}
&+C\mathbb{E}\left[\int_t^T \|\nabla\theta_u(s)\|\|\theta_u(s)\|ds\right]+C\mathbb{E}\left[\int_t^T \|\nabla\rho_u(s)\|\|\theta_u(s)\|ds\right]\\
&+C\mathbb{E}\left[\int_t^T \|\theta_v(s)\|\|\theta_u(s)\|ds\right]+C\mathbb{E}\left[\int_t^T \|\rho_v(s)\|\|\theta_u(s)\|ds\right].
\end{aligned}
$$

Then by using together a kick-back argument, Lemma 3.1 and (3.9) we get

$$
\begin{aligned}
&\mathbb{E}\left[\|\theta_u(t)\|^2\right]+\mathbb{E}\left[\int_t^T\|\tilde{\theta}_u(s)\|_1^2ds\right]+\mathbb{E}\left[\int_t^T\|\theta_v(s)\|^2ds\right]\\
\leq& C\left(\mathbb{E}\left[\int_t^T\|\tilde{\rho}_u(s)\|_1^2ds\right]+\mathbb{E}\left[\int_t^T\|\rho_u(s)\|^2ds\right]+\mathbb{E}\left[\int_t^T\|\rho_v(s)\|^2ds\right]\right)\\
&+C\mathbb{E}\left[\int_t^T\|\theta_u(s)\|^2ds\right]\\
\leq& C\left(h^2+\mathbb{E}\left[\int_t^T\|\theta_u(s)\|^2ds\right]\right).
\end{aligned}
$$

Then by the Gronwall inequality one gets

$$
\mathbb{E}\left[\|\theta_u(t)\|^2\right]+\mathbb{E}\left[\int_0^T\|\tilde{\theta}_u(s)\|_1^2ds\right]+\mathbb{E}\left[\int_0^T\|\theta_v(s)\|^2ds\right]\leq Ch^2. \tag{3.17}
$$

Similarly, we replace $\theta_u(t)+\rho_u(t)$ in (3.16) by $\tilde{\theta}_u(t)+\tilde{\rho}_u(t)$ to obtain

$$
\begin{aligned}
&\sup_{t\in[0,T]}\|\theta_u(t)\|^2\\
\leq& 2\int_0^T\left\|a(\tilde{\theta}_u(s),P_h\tilde{\rho}_u(s))\right\|ds+2\int_0^T\left\|\tilde{\theta}_u(s)\right\|_1^2ds\\
&+2\int_0^T\left\|\left(P_h\Big(f\big(s,\nabla u_h(s),u_h(s),v_h(s)\big)-f\big(s,\nabla u(s),u(s),v(s)\big)\Big),\theta_u(s)\right)\right\|ds\\
&+2\sup_{t\in[0,T]}\left\|\int_t^T(\theta_v(s),\theta_u(s))dW(s)\right\|,
\end{aligned}
$$

where we have used the following argument

$$
a\big(\tilde{\theta}_u(s),\rho_u(s)\big)-a\big(\tilde{\theta}_u(s),\tilde{\rho}_u(s)\big)=-a\big(\tilde{\theta}_u(s),P_h\tilde{\rho}_u(s)\big).
$$

Then by taking expectation on both sides of the above equation we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\|\theta_u(t)\|^2\right]\\
\leq& 2\mathbb{E}\left[\int_0^T\|a(\tilde{\theta}_u(s),P_h\tilde{\rho}_u(s))\|ds\right]+2\mathbb{E}\left[\int_0^T\|\tilde{\theta}_u(s)\|_1^2ds\right]
\end{aligned}
$$

$$
\begin{aligned}
&+2\mathbb{E}\left[\int_0^T\left\|\Big(P_h\big(f\big(s,\nabla u_h(s),u_h(s),v_h(s)\big)-f\big(s,\nabla u(s),u(s),v(s)\big)\big),\theta_u(s)\Big)\right\|ds\right]\\
&+2\mathbb{E}\left[\sup_{t\in[0,T]}\left\|\int_t^T(\theta_v(s),\theta_u(s))dW(s)\right\|\right].
\end{aligned}
$$

Now, by applying the Burkholder–Davis–Gundy inequality [6, 17] to the integral that involves $W(s)$ one gets

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\|\theta_u(t)\|^2\right]\\
\leq&2\mathbb{E}\left[\int_0^T\|a(\tilde{\theta}_u(s),P_h\tilde{\rho}_u(s))\|ds\right]+2\mathbb{E}\left[\int_0^T\|\tilde{\theta}_u(s)\|_1^2ds\right]\\
&+2\mathbb{E}\left[\int_0^T\left\|\Big(P_h\big(f\big(s,\nabla u_h(s),u_h(s),v_h(s)\big)-f\big(s,\nabla u(s),u(s),v(s)\big)\big),\theta_u(s)\Big)\right\|ds\right]\\
&+C\mathbb{E}\left[\left(\int_0^T\|\theta_v(s)\|^2\|\theta_u(s)\|^2ds\right)^{\frac{1}{2}}\right].
\end{aligned}
$$

Again, by using together Lemma 2.1, the Cauchy–Schwarz inequality, (3.8b), the Lipschitz condition on f and the triangle inequality, we get

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}\|\theta_u(t)\|^2\Big]\\
\leq&2\alpha\mathbb{E}\Big[\int_0^T\|\tilde{\theta}_u(s)\|_1\|\tilde{\rho}_u(s)\|_1ds\Big]+2\mathbb{E}\Big[\int_0^T\|\tilde{\theta}_u(s)\|_1^2ds\Big]\\
&+C\mathbb{E}\Big[\int_0^T\|\nabla\theta_u(s)\|\|\theta_u(s)\|ds\Big]+C\mathbb{E}\Big[\int_0^T\|\nabla\rho_u(s)\|\|\theta_u(s)\|ds\Big]\\
&+C\mathbb{E}\Big[\int_0^T\|\theta_u(s)\|^2ds\Big]+C\mathbb{E}\Big[\int_0^T\|\rho_u(s)\|\|\theta_u(s)\|ds\Big]\\
&+C\mathbb{E}\Big[\int_0^T\|\theta_v(s)\|\|\theta_u(s)\|ds\Big]+C\mathbb{E}\Big[\int_0^T\|\rho_v(s)\|\|\theta_u(s)\|ds\Big]\\
&+C\mathbb{E}\left[\left(\int_0^T\|\theta_v(s)\|^2\|\theta_u(s)\|^2ds\right)^{\frac{1}{2}}\right].
\end{aligned}
$$

Now, by the ε-inequality $ab\leq\varepsilon a^2+\frac{b^2}{4\varepsilon}$ we get

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}\|\theta_u(t)\|^2\right]\\
\leq&C\mathbb{E}\left[\int_0^T\|\rho_u(s)\|^2ds\right]+C\mathbb{E}\left[\int_0^T\|\tilde{\rho}_u(s)\|_1^2ds\right]+C\mathbb{E}\left[\int_0^T\|\rho_v(s)\|^2ds\right]
\end{aligned}
$$

$$+C\int_0^T \mathbb{E}\left[\|\theta_u(s)\|^2\right]ds+C\mathbb{E}\left[\int_0^T \|\tilde{\theta}_u(s)\|_1^2 ds\right]+C\mathbb{E}\left[\int_0^T \|\theta_v(s)\|^2 ds\right]$$
$$+\varepsilon\mathbb{E}\left[\sup_{t\in[0,T]}\|\theta_u(t)\|^2\right].$$

Then, by choosing a small $\varepsilon<1$, and using together Lemma 3.1, (3.9) and (3.17) we get

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|\theta_u(t)\|^2\right]\le Ch^2. \tag{3.18}$$

Consequently, by Lemma 3.1 we obtain

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|e_u(t)\|^2\right]\le\mathbb{E}\left[\sup_{t\in[0,T]}\|\theta_u(t)\|^2\right]+\mathbb{E}\left[\sup_{t\in[0,T]}\|\rho_u(t)\|^2\right]\le Ch^2. \tag{3.19}$$

Meanwhile, by (3.17) and Lemma 3.1 we also have

$$\mathbb{E}\left[\int_0^T \|e_u(t)\|_1^2 dt\right]\le\mathbb{E}\left[\int_0^T \|\tilde{\theta}_u(t)\|_1^2 dt\right]+\mathbb{E}\left[\int_0^T \|\tilde{\rho}_u(t)\|_1^2 dt\right]\le Ch^2, \tag{3.20a}$$
$$\mathbb{E}\left[\int_0^T \|e_v(t)\|^2 dt\right]\le\mathbb{E}\left[\int_0^T \|\theta_v(t)\|^2 dt\right]+\mathbb{E}\left[\int_0^T \|\rho_v(t)\|^2 dt\right]\le Ch^2. \tag{3.20b}$$

This completes the proof. □

4 A fully discrete scheme and its error estimates

In this section, we shall analyze the convergence of a fully discrete scheme for BSPDEs (2.3). To this end, we introduce the following partition for the time interval $[0,T]$:

$$\mathcal{T}=:\{0=t_0<\cdots<t_M=T\}$$

with $\Delta t_n=:t_{n+1}-t_n$ and $\Delta t=:\max_{0\le n\le M-1}\Delta t_n$. We assume that the time partition $\mathcal{T}$ has the following regularity:

$$\frac{\max\limits_{0\le n\le M-1}\Delta t_n}{\min\limits_{0\le n\le M-1}\Delta t_n}\le C_0, \tag{4.1}$$

where C_0 is a positive constant. Now, by the weak formulation we have (for $0\le n\le M-1$)

$$\begin{aligned}&\big(u(t_n),\phi\big)+\int_{t_n}^{t_{n+1}}a\big(u(s),\phi\big)ds\\ =&(u(t_{n+1}),\phi)+\int_{t_n}^{t_{n+1}}\Big(f\big(s,\nabla u(s),u(s),v(s)\big),\phi\Big)ds\\ &-\int_{t_n}^{t_{n+1}}\big(v(s),\phi\big)dW(s),\quad\forall\phi\in\dot{H}^1.\end{aligned} \tag{4.2}$$

Taking conditional expectation $\mathbb{E}_{t_n}^{y}[\cdot]=\mathbb{E}[\cdot|\mathcal{F}_{t_n},W(t_n)=y]$ on both sides of (4.2) we get

$$\begin{aligned}\big(u(t_n),\phi\big)=&\mathbb{E}_{t_n}^{y}\left[(u(t_{n+1}),\phi)\right]+\bigg(\int_{t_n}^{t_{n+1}}\mathbb{E}_{t_n}^{y}\Big[-Au(s)\\&+f\big(s,\nabla u(s),u(s),v(s)\big)\Big]ds,\phi\bigg).\end{aligned}\tag{4.3}$$

Notice that the integrand $\mathbb{E}_{t_n}^{y}\left[-Au(s)+f\big(s,\nabla u(s),u(s),v(s)\big)\right]$ on the right hand side of (4.3) is a deterministic smooth function of time s. Thus we can resort to numerical integration to approximate the integral in (4.3). In particular, we use the left rectangle rule to obtain

$$\begin{aligned}&\int_{t_n}^{t_{n+1}}\mathbb{E}_{t_n}^{y}\left[-Au(s)+f\big(s,\nabla u(s),u(s),v(s)\big)\right]ds\\=&\Delta t_n\left[-Au(t_n)+f\big(t_n,\nabla u(t_n),u(t_n),v(t_n)\big)\right]+R_u^n,\end{aligned}\tag{4.4}$$

where R_u^n is the approximation error defined by

$$\begin{aligned}R_u^n=&\int_{t_n}^{t_{n+1}}\mathbb{E}_{t_n}^{y}\left[-Au(s)+f\big(s,\nabla u(s),u(s),v(s)\big)\right]ds\\&-\Delta t_n\left[-Au(t_n)+f\big(t_n,\nabla u(t_n),u(t_n),v(t_n)\big)\right].\end{aligned}$$

By inserting (4.4) into (4.3) we obtain the following reference equation:

$$\begin{aligned}&\Big(u(t_n),\phi\Big)+\Delta t_n\Big(Au(t_n),\phi\Big)\\=&\mathbb{E}_{t_n}^{y}\Big[(u(t_{n+1}),\phi)\Big]+\Delta t_n\Big(f\big(t_n,\nabla u(t_n),u(t_n),v(t_n)\big),\phi\Big)+\big(R_u^n,\phi\big).\end{aligned}\tag{4.5}$$

Next, we set $\Delta W_s=W(s)-W(t_n)$ for $t_n\le s\le t_{n+1}$, then ΔW_s is a standard Wiener process with mean zero and variance $s-t_n$. By multiplying both sides of (4.2) with $\Delta W_{t_{n+1}}$ and taking the conditional expectation $\mathbb{E}_{t_n}^{y}[\cdot]$ on both sides of the derived equation, we obtain

$$\begin{aligned}0=&\mathbb{E}_{t_n}^{y}\left[(u(t_{n+1}),\phi)\Delta W_{t_{n+1}}\right]\\&+\Big(\int_{t_n}^{t_{n+1}}\mathbb{E}_{t_n}^{y}\left[-\big(Au(s)\Delta W_s+f\big(s,\nabla u(s),u(s),v(s)\big)\Delta W_s\right]ds,\phi\Big)\\&-\Big(\int_{t_n}^{t_{n+1}}\mathbb{E}_{t_n}^{y}\left[v(s)\right]ds,\phi\Big).\end{aligned}\tag{4.6}$$

Using again the left rectangle formula to discretize the two temporal integrals on the right hand side of (4.6) we obtain another reference equation

$$0=\mathbb{E}_{t_n}^{y}\Big[\Big((u(t_{n+1}),\phi\Big)\Delta W_{t_{n+1}}\Big]-\Delta t_n\Big(v(t_n),\phi\Big)+\Big(R_v^n,\phi\Big),\tag{4.7}$$

where the truncation error R_v^n is defined by

$$\begin{aligned}R_v^n = &\int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^y \big[-Au(s)\Delta W_s + f\big(s,\nabla u(s),u(s),v(s)\big)\Delta W_s\big] ds \\ &- \int_{t_n}^{t_{n+1}} \mathbb{E}_{t_n}^y \big[v(s) - v(t_n)\big] ds.\end{aligned}$$

Based on above two reference equations (4.5) and (4.7), we are now ready to present the fully discrete finite element method for BSPDEs (2.3) as follows.

Scheme 4.1. Given the terminal condition (u_h^M, v_h^M), we solve a pair of $\mathcal{F}_{t_n}$-adapted S_h-valued processes $\{(u_h^n, v_h^n)\}_{0\le n\le M-1}$ by

$$\begin{aligned}&\big(u_h^n,\phi_h\big) + \Delta t_n a\big(u_h^n,\phi_h\big) \\ &= \mathbb{E}_{t_n}^y \Big[\Big((u_h^{n+1},\phi_h \Big) \Big] + \Delta t_n \Big(f\big(t_n,\nabla u_h^n, u_h^n, v_h^n\big),\phi_h \Big), \qquad (4.8a)\\ &\mathbb{E}_{t_n}^y \Big[\Big((u_h^{n+1},\phi_h \Big) \Delta W_{t_{n+1}} \Big] - \Delta t_n \Big(v_h^n,\phi_h \Big) = 0. \qquad (4.8b)\end{aligned}$$

Since (4.8a) and (4.8b) hold for all $\phi_h \in S_h$, consequently, (4.8a) and (4.8b) can also be rewritten as

$$u_h^n + \Delta t_n A_h u_h^n = \mathbb{E}_{t_n}^y \Big[u_h^{n+1} \Big] + \Delta t_n P_h f\big(t_n,\nabla u_h^n,u_h^n,v_h^n\big), \qquad (4.9a)$$

$$\mathbb{E}_{t_n}^y \Big[u_h^{n+1}\Delta W_{t_{n+1}} \Big] - \Delta t_n v_h^n = 0. \qquad (4.9b)$$

4.1 Error estimates for the fully discrete scheme

In this section, we shall present the error estimates for the fully discrete scheme. To begin, let $(u(t_n),v(t_n))$ and (u_h^n,v_h^n) be the exact solution and the numerical solution of BSPDEs (2.3) at $t=t_n$, respectively. We then define

$$\begin{aligned}&e_u^n = u(t_n) - u_h^n = u(t_n) - P_h u(t_n) + P_h u(t_n) - u_h^n =: \rho_u^n + \theta_u^n,\\ &e_v^n = v(t_n) - v_h^n = v(t_n) - P_h v(t_n) + P_h v(t_n) - v_h^n =: \rho_v^n + \theta_v^n,\\ &e_f^n = f\big(t_n,\nabla u(t_n),u(t_n),v(t_n)\big) - f\big(t_n,\nabla u_h^n,u_h^n,v_h^n\big).\end{aligned}$$

By the definition of P_h, we have

$$(\rho_u^n,\phi_h) = 0 \quad \text{and} \quad (\rho_v^n,\phi_h) = 0, \quad \forall \phi_h \in S_h. \qquad (4.10)$$

We now present the following stability theorem that implies our error estimates.

Theorem 4.1. *Under similar assumptions as in Theorem* 2.6*, for sufficient small* Δt*, we have*

$$\begin{aligned}
&\mathbb{E}\Big[\|e_u^n\|^2\Big]+\sum_{i=n}^{M-1}\Delta t_n\mathbb{E}\Big[\Big\|e_u^i\Big\|_1^2\Big]+\sum_{i=n}^{M-1}\Delta t_n\mathbb{E}\Big[\Big\|e_v^i\Big\|^2\Big]\\
\lesssim&\mathbb{E}\Big[\Big\|e_u^M\Big\|^2\Big]+\sum_{i=n}^{M-1}\Delta t\mathbb{E}\Big[\Big\|\rho_u^i\Big\|_1^2\Big]+\sum_{i=n}^{M-1}\Delta t\mathbb{E}\Big[\Big\|\rho_u^i\Big\|_1^2\Big]+\sum_{i=n}^{M-1}\Delta t\mathbb{E}\Big[\Big\|\rho_v^i\Big\|^2\Big]\\
&+\sum_{i=n}^{M-1}\Delta t\Big(\mathbb{E}\Big[\Big\|\rho_u^i\Big\|^2\Big]+\mathbb{E}\Big[\Big\|\rho_v^i\Big\|^2\Big]\Big)+\sum_{i=n}^{M-1}\frac{1}{\Delta t}\mathbb{E}\Big[\Big\|R_v^i\Big\|^2\Big]\\
&+\sum_{i=n}^{M-1}\frac{1}{\Delta t}\mathbb{E}\Big[\Big\|R_u^i\Big\|^2\Big],
\end{aligned}\tag{4.11}$$

where $a\lesssim b$ *stands for* $a\leq Cb$ *with* C *being a positive constant.*

Proof. By (4.5) and (4.7), we have for $\forall\phi_h\in S_h$

$$\begin{aligned}
&\Big(u(t_n),\phi_h\Big)+\Delta t_n a\Big(u(t_n),\phi_h\Big)\\
&=\Delta t_n\Big(f\big(t_n,\nabla u(t_n),u(t_n),v(t_n)\big),\phi_h\Big)+\mathbb{E}_{t_n}^y\Big[\Big(u(t_{n+1}),\phi_h\Big)\Big]+\Big(R_u^n,\phi_h\Big),
\end{aligned}\tag{4.12a}$$

$$0=\mathbb{E}_{t_n}^y\Big[\Big((u(t_{n+1}),\phi_h\Big)\Delta W_{t_{n+1}}\Big]-\Delta t_n\Big(v(t_n),\phi_h\Big)+\Big(R_v^n,\phi_h\Big).\tag{4.12b}$$

Subtracting (4.8a) from (4.12a) leads to

$$\begin{aligned}
&\Big(u(t_n)-u_h^n,\phi_h\Big)+\Delta t_n a\Big(u(t_n)-u_h^n,\phi_h\Big)\\
=&\mathbb{E}_{t_n}^y\Big[\Big(u(t_{n+1})-u_h^{n+1},\phi_h\Big)\Big]+\Delta t_n\Big(e_f^n,\phi_h\Big)+\Big(R_u^n,\phi_h\Big),
\end{aligned}$$

which yields

$$\begin{aligned}
&\Big(\theta_u^n,\phi_h\Big)+\Delta t_n a\Big(\theta_u^n,\phi_h\Big)\\
=&-\Delta t_n a\Big(\rho_u^n,\phi_h\Big)+\mathbb{E}_{t_n}^y\Big[\Big(\theta_u^{n+1},\phi_h\Big)\Big]+\Delta t_n\Big(e_f^n,\phi_h\Big)+\Big(R_u^n,\phi_h\Big).
\end{aligned}\tag{4.13}$$

By setting $\phi_h=\theta_u^n$ in (4.13), and using together Lemma 2.1 and the property of conditional expectation we have

$$\begin{aligned}
\|\theta_u^n\|^2+\Delta t_n\beta\|\theta_u^n\|_1^2\leq&-\Delta t_n a\Big(\rho_u^n,\theta_u^n\Big)+\mathbb{E}_{t_n}^y\Big[\Big(\theta_u^{n+1},\theta_u^n\Big)\Big]+\Delta t_n\Big(e_f^n,\theta_u^n\Big)+\Big(R_u^n,\theta_u^n\Big)\\
\leq&-\Delta t_n a\Big(\rho_u^n,\theta_u^n\Big)+\Big(\mathbb{E}_{t_n}^y\Big[\theta_u^{n+1}\Big],\theta_u^n\Big)+\Delta t_n\Big(e_f^n,\theta_u^n\Big)+\Big(R_u^n,\theta_u^n\Big)\\
\leq&-\Delta t_n a\Big(\rho_u^n,\theta_u^n\Big)+\Big(\mathbb{E}_{t_n}^y\Big[\theta_u^{n+1}\Big]+\Delta t_n e_f^n+R_u^n,\theta_u^n\Big)\\
\leq&\Delta t_n\alpha\|\rho_u^n\|_1\cdot\|\theta_u^n\|_1+\Big\|\mathbb{E}_{t_n}^y\Big[\theta_u^{n+1}\Big]+\Delta t_n e_f^n+R_u^n\Big\|\cdot\|\theta_u^n\|.
\end{aligned}$$

Then, by using the inequality $ab\leq\frac{a^2}{2\beta}+\frac{\beta b^2}{2}$ we get

$$\begin{aligned}\|\theta_u^n\|^2+\beta\Delta t_n\|\theta_u^n\|_1^2\leq&\frac{\Delta t_n\alpha^2\|\rho_u^n\|_1^2}{2\beta}+\frac{\Delta t_n\beta\|\theta_u^n\|_1^2}{2}\\&+\frac{1}{2}\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]+\Delta t_n e_f^n+R_u^n\right\|^2+\frac{1}{2}\|\theta_u^n\|^2,\end{aligned}$$

which is equivalent to

$$\frac{1}{2}\|\theta_u^n\|^2+\frac{\Delta t_n\beta\|\theta_u^n\|_1^2}{2}\leq\frac{\Delta t_n\alpha^2\|\rho_u^n\|_1^2}{2\beta}+\frac{1}{2}\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]+\Delta t_n e_f^n+R_u^n\right\|^2. \tag{4.14}$$

Thus we have

$$\|\theta_u^n\|^2+\Delta t_n\beta\|\theta_u^n\|_1^2\leq\hat{C}\Delta t_n\|\rho_u^n\|_1^2+\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]+\Delta t_n e_f^n+R_u^n\|^2,\quad \hat{C}=\frac{\alpha^2}{\beta}. \tag{4.15}$$

By the inequalities $(a+b)^2\leq a^2+b^2+\gamma\Delta t a^2+\frac{1}{\gamma\Delta t}b^2$ and $(a+b)^2\leq 2(a^2+b^2)$, we derive

$$\begin{aligned}&\|\theta_u^n\|^2+\Delta t_n\beta\|\theta_u^n\|_1^2\\\leq&\hat{C}\Delta t_n\|\rho_u^n\|_1^2+(1+\gamma\Delta t)\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]\right\|^2+\left(1+\frac{1}{\gamma\Delta t}\right)\left\|\Delta t_n e_f^n+R_n^u\right\|^2\\\leq&\Delta t_n\|\rho_u^n\|_1^2+(1+\gamma\Delta t)\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]\right\|^2+\left(2+\frac{2}{\gamma\Delta t}\right)\left(\|\Delta t_n e_f^n\|^2+\|R_u^n\|^2\right).\end{aligned} \tag{4.16}$$

Now by applying the Lipschitz condition on f one gets

$$\begin{aligned}&\|\theta_u^n\|^2+\Delta t_n\beta\|\theta_u^n\|_1^2\\\leq&\hat{C}\Delta t_n\|\rho_u^n\|_1^2+(1+\gamma\Delta t)\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]\right\|^2\\&+\left(2+\frac{2}{\gamma\Delta t}\right)\left(3L^2(\Delta t_n)^2\left(\|e_u^n\|_1^2+\|e_u^n\|^2+\|e_v^n\|^2\right)+\|R_n^u\|^2\right)\\\leq&\hat{C}\Delta t_n\|\rho_u^n\|_1^2+(1+\gamma\Delta t)\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]\right\|^2+\left(2+\frac{2}{\gamma\Delta t}\right)\|R_u^n\|^2\\&+\left(6+\frac{6}{\gamma\Delta t}\right)L^2(\Delta t_n)^2\left(\|\rho_u^n+\theta_u^n\|_1^2+\|\rho_u^n+\theta_u^n\|^2+\|\rho_v^n+\theta_v^n\|^2\right)\\\leq&\hat{C}\Delta t_n\|\rho_u^n\|_1^2+(1+\gamma\Delta t)\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]\right\|^2+\left(2+\frac{2}{\gamma\Delta t}\right)\|R_u^n\|^2\\&+\left(12+\frac{12}{\gamma\Delta t}\right)L^2(\Delta t_n)^2\left(\|\theta_u^n\|_1^2+\|\theta_u^n\|^2+\|\theta_v^n\|^2\right)\\&+\left(12+\frac{12}{\gamma\Delta t}\right)L^2(\Delta t_n)^2\left(\|\rho_u^n\|_1^2+\|\rho_u^n\|^2+\|\rho_v^n\|^2\right).\end{aligned} \tag{4.17}$$

Similarly, subtracting (4.8b) from (4.12b) leads to

$$\Delta t_n\Big(v(t_n)-v_h^n,\phi_h\Big)=\mathbb{E}_{t_n}^{y}\Big[\Big((u(t_{n+1})-u_h^{n+1},\phi_h\Big)\Delta W_{t_{n+1}}\Big]+\Big(R_v^n,\phi_h\Big).$$

Then we have

$$\Delta t_n\Big(\theta_v^n,\phi_h\Big)=\mathbb{E}_{t_n}^{y}\Big[\Big(\theta_u^{n+1},\phi_h\Big)\Delta W_{t_{n+1}}\Big]+\Big(R_v^n,\phi_h\Big). \tag{4.18}$$

We set $\phi_h=\theta_v^n$ in the above equation to get

$$\begin{aligned}\Delta t_n\|\theta_v^n\|^2=&\mathbb{E}_{t_n}^{y}\Big[\Big(\theta_u^{n+1},\theta_v^n\Big)\Delta W_{t_{n+1}}\Big]+\Big(R_v^n,\theta_v^n\Big)\\ =&\Big(\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Delta W_{t_{n+1}}\Big],\theta_v^n\Big)+\Big(R_v^n,\theta_v^n\Big)\\ \le&\Big\|\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Delta W_{t_{n+1}}\Big]+R_v^n\Big\|\|\theta_v^n\|\\ \le&\frac{1}{\varepsilon}\Big\|\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Delta W_{t_{n+1}}\Big]+R_v^n\Big\|^2+\frac{\varepsilon}{4}\|\theta_v^n\|^2\\ \le&\frac{1}{\varepsilon}(1+\gamma)\Big\|\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Delta W_{t_{n+1}}\Big]\Big\|^2+\frac{1}{\varepsilon}\Big(1+\frac{1}{\gamma}\Big)\|R_v^n\|^2+\frac{\varepsilon}{4}\|\theta_v^n\|^2.\end{aligned} \tag{4.19}$$

In the above derivations, we have used the inequalities $ab\le\frac{1}{\varepsilon}a^2+\frac{\varepsilon}{4}b^2$ and $(a+b)^2\le(1+\gamma)a^2+(1+\frac{1}{\gamma})b^2$. Notice that

$$\begin{aligned}\Big\|\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Delta W_{t_{n+1}}\Big]\Big\|^2=&\Big\|\mathbb{E}_{t_n}^{y}\Big[\Big(\theta_u^{n+1}-\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Big]\Big)\Delta W_{t_{n+1}}\Big]\Big\|^2\\ \le&\Delta t_n\Big(\mathbb{E}_{t_n}^{y}\Big[\|\theta_u^{n+1}\|^2\Big]-\Big\|\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Big]\Big\|^2\Big).\end{aligned}$$

Then, by (4.19) we get

$$\begin{aligned}\Delta t_n\|\theta_v^n\|^2\le&\frac{1}{\varepsilon}(1+\gamma)\Delta t_n\Big(\mathbb{E}_{t_n}^{y}\Big[\|\theta_u^{n+1}\|^2\Big]-\Big\|\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Big]\Big\|^2\Big)\\ &+\frac{1}{\varepsilon}\Big(1+\frac{1}{\gamma}\Big)\|R_v^n\|^2+\frac{\varepsilon}{4}\|\theta_v^n\|^2.\end{aligned} \tag{4.20}$$

We set $\varepsilon=\Delta t_n$ in the above inequality to get

$$\begin{aligned}\Delta t_n\|\theta_v^n\|^2\le&(1+\gamma)\Big(\mathbb{E}_{t_n}^{y}\Big[\Big\|\theta_u^{n+1}\Big\|^2\Big]-\Big\|\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Big]\Big\|^2\Big)\\ &+\frac{1}{\Delta t_n}\Big(1+\frac{1}{\gamma}\Big)\|R_v^n\|^2+\frac{\Delta t_n}{4}\|\theta_v^n\|^2.\end{aligned}$$

Then we have

$$\frac{3\Delta t_n}{4}\|\theta_v^n\|^2\le(1+\gamma)\Big(\mathbb{E}_{t_n}^{y}\Big[\Big\|\theta_u^{n+1}\Big\|^2\Big]-\Big\|\mathbb{E}_{t_n}^{y}\Big[\theta_u^{n+1}\Big]\Big\|^2\Big)+\frac{1}{\Delta t_n}\Big(1+\frac{1}{\gamma}\Big)\|R_v^n\|^2,$$

which yields

$$\begin{aligned}\|\theta_v^n\|^2\le&\frac{4(1+\gamma)}{3\Delta t_n}\left(\mathbb{E}_{t_n}^y\left[\left\|\theta_u^{n+1}\right\|^2\right]-\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]\right\|^2\right)\\&+\frac{4}{3(\Delta t_n)^2}\left(1+\frac{1}{\gamma}\right)\|R_v^n\|^2.\end{aligned}\tag{4.21}$$

Dividing both sides of the inequality (4.21) by $\frac{4(1+\gamma)}{3\Delta t}$, we obtain

$$\begin{aligned}\frac{3\Delta t}{4(1+\gamma)}\|\theta_v^n\|^2\le&\frac{\Delta t}{\Delta t_n}\left(\mathbb{E}_{t_n}^y\left[\left\|\theta_u^{n+1}\right\|^2\right]-\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]\right\|^2\right)+\frac{\Delta t}{\gamma(\Delta t_n)^2}\|R_v^n\|^2\\\le&C_0\left(\mathbb{E}_{t_n}^y\left[\left\|\theta_u^{n+1}\right\|^2\right]-\left\|\mathbb{E}_{t_n}^y\left[\theta_u^{n+1}\right]\right\|^2\right)+\frac{\Delta t}{\gamma(\Delta t_n)^2}\|R_v^n\|^2.\end{aligned}\tag{4.22}$$

Now, we multiply (4.17) by C_0 and add it with (4.22) to obtain

$$\begin{aligned}&C_0\left\|\theta_u^n\right\|^2+C_0\beta\Delta t_n\left\|\theta_u^n\right\|_1^2+\frac{3\Delta t}{4(1+\gamma)}\left\|\theta_v^n\right\|^2\\\le&C_0\hat{C}\Delta t_n\|\rho_u^n\|_1^2+C_0(1+\gamma\Delta t)\mathbb{E}_{t_n}^y\left[\left\|\theta_u^{n+1}\right\|^2\right]\\&+C_0\left(12+\frac{12}{\gamma\Delta t}\right)L^2(\Delta t_n)^2\left(\|\theta_u^n\|_1^2+\|\theta_u^n\|^2+\|\theta_v^n\|^2\right)\\&+C_0\left(12+\frac{12}{\gamma\Delta t}\right)L^2(\Delta t_n)^2\left(\|\rho_u^n\|_1^2+\|\rho_u^n\|^2+\|\rho_v^n\|^2\right)\\&+C_0\left(2+\frac{2}{\gamma\Delta t}\right)\|R_u^n\|^2+\frac{\Delta t}{\gamma(\Delta t_n)^2}\|R_v^n\|^2,\end{aligned}$$

which implies

$$\begin{aligned}&C_0(1-C_1\Delta t)\left\|\theta_u^n\right\|^2+C_2\Delta t_n\left\|\theta_u^n\right\|_1^2+C_3\Delta t\left\|\theta_v^n\right\|^2\\\le&C_0(1+\gamma\Delta t)\mathbb{E}_{t_n}^y\left[\left\|\theta_u^{n+1}\right\|^2\right]+C_5\Delta t\|\rho_u^n\|_1^2+C_6\Delta t\left(\|\rho_u^n\|^2+\|\rho_v^n\|^2\right)\\&+C_0(2+2/(\gamma\Delta t))\|R_u^n\|^2+1/(\gamma\Delta t_n)\|R_v^n\|^2,\end{aligned}\tag{4.23}$$

where

$$\begin{aligned}&C_1=\left(12\Delta t+\frac{12}{\gamma}\right)L^2,\quad C_2=C_0\beta-\left(12C_0\Delta t_n+\frac{12C_0}{\gamma}\right)L^2,\\&C_3=\frac{3}{4(1+\gamma)}-\left(12C_0\Delta t+\frac{12C_0}{\gamma}\right)L^2,\\&C_5=C_0\hat{C}+C_0\left(12+\frac{12}{\gamma\Delta t}\right)L^2\Delta t_n,\quad C_6=C_0\left(12\Delta t+\frac{12}{\gamma}\right)L^2.\end{aligned}$$

Now we choose a large enough γ and sufficient small Δt such that it holds $C_1 \le C^*, \gamma \le C^*, 1-C^*\Delta t > 0$, where C^* is positive constant. Then we obtain by (4.23)

$$\begin{aligned}
&C_0(1-C^*\Delta t)\big\|\theta_u^n\big\|^2 + C_2\Delta t_n\big\|\theta_u^n\big\|_1^2 + C_3\Delta t\big\|\theta_v^n\big\|^2 \\
\le& C_0(1+C^*\Delta t)\mathbb{E}_{t_n}^y\Big[\big\|\theta_u^{n+1}\big\|^2\Big] + C_5\Delta t\big\|\rho_u^n\big\|_1^2 + C_6\Delta t\Big(\big\|\rho_u^n\big\|^2 + \big\|\rho_v^n\big\|^2\Big) \\
&+ C_0\left(2+\frac{2}{\gamma\Delta t}\right)\big\|R_u^n\big\|^2 + \frac{\Delta t}{\gamma(\Delta t_n)^2}\big\|R_v^n\big\|^2. \qquad (4.24)
\end{aligned}$$

Dividing both sides of the inequality (4.24) by $1-C^*\Delta t$ and taking mathematical expectation gives

$$\begin{aligned}
&C_0\mathbb{E}\Big[\big\|\theta_u^n\big\|^2\Big] + C_2\Delta t_n\mathbb{E}\Big[\big\|\theta_u^n\big\|_1^2\Big] + C_3\Delta t\mathbb{E}\Big[\big\|\theta_v^n\big\|^2\Big] \\
\le& \frac{1+C\Delta t}{1-C^*\Delta t}C_0\mathbb{E}\Big[\big\|\theta_u^{n+1}\big\|^2\Big] + \frac{C_5\Delta t}{1-C^*\Delta t}\mathbb{E}\Big[\big\|\rho_u^n\big\|_1^2\Big] \\
&+ \frac{C_6\Delta t}{1-C^*\Delta t}\Big(\mathbb{E}\Big[\big\|\rho_u^n\big\|^2\Big] + \mathbb{E}\Big[\big\|\rho_v^n\big\|^2\Big]\Big) \\
&+ \frac{2C_0(1+\gamma\Delta t)\mathbb{E}\Big[\big\|R_u^n\big\|^2\Big]}{\gamma(1-C^*\Delta t)\Delta t} + \frac{\Delta t\mathbb{E}\Big[\big\|R_v^n\big\|^2\Big]}{\gamma(1-C^*\Delta t)(\Delta t_n)^2}. \qquad (4.25)
\end{aligned}$$

Then by induction we get

$$\begin{aligned}
&\mathbb{E}\Big[\big\|\theta_u^n\big\|^2\Big] + \sum_{i=n}^{M-1}\left(\frac{1+C^*\Delta t}{1-C^*\Delta t}\right)^{i-n}\Delta t_n\mathbb{E}\Big[\big\|\theta_u^i\big\|_1^2\Big] + \sum_{i=n}^{M-1}\left(\frac{1+C^*\Delta t}{1-C^*\Delta t}\right)^{i-n}\Delta t\mathbb{E}\Big[\big\|\theta_v^i\big\|^2\Big] \\
\lesssim& \left(\frac{1+C^*\Delta t}{1-C^*\Delta t}\right)^{M-n}C_0\mathbb{E}\Big[\big\|\theta_u^M\big\|^2\Big] + \sum_{i=n}^{M-1}\left(\frac{1+C^*\Delta t}{1-C^*\Delta t}\right)^{i-n}\frac{C_5\Delta t}{1-C^*\Delta t}\mathbb{E}\Big[\big\|\rho_u^i\big\|_1^2\Big] \\
&+ \sum_{i=n}^{M-1}\left(\frac{1+C^*\Delta t}{1-C^*\Delta t}\right)^{i-n}\frac{C_6\Delta t}{1-C^*\Delta t}\Big(\mathbb{E}\Big[\big\|\rho_u^i\big\|^2\Big] + \mathbb{E}\Big[\big\|\rho_v^i\big\|^2\Big]\Big) \\
&+ \sum_{i=n}^{M-1}\left(\frac{1+C^*\Delta t}{1-C^*\Delta t}\right)^{i-n}\frac{2C_0(1+\gamma\Delta t)\mathbb{E}\Big[\big\|R_u^i\big\|^2\Big]}{\gamma(1-C^*\Delta t)\Delta t} + \sum_{i=n}^{M-1}\left(\frac{1+C^*\Delta t}{1-C^*\Delta t}\right)^{i-n}\frac{\Delta t\mathbb{E}\Big[\big\|R_v^i\big\|^2\Big]}{\gamma(1-C^*\Delta t)(\Delta t_n)^2} \\
\lesssim& \exp\Big\{2C^*C_0T\Big\}C_0\mathbb{E}\Big[\big\|\theta_u^M\big\|^2\Big] + \exp\Big\{2C^*C_0T\Big\}\sum_{i=n}^{M-1}\frac{C_5\Delta t}{1-C^*\Delta t}\mathbb{E}\Big[\big\|\rho_u^i\big\|_1^2\Big] \\
&+ \exp\Big\{2C^*C_0T\Big\}\sum_{i=n}^{M-1}\frac{C_6\Delta t}{1-C^*\Delta t}\Big(\mathbb{E}\Big[\big\|\rho_u^i\big\|^2\Big] + \mathbb{E}\Big[\big\|\rho_v^i\big\|^2\Big]\Big) \\
&+ \exp\Big\{2C^*C_0T\Big\}\sum_{i=n}^{M-1}\frac{2C_0(1+\gamma\Delta t)\mathbb{E}\Big[\big\|R_u^i\big\|^2\Big]}{\gamma(1-C^*\Delta t)\Delta t} + \exp\Big\{2C^*C_0T\Big\}\sum_{i=n}^{M-1}\frac{\Delta t\mathbb{E}\Big[\big\|R_v^i\big\|^2\Big]}{\gamma(1-C^*\Delta t)(\Delta t_n)^2}.
\end{aligned}$$

Thus we finally get

$$
\begin{aligned}
&\mathbb{E}\left[\|e_u^n\|^2\right]+\sum_{i=n}^{M-1}\left(\frac{1+C^*\Delta t}{1-C^*\Delta t}\right)^{i-n}\Delta t_n\mathbb{E}\left[\|e_u^i\|_1^2\right]+\sum_{i=n}^{M-1}\left(\frac{1+C^*\Delta t}{1-C^*\Delta t}\right)^{i-n}\Delta t_n\mathbb{E}\left[\|e_v^i\|^2\right]\\
\lesssim&\exp\left\{2C^*C_0T\right\}C_0\mathbb{E}\left[\|e_u^M\|^2\right]+\exp\left\{2C^*C_0T\right\}\sum_{i=n}^{M-1}\frac{C_5\Delta t}{1-C^*\Delta t}\mathbb{E}\left[\|\rho_u^i\|_1^2\right]\\
&+C_2\exp\left\{2C^*C_0T\right\}\sum_{i=n}^{M-1}\Delta t\mathbb{E}\left[\|\rho_u^i\|_1^2\right]+\exp\left\{2C^*C_0T\right\}\sum_{i=n}^{M-1}\frac{C_6\Delta t}{1-C^*\Delta t}\Big(\mathbb{E}\left[\|\rho_u^i\|^2\right]\\
&+C_3\exp\left\{2C^*C_0T\right\}\sum_{i=n}^{M-1}\Delta t\mathbb{E}\left[\|\rho_v^i\|^2\right]+\mathbb{E}\left[\|\rho_v^i\|^2\right]\Big)\\
&+\exp\left\{2C^*C_0T\right\}\sum_{i=n}^{M-1}\frac{2C_0(1+\gamma\Delta t)\mathbb{E}\left[\|R_u^i\|^2\right]}{\gamma(1-C^*\Delta t)\Delta t}\\
&+\exp\left\{2C^*C_0T\right\}\sum_{i=n}^{M-1}\frac{\Delta t\mathbb{E}\left[\|R_v^i\|^2\right]}{\gamma(1-C^*\Delta t)(\Delta t_n)^2}.
\end{aligned}
$$

Then the desired inequality (4.11) follows by noticing the boundedness of the involved constants. □

By using the Itô–Taylor expansion and the properties of the Wiener process, we have the following lemma, whose proof is quite similar to those certain lemmas in [14,29,31].

Lemma 4.1. *Suppose f and u_T are sufficiently smooth functions, let R_u^n and R_v^n be the truncation errors defined in* (4.5) *and* (4.7)*, respectively. Then, it holds that*

$$\|R_u^n\|\le C(\Delta t)^2 \quad \textit{and} \quad \|R_v^n\|\le C(\Delta t)^2. \tag{4.26}$$

The above lemma, together with the stability theorem leads to the following error estimate for the fully discrete scheme (4.8a)-(4.8b).

Theorem 4.2. *Suppose that the conditions in Theorem* 2.6 *and Lemma* 4.1 *hold, then for any $0\le n\le M-1$, we have*

$$\mathbb{E}\left[\|e_u^n\|^2\right]+\sum_{i=1}^{M-1}\Delta t_n\mathbb{E}\left[\|e_u^i\|_1^2\right]+\sum_{i=1}^{M-1}\Delta t_n\mathbb{E}\left[\|e_v^i\|^2\right]\le C\Big(h^2+(\Delta t)^2\Big). \tag{4.27}$$

Proof. By Lemma 3.1 and the estimate 3.9, together with Theorem 4.1 and Lemma 4.1, the desired result follows. □

Remark 4.1. The above error analysis for Theorem 3.1 and Theorem 4.2 can be extended immediately to finite elements of higher order, i.e., the finite element space S_h consists of piecewise polynomials of degree $r-1$, where $r\geq 2$ is an integer. In addition, suppose that the solution of (2.3) satisfies the following regularity

$$(u,v)\in\Big(L^2\big(\Omega;C([0,T];\dot{H}^{r-1})\big)\cap L^2_{\mathcal{F}}((0,T);\dot{H}^r)\Big)\times L^2_{\mathcal{F}}((0,T);\dot{H}^{r-1}).$$

Then, we can obtain the following estimates

$$\mathbb{E}\left[\sup_{t\in[0,T]}\|e_u(t)\|^2\right]+\mathbb{E}\left[\int_0^T\|e_u(t)\|_1^2dt\right]+\mathbb{E}\left[\int_0^T\|e_v(t)\|^2dt\right]\leq Ch^{2(r-1)}, \tag{4.28a}$$

$$\mathbb{E}\left[\|e_u^n\|^2\right]+\Delta t\sum_{i=1}^{M-1}\mathbb{E}\left[\|e_u^i\|_1^2\right]+\Delta t\sum_{i=1}^{M-1}\mathbb{E}\left[\|e_v^i\|^2\right]\leq C\big(h^{2(r-1)}+(\Delta t)^2\big). \tag{4.28b}$$

5 Conclusions

We proposed a spatial finite element semi-discrete scheme and a spatio-temporal full discrete scheme for solving nonlinear backward stochastic partial differential equations. In these schemes, finite element methods in the physical space and the implicit Euler method in time domain were used, respectively. We rigorously analyzed the errors of the schemes and obtained error estimates with convergence rates. In our future work, we shall investigate computational issues of the schemes for BSPDEs, and study other high accurate time discretization methods for BSPDEs.

Acknowledgements

This research is partially supported by the Science Challenge Project (No. TZ2018001), by National Key R&D Plan (No. 2018YFA0703903) and by the National Natural Science Foundations of China (under Grants Nos. 11901565, 11571206, 11831010 and 11871068). The authors would like to thank the referees for the helpful comments on the improvement of the present paper.

References

[1] A. R. AL-HUSSEIN, *Strong, mild and weak solutions of backward stochastic evolution equations,* Random Oper. Stochastic Equations, 13 (2005), pp. 129–138.

[2] C. BENDER AND J. ZHANG, *Time discretization and markovian iteration for coupled FBSDEs,* Ann. Appl. Probab., 18 (2008), pp. 143–177.

[3] A. BENSOUSSAN, *Stochastic maximum principle for distributed parameter systems,* J. Franklin Inst., 315 (1983), pp. 387–406.

[4] 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.

[5] M. CROUZEIX AND V. THOMÉE, *The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces,* Math. Comput., 48 (1987), pp. 521–532.

[6] G. DA PRATO AND J. ZABCZYK, Stochastic Equations in Infinite Dimensions, vol. 152 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, second ed., 2014.

[7] K. DU AND S. TANG, *Strong solution of backward stochastic partial differential equations in C^2 domains,* Probab. Theory Related Fields, 154 (2012), pp. 255–285.

[8] T. DUNST AND A. PROHL, *The forward-backward stochastic heat equation: numerical analysis and simulation,* SIAM J. Sci. Comput., 38 (2016), pp. A2725–A2755.

[9] M. FUHRMAN AND Y. HU, *Infinite horizon BSDEs in infinite dimensions with continuous driver and applications,* J. Evol. Equ., 6 (2006), pp. 459–484.

[10] E. GOBET, J. P. LEMMOR, AND X. WARIN, *A regression-based monte carlo method to solve backward stochastic differential equations,* Ann. Appl. Probab., 15 (2005), pp. 2172–2202.

[11] Y. HU AND S. G. PENG, *Maximum principle for semilinear stochastic evolution control systems,* Stochastics Stochastics Rep., 33 (1990), pp. 159–180.

[12] Y. HU AND S. G. PENG, *Adapted solution of a backward semilinear stochastic evolution equation,* Stochastic Anal. Appl., 9 (1991), pp. 445–459.

[13] R. KRUSE, Strong and Weak Approximation of Semilinear Stochastic Evolution Equations, vol. 2093 of Lecture Notes in Mathematics, Springer, Cham, 2014.

[14] 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.

[15] Y. LIU, Y. SUN, AND W. ZHAO, *A fully discrete explicit multistep scheme for solving coupled forward backward stochastic differential equations,* Adv. Appl. Math. Mech., 12 (2020), pp. 643–663.

[16] J. MA AND J. YONG, *Adapted solution of a degenerate backward SPDE, with applications,* Stochastic Process. Appl., 70 (1997), pp. 59–84.

[17] X. MAO, Stochastic differential equations and their applications, Horwood Publishing Series in Mathematics & Applications, Horwood Publishing Limited, Chichester, 1997.

[18] E. PARDOUX, *Stochastic partial differential equations and filtering of diffusion processes,* Stochastics, 3 (1979), pp. 127–167.

[19] E. PARDOUX AND S. G. PENG, *Adapted solution of a backward stochastic differential equation,* Systems Control Lett., 14 (1990), pp. 55–61.

[20] S. G. PENG, *Stochastic Hamilton-Jacobi-Bellman equations,* SIAM J. Control Optim., 30 (1992), pp. 284–304.

[21] M. RUIJTER AND C. OOSTERLEE, *A fourier cosine method for an efficient computation of solutions to BSDEs,* SIAM J. Sci. Comput., 37 (2015), pp. A859–A889.

[22] S. TANG AND W. WEI, *On the Cauchy problem for backward stochastic partial differential equations in Hölder spaces,* Ann. Probab., 44 (2016), pp. 360–398.

[23] T. TANG, W. ZHAO, AND T. ZHOU, *Deferred correction methods for forward backward stochastic differential equations,* Numer. Math. Theor. Meth. Appl., 10 (2017), pp. 222–242.

[24] G. TESSITORE, *Existence, uniqueness and space regularity of the adapted solutions of a backward SPDE,* Stochastic Anal. Appl., 14 (1996), pp. 461–486.

[25] V. THOMÉE, Galerkin Finite Element Methods for Parabolic Problems, vol. 25 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, second ed., 2006.

[26] Y. WANG, *A semidiscrete Galerkin scheme for backward stochastic parabolic differential equations,* Math. Control Relat. Fields, 6 (2016), pp. 489–515.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] W. ZHAO, W. ZHANG, AND L. JU, *A numerical method and its error estimates for the decoupled forward-backward stochastic differential equations,* Commun. Comput. Phys., 15 (2014), pp. 618–646.

[32] 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.

[33] X. Y. ZHOU, *A duality analysis on stochastic partial differential equations,* J. Funct. Anal., 103 (1992), pp. 275–293.

[34] X. Y. ZHOU, *On the necessary conditions of optimal controls for stochastic partial differential equations,* SIAM J. Control Optim., 31 (1993), pp. 1462–1478.