SIAM J. NUMER. ANAL.
Vol. 55, No. 6, pp. 2982–3005

AN EFFICIENT GRADIENT PROJECTION METHOD FOR STOCHASTIC OPTIMAL CONTROL PROBLEMS*

BO GONG†, WENBIN LIU‡, TAO TANG§, WEIDONG ZHAO¶, AND TAO ZHOU‖

Abstract. In this work, we propose a simple yet effective gradient projection algorithm for a class of stochastic optimal control problems. We first reduce the optimal control problem to an optimization problem for a convex functional by means of a projection operator. Then we propose a convergent iterative scheme for the optimization problem. The key issue in our iterative scheme is to compute the gradient of the objective functional by solving the adjoint equations that are given by backward stochastic differential equations (BSDEs). The Euler method is used to solve the resulting BSDEs. Rigorous convergence analysis is presented, and it is shown that the entire numerical algorithm admits a first order rate of convergence. Several numerical examples are carried out to support the theoretical finding.

Key words. stochastic optimal control, gradient projection methods, backward stochastic differential equations, conditional expectations

AMS subject classifications. 60H35, 65C20, 35Q99, 35R35

DOI. 10.1137/17M1123559

1. Introduction.

In recent years, stochastic optimal control has been extensively studied and has become an essential tool in various fields, such as financial mathematics and engineering. There exists a very extensive body of literature in both theoretical and practical studies of stochastic optimal control problems (SOCPs); see, e.g., [4, 5, 20, 28, 8, 11, 18, 17] and references therein.

In this work, we are concerned with numerical solutions of SOCPs. Theoretical investigations for SOCPs can be found in [4, 13, 20, 27, 3, 7, 15, 29, 35, 38]. For practical applications of SOCPs, one can refer to [7, 26, 29, 38, 40] for engineering applications, to [24, 25, 32, 42, 45] for applications in option pricing and portfolio optimization, to [1] for analysis of climate changes, and to [19] for biological and medical problems, to name a few.

In general, the SOCP does not admit explicitly closed form solutions, and thus efficient numerical algorithms have been widely studied in recent years. Roughly speaking, we can characterize numerical algorithms into four categories: (i) transferring the control problem into finite dimensional stochastic programming (see, e.g., [9, 15, 21, 22, 29, 38, 41, 43]); (ii) the dynamic programming principle (DPP) based

*Received by the editors March 30, 2017; accepted for publication (in revised form) September 1, 2017; published electronically November 28, 2017.
http://www.siam.org/journals/sinum/55-6/M112355.html
Funding: This work was supported by the National Natural Science Foundations of China under grants 91630312, 91630203, 11571351, 11171189, 11731006, and 11571206. The last author is also supported by NCMIS. The third and first authors are supported by the Hong Kong Research Grant Council and Hong Kong Baptist University.

†Department of Mathematics, The Hong Kong Baptist University, Hong Kong, China (13479245@life.hkbu.edu.hk).
‡Kent Business School, University of Kent, Canterbury CT2 7PE, UK (W.B.Liu@kent.ac.uk).
§Department of Mathematics, Southern University of Sciences and Technology, Shenzhen, Guangdong 518055, China (tangt@sustc.edu.cn).
¶School of Mathematics & Finance Institute, Shandong University, Jinan, Shandong 250100, China (wdzhao@sdu.edu.cn).
‖LSEC, Institute of Computational Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (tzhou@lsec.cc.ac.cn).

approach [6, 28], wherein one usually needs to solve the corresponding Hamilton–Jacobi–Bellman (HJB) equations, and this is one of the most widely used numerical methods [2, 4, 5, 10, 23]; (iii) martingale based methods [24, 25, 39]; and (iv) stochastic maximum principle (SMP) based methods (see, e.g., [20] and references therein).

Basically, the SMP procedure is to directly compute the directional derivative for the objective functional $J(\cdot)$ by introducing an adjoint process. Then by introducing an optimality condition for the control problem, a variational inequality coupled with the state and adjoint equations forms an optimality condition system (we call it the SMP system) that can be used to solve the optimal control problem. While SMP is a popular tool for theoretical studies of stochastic optimal control (see, e.g., [42, 45]), it has not been widely used in the numerical setting.

In this work, we propose a simple yet effective gradient projection algorithm for SOCPs. We first reduce the optimal control problems to an optimization problem for a convex functional by means of a projection operator. Then we propose a convergent iterative scheme for the optimization problem. The key idea in our iterative scheme is to compute the gradient of the objective functional in an efficient way, and this is done by solving the adjoint equations that are given by backward stochastic differential equations (BSDEs). Our approach belongs to the SMP based approach category, and it relies on solving the SMP system in an efficient way. To this end, we propose a simple yet effective Euler-type method for solving the resulting BSDEs. Furthermore, we perform a sharp convergence analysis, and we show that our numerical method admits a first order rate of convergence. Several numerical examples are presented to support the theoretical finding.

The rest of the paper is organized as follows. In section 2 we set up the SOCP and provide some assumptions. The gradient projection method is presented in section 3. Section 4 is devoted to convergence analysis of the proposed numerical approach. Several numerical experiments are presented to show the effectiveness of the proposed numerical method in section 5. We finally give some conclusions in section 6.

2. Problem setup.

For notational simplicity, we shall narrow our discussion to the one dimensional case; however, the whole framework applies easily to multi-dimensional cases. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leqslant t\leqslant T}, \mathrm{P})$ be a complete probability space with filtration $\mathcal{F}_t$ generated by the Brownian motion $\{W_s\}_{0\leqslant s\leqslant t}$. Here T is the terminal time. We denote by $U = L^2([0,T];\mathbb{R})$ the space of all square integrable functions $x : [0,T] \mapsto \mathbb{R}$ and denote by $U_{\mathcal{F}} = \mathbb{L}^2_{\mathcal{F}}([0,T]\times\Omega;\mathbb{R})$ the space of all adapted stochastic processes $x = x_t : [0,T]\times\Omega \mapsto \mathbb{R}$ that satisfy

$$\mathbb{E}\Big[\int_0^T (x_t)^2\,dt\Big] < +\infty.$$

We let $\mathcal{C} \subset \mathbb{R}$ be a nonempty, convex, and closed subset, and we define the following control set:

$$K = \big\{u \in U \big|\ u(t) \in \mathcal{C} \text{ a.e.}\big\}.$$

Note that we have assumed that the control u is deterministic. We remark that a deterministic control can still be useful for future planning as discussed, e.g., in [7, 29] for engineering applications, in [9] for financial applications, and in [38] for an application in stochastic hybrid systems. Moreover, stochastic control (i.e., $u \in U_{\mathcal{F}}$) can also be included in our approach, and this will be discussed in section 5 via numerical examples.

Given $u \in K$, the controlled state process x_t^u is governed by the stochastic differential equation (SDE)

$$dx_t^u = b\big(x_t^u, u(t)\big)dt + \sigma\big(x_t^u, u(t)\big)dW_t, \quad t \in (0, T], \quad x|_{t=0} = x_0 \in \mathbb{R}. \tag{1}$$

The considered cost functional is given by

$$J(u) = \mathbb{E}\left[\int_0^T \Big(h(x_t^u) + j\big(u(t)\big)\Big)dt + k(x_T^u)\right], \tag{2}$$

where $h(\cdot), j(\cdot), k(\cdot)$ are given functions and x_t^u is the solution of (1). We now state our SOCP as follows:

$$\text{Find} \quad u^* \in K \quad \text{such that} \quad J(u^*) = \min_{u \in K} J(u). \tag{3}$$

Throughout the paper, we shall make the following assumption.

Assumption 2.1.

- *The functions $b = b(x, u)$ and $\sigma = \sigma(x, u)$ are continuously differentiable with respect to x and u and have bounded derivatives.*
- *The functions h, j, and k are continuously differentiable, and their derivatives have at most a linear growth with respect to the underlying variables.*

Notice that under Assumption 2.1, the solution x_t^u of (1) and the cost functional $J(u)$ are all well defined for $u \in K$.

3. The gradient projection method.

In this section, we will present details of our gradient projection method. For the SOCP (1)–(3), it is well known that for the optimal control u^* it holds that

$$(J'(u^*), v - u^*) \geqslant 0 \qquad \forall v \in K, \tag{4}$$

where $(J'(u), v)$ is the variation of $J(u)$ along the direction v; i.e., for $v \in U$ such that $u + v \in K$, we have

$$(J'(u), v) = \lim_{\rho \downarrow 0} \frac{J(u + \rho v) - J(u)}{\rho}. \tag{5}$$

The existence of such derivatives has been discussed in [13, 35, 42]. Here we slightly abuse the notation by referring to $J'(u)$ as the associated element in U (its respresentation in U by the unique mapping), as U is a Hilbert space.

Next, we propose a gradient projection method for solving the optimality condition (4). To this end, let $\|\cdot\|$ be the norm of U. We define the projection operator $P_K\colon \omega \mapsto P_K\omega$ as

$$\big\|P_K\omega - \omega\big\| = \min_{u \in K} \|u - \omega\|. \tag{6}$$

Notice that the projection problem (6) is equivalent to the inequality

$$(P_K\omega - \omega, v - P_K\omega) \geqslant 0 \qquad \forall v \in K. \tag{7}$$

For any positive constant ρ, the variational inequality (4) is equivalent to the following inequality:

$$\big(u^* - (u^* - \rho J'(u^*)), v - u^*\big) \geqslant 0 \qquad \forall v \in K. \tag{8}$$

By the fact of well-posedness of convex optimizations and by comparing the above inequality with the inequality (7), we conclude that for the optimal control u^*, it holds that

$$u^* = P_K(u^* - \rho J'(u^*)). \tag{9}$$

That is, the optimal control u^* is the fixed point of $P_K(u - \rho J'(u))$ on K.

We shall approximate the control u^* numerically by step functions. To this end, we introduce the following uniform time partition:

$$0 = t_0^N < t_1^N < \cdots < t_N^N = T, \quad t_{n+1}^N - t_n^N = T/N =: \Delta t. \tag{10}$$

We will denote by I_n^N the intervals $[t_{n-1}^N, t_n^N)$ for $1 \leqslant n \leqslant N-1$, and by I_N^N the interval $[t_{N-1}^N, t_N^N]$. In the context where N is fixed, we shall omit the superscript N of t_n^N. We also define the associated space of piecewise constant functions by

$$U_N = \left\{ u \in U |\ u = \sum_{n=1}^{N} \alpha_n \mathcal{X}_{I_n^N} \text{ a.e., } \alpha_n \in \mathbb{R} \right\}.$$

Let $K_N = K \cap U_N$; then it is clear that K_N is also convex and closed. Now, we define the approximated problem of (3) by

$$J(u^{*,N}) = \min_{u \in K_N} J(u).$$

Using similar arguments, one can show that

$$u^{*,N} = P_{K_N}\left(u^{*,N} - \rho J'(u^{*,N})\right). \tag{11}$$

Based on the above optimality condition, we propose the following fixed-point iteration scheme to get the approximated optimal control:

$$u^{i+1,N} = P_{K_N}\big(u^{i,N} - \rho_i J_N'(u^{i,N})\big), \qquad i = 1, 2, \ldots, \tag{12}$$

where ρ_i is a positive constant. Notice that in the above equation we have changed $J'(\cdot)$ in (11) to $J_N'(\cdot)$, as one cannot compute $J'(\cdot)$ exactly in general, and thus it is obtained by numerical approaches. The iteration procedure in (11) is a projected gradient iteration method for solving the variation inequality (4).

It is clear that $J_N'(\cdot)$ depends on particular numerical schemes, and we shall discuss the numerical approximation of $J_N'(\cdot)$ in later sections. We will denote the error between $J'(\cdot)$ and $J_N'(\cdot)$ by

$$\epsilon_N = \sup_i \left\| J'(u^{i,N}) - J_N'(u^{i,N}) \right\|. \tag{13}$$

Next, we show in Theorem 3.1 the convergence property of the iteration scheme (12).

Theorem 3.1. *Assume that $J'(\cdot)$ is Lipschitz and uniformly monotone around u^* and $u^{*,N}$ in the sense that there exist positive constants c and C such that*

$$\begin{aligned}
&\|J'(u^*) - J'(v)\| \leqslant C\|u^* - v\| \quad \forall v \in K,\\
&\big(J'(u^*) - J'(v), u^* - v\big) \geqslant c\|u^* - v\|^2 \quad \forall v \in K,\\
&\|J'(u^{*,N}) - J'(v)\| \leqslant C\|u^{*,N} - v\| \quad \forall v \in K_N,\\
&\big(J'(u^{*,N}) - J'(v), u^{*,N} - v\big) \geqslant c\|u^{*,N} - v\|^2 \quad \forall v \in K_N.
\end{aligned}$$

Moreover, we assume that

$$\epsilon_N = \sup_i \left\| J'(u^{i,N}) - J_N'(u^{i,N}) \right\| \to 0, \quad N \to \infty.$$

Suppose that ρ_i *is chosen such that* $0 < 1 - 2c\rho_i + (1+2C)\rho_i^2 \leqslant \delta^2$ *for some constant* $0 < \delta < 1$. *Then, the iteration scheme* (12) *is convergent; more precisely, we have*

$$\|u^* - u^{i,N}\| \to 0, \quad i, N \to \infty.$$

Proof. By (11) and (12), we have

$$\begin{aligned}\left\|u^{*,N} - u^{i+1,N}\right\|^2 &\leqslant \left\|u^{*,N} - u^{i,N} - \rho_i\big(J'(u^{*,N}) - J_N'(u^{i,N})\big)\right\|^2 \\ &= \left\|u^{*,N} - u^{i,N}\right\|^2 - 2\rho_i\big(u^{*,N} - u^{i,N}, J'(u^{*,N}) - J_N'(u^{i,N})\big) \\ &\quad + \rho_i^2 \left\|J'(u^{*,N}) - J_N'(u^{i,N})\right\|^2.\end{aligned}$$

By the Lipschitz condition and the monotonicity property of $J'(\cdot)$, we have

$$\begin{aligned}&- 2\rho_i\big(u^{*,N} - u^{i,N}, J'(u^{*,N}) - J_N'(u^{i,N})\big) \\ =&- 2\rho_i\big(u^{*,N} - u^{i,N}, J'(u^{*,N}) - J'(u^{i,N}) + J'(u^{i,N}) - J_N'(u^{i,N})\big) \\ \leqslant&- 2c\rho_i \left\|u^{*,N} - u^{i,N}\right\|^2 + \rho_i^2 \left\|u^{*,N} - u^{i,N}\right\|^2 + \epsilon_N^2.\end{aligned}$$

Moreover, we have

$$\begin{aligned}\rho_i^2 \left\|J'(u^{*,N}) - J_N'(u^{i,N})\right\|^2 &= \rho_i^2 \left\|J'(u^{*,N}) - J'(u^{i,N}) + J'(u^{i,N}) - J_N'(u^{i,N})\right\|^2 \\ &\leqslant 2C\rho_i^2 \left\|u^{*,N} - u^{i,N}\right\|^2 + 2\rho_i^2\epsilon_N^2.\end{aligned}$$

It is easy to show that for sufficiently small ρ_i, there is a constant $0 < \delta < 1$ such that $0 < 1 - 2c\rho_i + (1+2C)\rho_i^2 \leqslant \delta^2$; then we get

$$\left\|u^{*,N} - u^{i+1,N}\right\|^2 \leqslant \delta^2 \left\|u^{*,N} - u^{i,N}\right\|^2 + (1 + 2\rho_i^2)\epsilon_N^2. \tag{14}$$

Then, there exists a constant C_1 that is independent of N and i such that

$$\left\|u^{*,N} - u^{i,N}\right\| \leqslant \delta^i \left\|u^{*,N} - u^{0,N}\right\| + C_1\epsilon_N.$$

Under the assumption $\epsilon_N \to 0$, we get

$$\|u^{*,N} - u^{i,N}\| \to 0 \quad (N, i \to \infty). \tag{15}$$

On the other hand, using similar arguments as for deriving (14), we obtain

$$\begin{aligned}\left\|u^* - u^{*,N}\right\| &= \left\|u^* - P_{K_N}(u^* - \rho J'(u^*)) + P_{K_N}\big(u^* - \rho J'(u^*)\big) - u^{*,N}\right\| \\ &\leqslant \|u^* - P_{K_N}(u^* - \rho J'(u^*))\| + \sqrt{1 - 2c\rho + C\rho^2}\left\|u^* - u^{*,N}\right\|.\end{aligned}$$

Let $\rho = c/C$, $C_2 = \big(1 - \sqrt{1 - 2c\rho + C\rho^2}\big)^{-1}$; we have

$$\left\|u^* - u^{*,N}\right\| \leqslant C_2 \left\|u^* - P_{K_N}(u^* - \rho J'(u^*))\right\|.$$

Since $\mathcal{C}$ is invariant in time, for $v \in U_N$, it holds that $P_K v \in U_N$. Thus we have $P_K v \in K_N$, and then we have $P_K v = P_{K_N} v$. Now, denoting $\omega := u^* - \rho J'(u^*)$, we have

$$
\begin{aligned}
\left\|u^* - u^{*,N}\right\| &\leqslant C_2 \left\|u^* - P_{K_N}(u^* - \rho J'(u^*))\right\| = C_2 \left\|P_K \omega - P_{K_N}\omega\right\| \\
&\leqslant C_2 \left(\|P_K\omega - P_K P_{U_N}\omega\| + \|P_K P_{U_N}\omega - P_{K_N}\omega\|\right) \\
&= C_2 \left(\|P_K\omega - P_K P_{U_N}\omega\| + \|P_{K_N} P_{U_N}\omega - P_{K_N}\omega\|\right) \\
&\leqslant 2C_2 \left\|\omega - P_{U_N}\omega\right\|.
\end{aligned}
$$

As U_N is dense in U, we have $\|\omega - P_{U_N}\omega\| \to 0$, and thus $\|u^* - u^{*,N}\| \to 0$. Then, the conclusion follows from this argument and (15). ∎

In Theorem 3.1 we have shown the convergence of $\|u^{*,N} - u^{i,N}\|$ under the assumption $\epsilon_N \to 0$. Note that this is a reasonable assumption. In fact, under certain regularity requirements, and by designing suitable numerical approaches for $J'_N(\cdot)$, one could further expect that $\epsilon_N \sim \mathcal{O}(\Delta t)$. In such a case, we could expect a first order rate of convergence of our iteration scheme (12), as illustrated in the following corollary.

Corollary 3.2. *We suppose that the conditions in Theorem* 3.1 *hold, and furthermore, we assume that u^* and $J'(u^*)$ are both Lipschitz continuous functions in U. Then under the condition $\epsilon_N \sim \mathcal{O}(\Delta t)$ we have*

$$
\left\|u^* - u^{i,N}\right\| \sim \mathcal{O}(\Delta t), \quad i \to \infty.
$$

The iteration scheme (12) is the starting point of our numerical approach for SOCPs. In the following sections, we shall show how to get the numerical approximation $J'_N(u)$ of $J'(u)$ in each iteration.

3.1. The representation of $J'(u)$. Note that the iteration scheme (12) involves the computation of $J'(u)$. In this section, we will derive a new formula of $J'(u)$ for fixed $u \in K$ by introducing a pair of adjoint processes. Again, in all of our arguments, $J'(u)$ is referred to as its representation in U.

By the definition (5), we have

$$
\begin{aligned}
(J'(u), v) &= \lim_{\rho \downarrow 0} \frac{J(u+\rho v) - J(u)}{\rho} \\
&= \mathbb{E}\left[\int_0^T h'(x_t^u) Dx_t^u(v)\,dt + \int_0^T j'\big(u(t)\big)v(t)\,dt + k'(x_T^u) Dx_T^u(v)\right],
\end{aligned} \tag{16}
$$

where x_t^u is the solution of the SDE (1), and

$$
Dx_t^u(v) := \lim_{\rho \downarrow 0} \frac{x_t^{u+\rho v} - x_t^u}{\rho}.
$$

Under Assumption 2.1, the process $Dx_t^u(v)$ satisfies the SDE

$$
\begin{aligned}
dDx_t^u(v) &= \Big(b'_x\big(x_t^u, u(t)\big)Dx_t^u(v) + b'_u\big(x_t^u, u(t)\big)v(t)\Big)dt \\
&\quad + \Big(\sigma'_x\big(x_t^u, u(t)\big)Dx_t^u(v) + \sigma'_u\big(x_t^u, u(t)\big)v(t)\Big)dW_t, \quad Dx_0^u(v) = 0.
\end{aligned} \tag{17}
$$

Notice that one can resort to the above equation to get $J'(u)$; however, this would involve very complicated numerical schemes for solving (17) (see, e.g., [7]). To overcome this, we shall introduce a pair of adjoint processes (p^u, q^u) that solves the following BSDE:

$$
-\,dp_t^u = f\big(x_t^u, p_t^u, q_t^u, u(t)\big)dt - q_t^u\, dW_t, \quad p_T^u = g(x_T^u) = k'(x_T^u), \tag{18}
$$

where f is defined as

$$
f(x,p,q,u) = h'(x) + p\, b_x'(x,u) + q\,\sigma_x'(x,u).
$$

Notice that by the standard BSDE theory, under Assumption 2.1, the BSDE (18) admits an unique solution (p_t^u, q_t^u) for $u \in K$. We remark that theoretical study of BSDEs has been a hot topic recently. In particular, the well-posedness of our adjoint equation, i.e., the BSDE (18), has been well discussed under mild assumptions. One can refer to [34, 37] and [30] for more details on BSDE theory.

We shall show in the following that by introducing the solution pair (p_t^u, q_t^u) of the BSDE (18), the involvement of terms $Dx_t^u(v)$ in (16) will be canceled. By (17), (18), and Itô's formula, we deduce

$$
\begin{aligned}
d\big(p_t^u\, Dx_t^u(v)\big) =& dp_t^u\, Dx_t^u(v) + p_t^u\, dDx_t^u(v)\\
&+ q_t^u \Big(\sigma_x'\big(x_t^u, u(t)\big)Dx_t^u(v) + \sigma_u'\big(x_t^u, u(t)\big)v(t)\Big)dt\\
=& (-(h'(x_t^u) + p_t^u\, b_x'(x_t^u, u(t)) + q_t^u\,\sigma_x'(x_t^u, u(t)))dt + q_t^u dW_t)\, Dx_t^u(v)\\
&+ p_t^u \Big(\Big(b_x'\big(x_t^u, u(t)\big)Dx_t^u(v) + b_u'\big(x_t^u, u(t)\big)v(t)\Big)dt\\
&\quad + \Big(\sigma_x'\big(x_t^u, u(t)\big)Dx_t^u(v) + \sigma_u'\big(x_t^u, u(t)\big)v(t)\Big)dW_t\Big)\\
&+ q_t^u \Big(\sigma_x'\big(x_t^u, u(t)\big)Dx_t^u(v) + \sigma_u'\big(x_t^u, u(t)\big)v(t)\Big)dt\\
=& -h'(x_t^u)Dx_t^u(v)dt\\
&+ \big(p_t^u b_u'\big(x_t^u, u(t)\big)v(t) + q_t^u\sigma_u'\big(x_t^u, u(t)\big)v(t)\big)\, dt\\
&\quad + \Big(p_t^u\sigma_x'\big(x_t^u, u(t)\big)Dx_t^u(v) + p_t^u\sigma_u'\big(x_t^u, u(t)\big)v(t) + q_t^u Dx_t^u(v)\Big)dW_t.
\end{aligned}
$$

Then, we have

$$
\begin{aligned}
\int_0^T h'(x_t^u)Dx_t^u(v)dt =& -p_T^u\, Dx_T^u(v) + p_0^u\, Dx_0^u(v)\\
&+ \int_0^T \Big(p_t^u\, b_u'\big(x_t^u, u(t)\big) + q_t^u\,\sigma_u'\big(x_t^u, u(t)\big)\Big)v(t)dt\\
&+ \int_0^T \Big(p_t^u\,\sigma_x'\big(x_t^u, u(t)\big)Dx_t^u(v) + p_t^u\sigma_u'\big(x_t^u, u(t)\big)v(t)\\
&\qquad + q_t^u\, Dx_t^u(v)\Big)dW_t.
\end{aligned} \tag{19}
$$

Then, by inserting (19) into (16) and using the initial condition $Dx_t^u(v) = 0$ and the terminal condition $p_T^u = k'(x_T^u)$, we obtain

$$
(J'(u), v) = \int_0^T \Big(\mathbb{E}\big[p_t^u\, b_u'\big(x_t^u, u(t)\big) + q_t^u\,\sigma_u'\big(x_t^u, u(t)\big)\big] + j'\big(u(t)\big)\Big)v(t)dt.
$$

Then, we can redefine $J'(u)$ by

$$
J'(u)|_t = \mathbb{E}\big[p_t^u\, b_u'\big(x_t^u, u(t)\big) + q_t^u\,\sigma_u'\big(x_t^u, u(t)\big)\big] + j'\big(u(t)\big). \tag{20}
$$

Here $J'(u)|_t$ represents $J'(u)$, as an element of U, valued at t.

To simplify the expression of $J'(u)$, we have introduced a pair of adjoint processes (p^u, q^u) that satisfies the BSDE (18), to get rid of the term $Dx_t^u(\cdot)$. Then, by solving the BSDE (18), we can get the solution pair (p^u, q^u) numerically, and then further get an approximated $J'_N(u)$ of $J'(u)$ by using (20). In the next section, we shall propose an Euler-type method for solving the adjoint BSDE (18).

Remark 3.3. We remark that the authors in [12] also introduced an adjoint equation to cancel the term $Dx_t^u(\cdot)$. The adjoint equation therein is an anticipating integrand SDE, where the solution is required to be *backward-adapted* instead of the classic forward-adapted. However, such a requirement is not true in general. In other words, the well-posedness of the adjoint equation in [12] is unclear for general situations.

3.2. Numerical approximations for adjoint equations. By (18), we notice that the solution pair (p_t^u, q_t^u) depends on the forward process x_t^u. Hence, we need to solve (for $t \in [0,T]$) the following forward-backward stochastic differential equations (FBSDEs):

$$\begin{cases} dx_t^u = b\big(x_t^u, u(t)\big)dt + \sigma\big(x_t^u, u(t)\big)dW_t, & x_{t=0} = x_0, \\ -dp_t^u = f\big(x_t^u, p_t^u, q_t^u, u(t)\big)dt - q_t^u dW_t, & p_T^u = g(x_T^u). \end{cases} \tag{21}$$

Next, we shall discuss how to solve the above FBSDEs numerically with a given $u \in K$. For notational simplicity, we shall omit the superscript u in this section, such as $x_t = x_t^u$, $p_t = p_t^u$, $q_t = q_t^u$.

Under mild assumptions, it is well known that the above backward equation is well-posed [36]. Moreover, the solutions p_t and q_t have the representations

$$p_t = \eta(t, x_t), \qquad q_t = \sigma\big(x_t, u(t)\big)\partial_x \eta(t, x_t), \tag{22}$$

where $\eta(t,x) : [0,T] \times \mathbb{R} \to \mathbb{R}$ is the solution of the following parabolic PDE:

$$\mathcal{L}^0 \eta(t,x) = -f\Big(x, \eta(t,x), \sigma\big(x, u(t)\big)\partial_x \eta(t,x), u(t)\Big), \qquad \eta(T,x) = g(x), \tag{23}$$

with

$$\mathcal{L}^0 \eta(t,x) = \partial_t \eta(t,x) + b\big(x, u(t)\big)\partial_x \eta(t,x) + \frac{1}{2}\sigma\big(x, u(t)\big)^2 \partial_{xx}\eta(t,x).$$

The representation in (22) is the so-called nonlinear Feynman–Kac formula [36].

We remark that numerical methods for FBSDEs have been a hot topic recently, and one can refer to [16, 33, 44, 47, 48, 49] and references therein for variable numerical approaches. In the present paper, we shall introduce a simple scheme, namely the Euler scheme, for solving the FBSDEs (21).

3.3. The Euler scheme for FBSDEs. We now closely follow the works [47] and [48] to introduce the Euler method for solving the FBSDEs (21). The time partition was defined in (10). By integrating both sides of the backward equation on $[t_n, t_{n+1}]$ we obtain

$$p_{t_n} = p_{t_{n+1}} + \int_{t_n}^{t_{n+1}} f(x_t, p_t, q_t, u(t))\, dt - \int_{t_n}^{t_{n+1}} q_t\, dW_t. \tag{24}$$

Then, by taking conditional expectation $\mathbb{E}^x_{t_n}[\cdot] = \mathbb{E}[\cdot \,|\, \mathcal{F}_{t_n}, x_{t_n} = x]$ on both sides of (24) and applying the left-point rectangular rule, we have

$$p^x_{t_n} = \mathbb{E}^x_{t_n}\big[p_{t_{n+1}}\big] + \Delta t\, f\big(x, p^x_{t_n}, q^x_{t_n}, u(t_n)\big) + \bar{R}^x_{p,n}, \tag{25}$$

where

$$\bar{R}^x_{p,n} = \int_{t_n}^{t_{n+1}} \mathbb{E}^x_{t_n}\big[f\big(x_t, p_t, q_t, u(t)\big)\big]\, dt - \Delta t\, f\big(x, p^x_{t_n}, q^x_{t_n}, u(t_n)\big)$$

is the truncation error due to the left-point rectangular rule. Equation (25) is our semidiscrete equation for solving p.

Next, we aim to derive another discrete equation for solving q. To this end, by multiplying (24) by $\Delta W_{n+1} := W_{t_{n+1}} - W_{t_n}$ and taking conditional expectation $\mathbb{E}^x_{t_n}[\cdot]$ on both sides of the derived equation and then again applying the left-point rectangular rule, we obtain

$$q^x_{t_n} = \frac{1}{\Delta t}\Big(\mathbb{E}^x_{t_n}\big[p_{t_{n+1}}\Delta W_{n+1}\big] + \bar{R}^x_{q,n}\Big), \tag{26}$$

where

$$\bar{R}^x_{q,n} = \int_{t_n}^{t_{n+1}} \mathbb{E}^x_{t_n}\big[f\big(x_t, p_t, q_t, u(t)\big)\Delta W_{n+1}\big]\, dt - \int_{t_n}^{t_{n+1}} \mathbb{E}^x_{t_n}[q_t]\, dt + \Delta t\, q^x_{t_n}$$

is again the corresponding truncation error.

By removing the error terms $\bar{R}^x_{p,n}$ and $\bar{R}^x_{q,n}$ in (25) and (26), we get the following semidiscretization scheme for the BSDE in (21): impose the initial value of $p^x_N = g(x)$ on $x \in \mathbb{R}$, and then for $n = N-1, \dots, 1, 0$, compute $p^x_n = p_n(x)$ and $q^x_n = q_n(x)$ with $x \in \mathbb{R}$ in a backward way by

$$p^x_n = \mathbb{E}^x_{t_n}\big[p_{n+1}\big] + \Delta t\, f\big(x, p^x_n, q^x_n, u(t_n)\big), \tag{27}$$

$$q^x_n = \frac{1}{\Delta t}\mathbb{E}^x_{t_n}\big[p_{n+1}\Delta W_{n+1}\big]. \tag{28}$$

Notice that in the above semidiscretization schemes (27) and (28), solving p^x_n and q^x_n for each $x \in \mathbb{R}$ may involve the knowledge of p_{n+1} on the whole space region $\mathbb{R}$. To apply this scheme in practice, the spacial discretization of $\mathbb{R}$ and the approximations of the conditional expectation $\mathbb{E}^x_{t_n}[\cdot]$ are required.

To do this, we introduce a uniform partition $\mathbb{R}_h$ of the $\mathbb{R}$ as

$$\mathbb{R}_h = \big\{x_k \,\big|\, k = 0, \pm 1, \pm 2, \dots\big\}, \quad \text{with} \quad \Delta x = x_{k+1} - x_k.$$

We shall denote $I_k =: [x_k, x_{k+1}]$. Notice that the above partition involves infinite grid points; however, this is unnecessary in practical applications, as we are always interested in the final information ($t = 0$) in a finite interval. This means that we can consider a finite partition with $|k| \le P$, with P being a positive integer (which can be very large and problem dependent). In what follows, we shall consider a finite partition with the parameter P. We remark that choosing a reasonable P is not a trivial task, and we refer the reader to [48] for further discussion.

On the partition $\mathbb{R}_h$, we introduce a continuous piecewise linear function space V_h, the element of which $v \in V_h$ can be represented as follows:

$$v(x) = \sum_{|k| \le P} v_k(x_k)\phi_k(x), \quad \text{with} \quad \phi_k(x) = \begin{cases} \frac{x - x_{k-1}}{x_k - x_{k-1}}, & x \in I_{k-1}, \\ \frac{x_{k+1} - x}{x_{k+1} - x_k}, & x \in I_k, \\ 0 & \text{otherwise.} \end{cases}$$

For a continuous function $f(x)$, we now introduce the associated interpolation operator $\mathcal{I}_h$ by

$$\mathcal{I}_h f(x) = \sum_{|k|\leq P} f(x_k)\phi_k(x);$$

i.e., a function in V_h is determined by its values at the grid points in $\mathbb{R}_h$.

3.3.1. The approximation of conditional expectations. We now discuss the approximation of conditional expectations. Let $\tilde{x}_{t_{n+1}}^{t_n,x}$ be the Euler approximation of the state $x_{t_{n+1}}^{t_n,x}$, namely,

$$\begin{aligned}\tilde{x}_{t_{n+1}}^{t_n,x} &= x + b\big(x,u(t_n)\big)\Delta t + \sigma\big(x,u(t_n)\big)\Delta W_{n+1} \\ &= x + b\big(x,u(t_n)\big)\Delta t + \sigma\big(x,u(t_n)\big)\sqrt{\Delta t}\,\zeta,\end{aligned} \tag{29}$$

where $\zeta \sim \mathcal{N}(0,1)$ is a normal random variable.

We choose $\tilde{p}_{t_{n+1}}^{t_n,x} = p_{t_{n+1}}(\tilde{x}_{t_{n+1}}^{t_n,x}) = \eta(t_{n+1},\tilde{x}_{t_{n+1}}^{t_n,x})$ to approximate $p_{t_{n+1}}^{t_n,x}$, where $\eta(t,x)$ is the solution of problem (22). As a result, $\tilde{p}_{t_{n+1}}^{t_n,x_k}$ is a function of $\tilde{x}_{t_{n+1}}^{t_n,x_k}$ and thus a function of the increment ΔW_{n+1}. Therefore, the conditional expectation $\mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}]$ (as well as $\mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}\Delta W_{n+1}]$) can be written into an integral on $\mathbb{R}$ with the Gaussian probability density function $\rho(\xi) = \frac{1}{\sqrt{2\pi}}e^{-\xi^2/2}$. Hence we propose the Gauss–Hermite quadrature rule to approximate these conditional expectations. The L-point Gauss–Hermite quadrature rule for a function f writes as

$$\mathbb{E}[f(\zeta)] = \int_{\mathbb{R}} f(\xi)\rho(\xi)d\xi \approx \sum_{\ell=1}^{L} f(\xi_\ell)\omega_\ell, \tag{30}$$

where $\{\xi_\ell\}$ and $\{\omega_\ell\}$ are the Gaussian–Hermite quadrature points and the associated weights, respectively.

Consider, for example, the approximation of the conditional expectation $\mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}]$; we have

$$\begin{aligned}\mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}] &= \mathbb{E}[p_{t_{n+1}}(\tilde{x}_{t_{n+1}}^{t_n,x})] \\ &= \mathbb{E}\big[p_{t_{n+1}}\big(x + b\big(x,u(t_n)\big)\Delta t + \sigma\big(x,u(t_n)\big)\sqrt{\Delta t}\,\zeta\big)\big] \\ &\approx \sum_{\ell=1}^{L} p_{t_{n+1}}\big(x + b\big(x,u(t_n)\big)\Delta t + \sigma\big(x,u(t_n)\big)\sqrt{\Delta t}\,\xi_\ell\big)\omega_\ell.\end{aligned} \tag{31}$$

We shall denote by $\tilde{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}]$ the approximation of $\mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}]$; more precisely,

$$\tilde{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}] = \sum_{\ell=1}^{L} p_{t_{n+1}}\big(x + b\big(x,u(t_n)\big)\Delta t + \sigma\big(x,u(t_n)\big)\sqrt{\Delta t}\,\xi_\ell\big)\omega_\ell. \tag{32}$$

Similarly, we denote by $\tilde{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}]$ the approximation of $\mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}\Delta W_{n+1}]$:

$$\tilde{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}] = \sum_{\ell=1}^{L} p_{t_{n+1}}\big(x + b\big(x,u(t_n)\big)\Delta t + \sigma\big(x,u(t_n)\big)\sqrt{\Delta t}\,\xi_\ell\big)\sqrt{\Delta t}\,\xi_\ell\,\omega_\ell. \tag{33}$$

In the quadrature rule (32), we notice that $\hat{x} = x + b(x,u(t_n))\Delta t + \sigma(x,u(t_n))\sqrt{\Delta t}\,\xi_\ell$ may not be on the partition $\mathbb{R}_h$. Therefore, we shall resort to the linear interpolation

$\mathcal{I}_h$ to get the desired information. To this end, we define

$$\hat{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}] = \sum_{\ell=1}^{L} \mathcal{I}_h p_{t_{n+1}}\big(x + b\big(x, u(t_n)\big)\Delta t + \sigma\big(x, u(t_n)\big)\sqrt{\Delta t}\,\xi_\ell\big)\omega_\ell. \tag{34}$$

Similarly, we define $\hat{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}]$ as

$$\begin{aligned}\hat{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}] = \sum_{\ell=1}^{L} \mathcal{I}_h p_{t_{n+1}}\big(x &+ b\big(x, u(t_n)\big)\Delta t \\ &+ \sigma\big(x, u(t_n)\big)\sqrt{\Delta t}\,\xi_\ell\big)\sqrt{\Delta t}\,\xi_\ell\,\omega_\ell.\end{aligned} \tag{35}$$

Notice that the approximated expectation $\hat{\mathbb{E}}_{t_n}^x[\cdot]$ is a function of x. In the partition space $\mathbb{R}_h$, we denote $\hat{\mathbb{E}}_{t_n}[\cdot] := \hat{\mathbb{E}}_{t_n}^{x_{t_n}}[\cdot]$. In addition, for functions $f \in V_h$, we have

$$\begin{aligned}\hat{\mathbb{E}}_{t_n}^x[f(\tilde{x}_{t_{n+1}}^{t_n,x})] &= \tilde{\mathbb{E}}_{t_n}^x[f(\tilde{x}_{t_{n+1}}^{t_n,x})],\\ \hat{\mathbb{E}}_{t_n}^x[f(\tilde{x}_{t_{n+1}}^{t_n,x})\Delta W_{n+1}] &= \tilde{\mathbb{E}}_{t_n}^x[f(\tilde{x}_{t_{n+1}}^{t_n,x})\Delta W_{n+1}].\end{aligned}$$

Based on the above observations, we finally get the following approximations $\hat{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}]$ and $\hat{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}]$ of $\mathbb{E}_{t_n}^x[p_{t_{n+1}}]$ and $\mathbb{E}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}]$:

$$\begin{aligned}\mathbb{E}_{t_n}^x[p_{t_{n+1}}] &= \hat{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}] + \hat{R}_{p,n}^x,\\ \mathbb{E}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}] &= \hat{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}] + \hat{R}_{q,n}^x,\end{aligned} \tag{36}$$

where $\hat{R}_{p,n}^k$ and $\hat{R}_{q,n}^k$ are the truncation errors,

$$\begin{aligned}\hat{R}_{p,n}^x &= \tilde{R}_{p,n}^x + R_{E,p,n}^x + R_{I_h,p,n}^x,\\ \hat{R}_{q,n}^x &= \tilde{R}_{q,n}^x + R_{E,q,n}^x + R_{I_h,q,n}^x,\end{aligned} \tag{37}$$

with

$$\begin{aligned}\tilde{R}_{p,n}^x &= \mathbb{E}_{t_n}^x[p_{t_{n+1}}] - \mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}], &\quad \tilde{R}_{q,n}^x &= \mathbb{E}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}] - \mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}\Delta W_{n+1}],\\ R_{E,p,n}^x &= \mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}] - \tilde{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}], &\quad R_{E,q,n}^x &= \mathbb{E}_{t_n}^x[\tilde{p}_{t_{n+1}}\Delta W_{n+1}] - \tilde{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}],\\ R_{I_h,p,n}^x &= \tilde{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}] - \hat{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}], &\quad R_{I_h,q,n}^x &= \tilde{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}] - \hat{\mathbb{E}}_{t_n}^x[p_{t_{n+1}}\Delta W_{n+1}].\end{aligned}$$

3.3.2. The fully discrete scheme. By the semidiscrete equations (25) and (26) and the approximations of the conditional expectations in (34), we get the following two equations:

$$p_{t_n}^x = \hat{\mathbb{E}}_{t_n}^x\left[p_{t_{n+1}}\right] + \Delta t\, f\big(x, p_{t_n}^x, q_{t_n}^x, u(t_n)\big) + R_{p,n}^x, \quad p_{t_N}^x = g(x), \tag{38}$$

$$q_{t_n}^x = \frac{1}{\Delta t}\Big(\hat{\mathbb{E}}_{t_n}^x\left[p_{t_{n+1}}\Delta W_{n+1}\right] + R_{q,n}^x\Big), \tag{39}$$

where $R_{p,n}^x$ and $R_{q,n}^x$ are the total truncation errors defined by

$$R_{p,n}^x = \bar{R}_{p,n}^x + \hat{R}_{p,n}^x, \quad R_{q,n}^x = \bar{R}_{q,n}^x + \hat{R}_{q,n}^x, \tag{40}$$

with $\bar{R}_{p,n}^x$ and $\bar{R}_{q,n}^x$ defined as in (25) and (26), and $\hat{R}_{p,n}^x$ and $\hat{R}_{q,n}^x$ defined as in (37).

Based on the two fully discrete equations (38) and (39), we propose a fully discrete numerical scheme for solving the FBSDEs (21) as follows: Given the terminal condition $p_N = \sum_{|k|\le P} g(x_k)\phi_k(\cdot) \in V_h$, for $n = N-1,\dots,1,0$, and each $x_k \in \mathbb{R}_h$, solve $p_n = \sum_k p_n^k \phi_k \in V_h$ and $q_n = \sum_k q_n^k \phi_k \in V_h$ by

$$p_n^k = \hat{\mathbb{E}}_{t_n}^{x_k}\big[p_{n+1}\big] + \Delta t\, f\big(x_k, p_n^k, q_n^k, u(t_n)\big), \tag{41}$$

$$q_n^k = \frac{1}{\Delta t}\hat{\mathbb{E}}_{t_n}^{x_k}\big[p_{n+1}\Delta W_{n+1}\big]. \tag{42}$$

3.4. Summary of the numerical approach. We now summarize the entire algorithm of our gradient projection method. In the fixed-point iteration (12), we have introduced $J_N'(\cdot)$ as the approximation of $J'(\cdot)$. As the relation between $J'(\cdot)$ and the adjoint processes (p,q) has been revealed in (20), it is natural to define the approximation $J_N'(\cdot)$ by replacing p, q, and $\mathbb{E}[\cdot]$ in (20) with the associated numerical approximations. More precisely, we define

$$J_N'(u)|_{t_n} = \hat{\mathbb{E}}\big[p_n b_u'\big(\cdot, u(t_n)\big) + q_n \sigma_u'\big(\cdot, u(t_n)\big)\big] + j'\big(u(t_n)\big), \tag{43}$$

where $\hat{\mathbb{E}}[\cdot]$ is defined by

$$\hat{\mathbb{E}}[\phi_{t_0}] = \phi_{t_0}, \quad \hat{\mathbb{E}}[\phi_{t_n}] = \hat{\mathbb{E}}_{t_0}^{x_0}[\hat{\mathbb{E}}_{t_1}[\cdots\hat{\mathbb{E}}_{t_{n-1}}[\phi_{t_n}]]], \quad n \geqslant 1. \tag{44}$$

To make sure that $J_N'(\cdot) \in U_N$, we define

$$J_N'(u)|_t = \sum_{n=0}^{N-1} J_N'(u)|_{t_n}\, \mathcal{X}_{I_n^N}(t). \tag{45}$$

Then, the gradient projection method is summarized in Algorithm 1.

Algorithm 1 Gradient projection method

Set the initial guess of the control $u_0 \in U_N$ and the error tolerance ϵ_0;
1. Set the terminal condition: $p_N^k = g(x_k)$, $x_k \in \mathbb{R}_h$;
2. For $n = N-1,\dots,1,0$, solve (p_n, q_n) by (41)–(42);
3. Compute $J_N'(u)|_{t_n}$ by (43);
4. Update u by (12);

Repeat the above steps until the error $\|u^{i+1,N} - u^{i,N}\|$ reaches the tolerance ϵ_0.

To be more clear, we state the procedure for updating u by (12) as follows: Given $\{u^{0,N}(t_n)\}_{n=1}^N$, for $i = 0,1,\dots,$

1. solve the forward SDE (1) with $u(t) = \sum_{n=1}^N u^{i+1,N}(t_n)\mathcal{X}_{I_n^N}(t)$,
2. solve p_n and q_n by (41) and (42),
3. calculate $J_u'|_{t_n}$ by (43),
4. update u by (12). To be specific, given $u^{i,N}$ as the control in the ith iteration, we update the control $u = u^{i+1,N}$ by

$$u^{i+1,N} = \sum_{n=1}^N u^{i+1,N}(t_n)\mathcal{X}_{I_n^N},$$

where the coefficients $u^{i+1,N}(t_n)$ are computed by the projection

$$u^{i+1,N}(t_n) = P_{\mathcal{C}}\big(u^{i,N}(t_n) - \rho_i J_N'(u^{i,N})|_{t_n}\big).$$

We remark that the exact projection $P_{\mathcal{C}}$ is used in this paper (as for problems with regular domains, this projection can be obtained exactly). The effect of the approximation error of $P_{\mathcal{C}}$ (when the projection cannot be computed exactly) will be investigated in our future work.

4. Error estimates.

In this section, we shall perform a rigorous error analysis for our gradient projection method. As concluded in Corollary 3.2, the first order rate of convergence relies on the estimate $\epsilon_N = \mathcal{O}(\Delta t)$. By observing the definition of (20) and (43), we see that the error ϵ_N contains two parts: the numerical error of (p_n^k, q_n^k) and the approximation error of $\hat{\mathbb{E}}[\cdot]$. In the following sections, we shall estimate the two parts one by one.

4.1. Preliminary results of the discrete operator $\hat{\mathbb{E}}[\cdot]$.

In this subsection, we first show some basic properties of the approximated conditional expectations $\hat{\mathbb{E}}_{t_n}^{x_k}[\cdot]$ and $\hat{\mathbb{E}}[\cdot]$ which are defined in (34) and (44), respectively.

Notice that the weights of the quadrature rule $\{\omega_\ell\}$ are all positive, and it holds that

$$\sum_\ell \omega_\ell = 1.$$

Moreover, the L-point Gauss–Hermite quadrature rule is exact for polynomials with degree less than or equal to $2L-1$. We now state some basic properties of $\hat{\mathbb{E}}_{t_n}^{x_k}[\cdot]$ in the following proposition.

PROPOSITION 4.1. *Given variables $\phi_{t_{n+1}} = \bar{\phi}(t_{n+1}, x_{t_{n+1}})$, for $L \geqslant 2$, we have the following:*

- $\hat{\mathbb{E}}[\hat{\mathbb{E}}_{t_n}[\phi_{t_{n+1}}]] = \hat{\mathbb{E}}[\phi_{t_{n+1}}]$.
- *If, for any x, it holds that $\phi_{t_{n+1}}^x \geqslant 0$, then we have $\hat{\mathbb{E}}_{t_n}^{x_k}[\phi_{t_{n+1}}] \geqslant 0$, $\hat{\mathbb{E}}[\phi_{t_{n+1}}] \geqslant 0$.*
- $(\hat{\mathbb{E}}_{t_n}^{x_k}[\phi_{t_{n+1}}])^2 \leqslant \hat{\mathbb{E}}_{t_n}^{x_k}[(\phi_{t_{n+1}})^2]$, $(\hat{\mathbb{E}}[\phi_{t_n}])^2 \leqslant \hat{\mathbb{E}}[(\phi_{t_n})^2]$, $(\hat{\mathbb{E}}_{t_n}^{x_k}[\phi_{t_{n+1}}\Delta W_{t_{n+1}}])^2 \leqslant (\hat{\mathbb{E}}_{t_n}^{x_k}[(\phi_{t_{n+1}})^2] - (\hat{\mathbb{E}}_{t_n}^{x_k}[\phi_{t_{n+1}}])^2)\Delta t$.

The points of the above proposition are all well known and easy to prove. It is also known that under Assumption 2.1, for $m \geqslant 1$ it holds that

$$\mathbb{E}[|x_t|^m] \leqslant C(|x_0|^m + 1).$$

In the following, we shall provide a similar result for the approximated expectation $\hat{\mathbb{E}}[\cdot]$. Notice that in what follows, C shall stand for a constant that is independent of Δt, Δx, n, and k, while its value may vary from place to place.

PROPOSITION 4.2. *Under Assumption 2.1, for $m \geqslant 2$, $L \geqslant 2$, and $\Delta x = \mathcal{O}(\sqrt{\Delta t})$, it holds that*

$$\hat{\mathbb{E}}[|x_{t_n}|^m] \leqslant C(|x_0|^m + 1).$$

Proof. We denote by $\mathcal{I}_h|x|^m$ the linear interpolation of the function $|\cdot|^m$ at x. By the interpolation theory, there exists $\theta \in [x^-, x^+]$ (where $x^- < x^+$ are two grid points around x) such that for Δx sufficiently small, it holds that

$$
\begin{aligned}
\mathcal{I}_h|x|^m &\leqslant |x|^m + \frac{1}{8}m(m-1)|\theta|^{m-2}(\Delta x)^2 \\
&\leqslant |x|^m + \frac{1}{8}m(m-1)\big((|x|+\Delta x)^m+1\big)(\Delta x)^2 \\
&\leqslant |x|^m + C\big(|x|^m + C\Delta x(|x|^m+1)+1\big)(\Delta x)^2 \\
&\leqslant \big(1+C(\Delta x)^2\big)|x|^m + C(\Delta x)^2. \qquad (46)
\end{aligned}
$$

For fixed k and ℓ, let $a_1 = x_k + b(x_k, u(t_n))\Delta t$ and $a_2 = \sigma(x_k, u(t_n))\sqrt{\Delta t}\,\xi_\ell$; then there exists $\theta \in [a_1, a_1 + a_2]$ such that

$$
\begin{aligned}
|x_{k,\ell}|^m = |a_1+a_2|^m &= |a_1|^m + m|a_1|^{m-1}\mathrm{sgn}(a_1)a_2 + m(m-1)|\theta|^{m-2}(a_2)^2 \\
&\leqslant |a_1|^m + m|a_1|^{m-1}\mathrm{sgn}(a_1)a_2 + m(m-1)(|a_1|+|a_2|)^{m-2}(a_2)^2. \qquad (47)
\end{aligned}
$$

By the assumptions on b and σ, for sufficiently small Δt we have

$$
\begin{aligned}
|a_1|^m &\leqslant \big((1+C\Delta t)|x_k| + C\Delta t\big)^m \\
&\leqslant (1+C\Delta t)^m|x_k|^m + C\Delta t(|x_k|^m+1) \\
&\leqslant (1+C\Delta t)|x_k|^m + C\Delta t, \\
|a_1| + |a_2| &\leqslant C(|x_k|+1).
\end{aligned}
$$

Using (46)–(47) and the definition

$$
x_{k,\ell} = x_k + b(x_k, u(t_n))\Delta t + \sigma(x_k, u(t_n))\sqrt{\Delta t}\,\xi_\ell,
$$

we have

$$
\begin{aligned}
&\hat{\mathbb{E}}_{t_n}^{x_k}[|x_{t_{n+1}}|^m] = \sum_{\ell=1}^{L} \mathcal{I}_h|x_{k,\ell}|^m\omega_\ell \\
&\leqslant \big(1+C(\Delta x)^2\big)\sum_{\ell=1}^{L}|x_{k,\ell}|^m\omega_\ell + C(\Delta x)^2 \\
&\leqslant \big(1+C(\Delta x)^2\big)\Big((1+C\Delta t)|x_k|^m + C\Delta t + C(|x_k|^m+1)\Delta t\Big) + C(\Delta x)^2 \\
&\leqslant \big(1+C(\Delta x)^2\big)\big((1+C\Delta t)|x_k|^m + C\Delta t\big) + C(\Delta x)^2.
\end{aligned}
$$

Consequently, by the definition (44) and the assumption $\Delta x = \mathcal{O}(\sqrt{\Delta t})$, we have

$$
\begin{aligned}
&\hat{\mathbb{E}}\left[|x_{t_{n+1}}|^m\right] = \hat{\mathbb{E}}\big[\hat{\mathbb{E}}_{t_n}[|x_{t_{n+1}}|^m]\big] \\
&\leqslant \big(1+C(\Delta x)^2\big)\Big((1+C\Delta t)\hat{\mathbb{E}}[|x_{t_n}|^m] + C\Delta t\Big) + C(\Delta x)^2 \\
&\leqslant \big(1+C(\Delta x)^2\big)^{n+1}(1+C\Delta t)^{n+1}\Big(|x_0|^m + (n+1)C\Delta t + (n+1)C(\Delta x)^2\Big) \\
&\leqslant C(|x_0|^m+1).
\end{aligned}
$$

This completes the proof. ∎

Next, by the variational arguments, we can easily present an approximation property for the expectation $\hat{\mathbb{E}}[\cdot]$.

LEMMA 4.3. *Assume that* $b, \sigma \in C_b^{0,4}$. *For* $\phi_t = \bar{\phi}(t, x_t)$ *with* $\bar{\phi} \in C_b^{0,4}$, *we define* $\Phi_{t_i}(x) = \mathbb{E}_{t_i}^x[\phi_{t_n}]$; *then it holds that* $\Phi_{t_i} \in C_b^{0,4}$, *and furthermore we have*

$$\mathbb{E}[\phi_{t_n}] = \hat{\mathbb{E}}[\phi_{t_n}] + \sum_{i=0}^{n-1} \hat{\mathbb{E}}[\hat{R}_{\Phi,i}],$$

$$\textit{with} \quad \hat{R}_{\Phi,i} = \mathbb{E}_{t_i}^{x_{t_i}}[\Phi_{t_{i+1}}] - \hat{\mathbb{E}}_{t_i}^{x_{t_i}}[\Phi_{t_{i+1}}], \quad 1 \leqslant i \leqslant n.$$

4.2. The error estimates of (p_n^k, q_n^k). We let

$$\mu_n = p_{t_n} - p_n, \qquad \nu_n = q_{t_n} - q_n,$$

where (p_t, q_t) and (p_n, q_n) are the exact solutions of the FBSDEs (21) and numerical solutions of the scheme (41)–(42), respectively. Notice that

$$p_n(x) = \sum_{k=-\infty}^{\infty} p_n^k \phi_k(x), \quad q_n(x) = \sum_{k=-\infty}^{\infty} q_n^k \phi_k(x),$$

where (p_n^k, q_n^k) are numerical solutions by scheme (41)–(42), and for $x_k \in \mathbb{R}_h$ we have $(p_n^k, q_n^k) = (p_n(x_k), q_n(x_k))$. We now define

$$\mu_n^k = p_{t_n}^{x_k} - p_n^k, \qquad \nu_n^k = q_{t_n}^{x_k} - q_n^k.$$

Then, by subtracting (41) from (38), and (42) from (39), respectively, we deduce that

$$\mu_n^k = \hat{\mathbb{E}}_{t_n}^{x_k}[\mu_{n+1}] + \Delta t\, \delta f_n^k + R_{p,n}^k, \quad \mu_N^k = p_{t_N}^{x_k} - p_N^k, \tag{48}$$

$$\nu_n^k = \frac{1}{\Delta t}\Big(\hat{\mathbb{E}}_{t_n}^{x_k}[\mu_{n+1}\Delta W_{n+1}] + R_{q,n}^k\Big), \tag{49}$$

where

$$\delta f_n^k = f(x_k, p_{t_n}^{x_k}, q_{t_n}^{x_k}, u(t_n)) - f(x_k, p_n^k, q_n^k, u(t_n)), \quad R_{p,n}^k = R_{p,n}^{x_k}, \quad R_{q,n}^k = R_{q,n}^{x_k}.$$

Now, we are ready to give the estimates of μ_n^k and ν_n^k in the following lemma. The estimates also imply the stability of the scheme (41)–(42) and will be used in our final error estimates.

LEMMA 4.4. *Under Proposition* 4.1, *namely, assume that* $f(x, p, q, u)$ *is Lipschitz continuous with respect to* p *and* q, *uniformly in* x *and* u; *then there holds*

$$\hat{\mathbb{E}}[(\mu_n)^2] + \Delta t \sum_{n=0}^{N-1} \hat{\mathbb{E}}[(\nu_n)^2] \leqslant C\, \hat{\mathbb{E}}[(\mu_N)^2] + \frac{C}{\Delta t} \sum_{n=0}^{N-1} \hat{\mathbb{E}}\left[(R_{p,n})^2 + (R_{q,n})^2\right].$$

Proof. By taking square of (48)–(49) and using Proposition 4.1 and the inequality $(a+b)^2 \leqslant (1+\epsilon)a^2 + (1+1/\epsilon)b^2$, we get

$$\begin{aligned}(\mu_n^k)^2 &\leqslant (1+\gamma\Delta t)\big(\hat{\mathbb{E}}_{t_n}^{x_k}[\mu_{n+1}]\big)^2 \\ &\quad + C\Big(1 + \frac{1}{\gamma \Delta t}\Big)\Big(\Delta t^2\big((\mu_n^k)^2 + (\nu_n^k)^2\big) + (R_{p,n}^k)^2\Big), \\ (\nu_n^k)^2 &\leqslant \frac{C}{\Delta t}\Big(\hat{\mathbb{E}}_{t_n}^{x_k}[(\mu_{n+1})^2] - \big(\hat{\mathbb{E}}_{t_n}^{x_k}[\mu_{n+1}]\big)^2\Big) + \frac{C}{(\Delta t)^2}\left(R_{q,n}^k\right)^2.\end{aligned}$$

Let $\Delta t \leqslant 1/\gamma$, $\gamma = 4C^2$, and add up the above inequalities to get

$$(\mu_n^k)^2 + \frac{\Delta t}{2C}(\nu_n^k)^2 \leqslant (1+\gamma\Delta t)\hat{\mathbb{E}}_{t_n}^{x_k}[(\mu_{n+1})^2] + \frac{\Delta t}{2C}(\mu_n^k)^2 + \frac{1}{2\Delta t}\Big((R_{p,n}^k)^2 + (R_{q,n}^k)^2\Big),$$

which yields

$$(\mu_n^k)^2 + C\Delta t(\nu_n^k)^2 \leqslant (1+C\Delta t)\hat{\mathbb{E}}_{t_n}^{x_k}[(\mu_{n+1})^2] + \frac{1}{\Delta t}\Big((R_{p,n}^k)^2 + (R_{q,n}^k)^2\Big).$$

By taking discrete expectation on the above inequality, we have

$$\hat{\mathbb{E}}[(\mu_n)^2] + C\Delta t\,\hat{\mathbb{E}}[(\nu_n)^2] \leqslant (1+C\Delta t)\hat{\mathbb{E}}[(\mu_{n+1})^2] + \frac{1}{\Delta t}\hat{\mathbb{E}}\big[(R_{p,n})^2 + (R_{q,n})^2\big]. \tag{50}$$

Then, we get

$$\hat{\mathbb{E}}[(\mu_n)^2] \leqslant C\,\hat{\mathbb{E}}[(\mu_N)^2] + \frac{C}{\Delta t}\sum_{n=0}^{N-1}\hat{\mathbb{E}}\big[(R_{p,n})^2 + (R_{q,n})^2\big]. \tag{51}$$

Taking the summation of (50) from $n = 0$ to $N-1$, we get

$$\begin{aligned} C\Delta t\sum_{n=0}^{N-1}\hat{\mathbb{E}}[(\nu_n)^2] &\leqslant \sum_{n=0}^{N-1}\Big(C\Delta t\,\hat{\mathbb{E}}[(\mu_{n+1})^2] + \frac{1}{\Delta t}\hat{\mathbb{E}}\big[(R_{p,n})^2 + (R_{q,n})^2\big]\Big) \\ &\leqslant C\,\hat{\mathbb{E}}[(\mu_N)^2] + \frac{C}{\Delta t}\sum_{n=0}^{N-1}\hat{\mathbb{E}}\big[(R_{p,n})^2 + (R_{q,n})^2\big]. \end{aligned} \tag{52}$$

Then, the proof is complete. ∎

We now provide the following lemma for estimating the truncation errors, and the proof is somewhat standard using the arguments of approximation theory.

LEMMA 4.5. *Suppose that Assumption* 2.1 *holds, and, moreover, we assume that* $b(\cdot,w), \sigma(\cdot,w) \in C_b^4$ *and* $f(\cdot,\cdot,\cdot,w) \in C_b^{2,2,2}$ *hold uniformly in* $w \in \mathcal{C}$, $\eta \in C_b^{1,4}$. *Then, we have*

$$\frac{1}{\Delta t}\sum_{n=0}^{N-1}\hat{\mathbb{E}}\big[(R_{p,n})^2 + (R_{q,n})^2\big] = \mathcal{O}\big((\Delta t)^2\big) + \mathcal{O}\big((\Delta x)^4/(\Delta t)^2\big).$$

Proof. As shown in (37), (38), and (39), $R_{p,n}^k$ and $R_{q,n}^k$ consist of the following parts of errors:

$$\begin{aligned} R_{p,n}^k &= \bar{R}_{p,n}^k + \tilde{R}_{p,n}^k + R_{E,p,n}^k + R_{I_h,p,n}^k, \\ R_{q,n}^k &= \bar{R}_{q,n}^k + \tilde{R}_{q,n}^k + R_{E,q,n}^k + R_{I_h,q,n}^k. \end{aligned} \tag{53}$$

By the interpolation theory, we have the following estimate:

$$R_{I_h,p,n}^k = \mathcal{O}\big((\Delta x)^2\big), \quad R_{I_h,q,n}^k = \mathcal{O}\big((\Delta x)^2\big).$$

Also, by the error estimate of the Gauss quadrature rule [31], for $f \in C^r$ and $0 < \epsilon < 1$ we have

$$\left|\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} f(\xi)e^{-\xi^2/2}d\xi - \sum_{\ell=1}^{L} f(\xi_\ell)\omega_\ell\right| \leqslant \frac{CL^{-r/2}}{\sqrt{2\pi}}\int_{\mathbb{R}}|f^{(r)}(\xi)e^{-(1-\epsilon)\xi^2/2}|d\xi,$$

where the constant C is dependent on r while it is independent of L and f. Therefore, we have that

$$\begin{aligned}\big|R^k_{E,p,n}\big| &\leqslant C\sigma\big(x_k,u(t_n)\big)^4(\Delta t)^2,\\ \big|R^k_{E,q,n}\big| &\leqslant C\sigma\big(x_k,u(t_n)\big)^4(\Delta t)^{5/2}+C\sigma\big(x_k,u(t_n)\big)^3(\Delta t)^2.\end{aligned}$$

Then, by Proposition 4.2, we have

$$\hat{\mathbb{E}}[(R_{E,p,n})^2]=\mathcal{O}\big((\Delta t)^4\big),\quad \hat{\mathbb{E}}[(R_{E,q,n})^2]=\mathcal{O}\big((\Delta t)^4\big).$$

For $\tilde{R}^k_{p,n}$ and $\tilde{R}^k_{q,n}$, a rough estimation follows by the Taylor expansion:

$$\mathbb{E}^{x_k}_{t_n}[p_{t_{n+1}}]=p^{x_k}_{t_{n+1}}+\Delta t\,\mathcal{L}\,\eta(t_{n+1},x_k)+\int_{t_n}^{t_{n+1}}\int_{t_n}^{s}\mathbb{E}^{x_k}_{t_n}[\mathcal{L}\mathcal{L}\,\eta(t_{n+1},x_r)]drds, \tag{54}$$

$$\mathbb{E}^{x_k}_{t_n}[\tilde{p}_{t_{n+1}}]=p^{x_k}_{t_{n+1}}+\Delta t\,\tilde{\mathcal{L}}\,\eta(t_{n+1},x_k)+\int_{t_n}^{t_{n+1}}\int_{t_n}^{s}\mathbb{E}^{x_k}_{t_n}[\tilde{\mathcal{L}}\tilde{\mathcal{L}}\,\eta(t_{n+1},\tilde{x}_r)]drds; \tag{55}$$

then we can deduce that $\tilde{R}^k_{p,n}=\mathcal{O}\big((\Delta t)^2\big)$, where

$$\begin{aligned}\mathcal{L}\,\phi(x_r)&=b\big(x_r,u(r)\big)\phi'(x_r)+\frac{1}{2}\sigma\big(x_r,u(r)\big)^2\phi''(x_r),\\ \tilde{\mathcal{L}}\,\phi(\tilde{x}_r)&=b\big(x_{t_n},u(t_n)\big)\phi'(\tilde{x}_r)+\frac{1}{2}\sigma\big(x_{t_n},u(t_n)\big)^2\phi''(\tilde{x}_r).\end{aligned}$$

Similarly, for $\tilde{R}^k_{q,n}$ we can derive that $\tilde{R}^k_{q,n}=\mathcal{O}\big((\Delta t)^2\big)$. Finally, for the semi-discretization error, we have

$$\begin{aligned}\bar{R}^{x_k}_{p,n}&=\int_{t_n}^{t_{n+1}}\int_{t_n}^{t}\mathbb{E}^{x_k}_{t_n}[\mathcal{L}^0\bar{f}(s,x_s)]dsdt,\\ \bar{R}^{x_k}_{q,n}&=\int_{t_n}^{t_{n+1}}\int_{t_n}^{t}\mathbb{E}^{x_k}_{t_n}[\mathcal{L}^0\bar{f}(s,x_s)\Delta W_{n+1}+\mathcal{L}^1\bar{f}(s,x_s)-\mathcal{L}^0\zeta(s,x_s)]dsdt,\end{aligned}$$

where $\bar{f}(t,x)=f\big(x,\eta(t,x),\zeta(t,x),u(t)\big)$. Thus, by recalling that $u\in U_N$, we have

$$\bar{R}^{x_k}_{p,n}=\mathcal{O}\big((\Delta t)^2\big),\quad \bar{R}^{x_k}_{q,n}=\mathcal{O}\big((\Delta t)^2\big).$$

Then, the desired result follows by combining all the estimates above. ∎

By the above arguments (Lemmas 4.3–4.5), we can finally get the following error estimates for our numerical schemes.

THEOREM 4.6. *Under Assumption* 2.1 *and the conditions in Lemmas* 4.3–4.5*, it holds that*

$$\begin{aligned}\hat{\mathbb{E}}[(\mu_n)^2]+\Delta t\sum_{n=0}^{N-1}\hat{\mathbb{E}}\big[(\nu_n)^2\big]&=\mathcal{O}\big((\Delta t)^2\big)+\mathcal{O}\big((\Delta x)^4/(\Delta t)^2\big),\\ \sup_i\|J'(u^{N,i})-J'_N(u^{N,i})\|&=\mathcal{O}(\Delta t)+\mathcal{O}\big((\Delta x)^2/\Delta t\big).\end{aligned}$$

In particular, if we have $\Delta x=\Delta t$, and we suppose that the assumptions in Theorem 3.1 *and Corollary* 3.2 *hold, then it holds that*

$$\sup_i\|J'(u^{N,i})-J'_N(u^{N,i})\|=\mathcal{O}(\Delta t),\quad \|u^*-u^{N,i}\|=\mathcal{O}(\Delta t),\quad i\to\infty.$$

Proof. Given $u \in U_N$, we define

$$\begin{aligned}\phi_t &= p_t\, b'_u\big(x_t, u(t)\big) + q_t\, \sigma'_u\big(x_t, u(t)\big) + j'\big(u(t)\big),\\ \phi_n^k &= p_n^k b'_u\big(x_k, u(t_n)\big) + q_n^k \sigma'_u\big(x_k, u(t_n)\big) + j'\big(u(t_n)\big).\end{aligned}$$

Then by the assumptions, we have $\bar{\phi} \in C_b^{1,4}$ in $[t_n, t_{n+1}) \times \mathbb{R}$, where $\bar{\phi}$ is such that $\phi_t = \bar{\phi}(t, x_t)$. Moreover, $J'(u)|_t = \mathbb{E}[\phi_t]$, $J'_N(u)|_{t_n} = \hat{\mathbb{E}}[\phi_n]$. Then, we have

$$\begin{aligned}&\|J'(u) - J'_N(u)\|^2\\ &\leqslant C \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \big(J'(u)|_t - J'(u)|_{t_n}\big)^2 + \big(J'(u)|_{t_n} - J'_N(u)|_{t_n}\big)^2 dt\\ &= C \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \Big(\int_{t_n}^{t} \frac{d}{dr}\mathbb{E}[\phi_r]\Big|_{r=s} ds\Big)^2 dt \,+ C\Delta t \sum_{n=0}^{N-1} \big(\mathbb{E}[\phi_{t_n}] - \hat{\mathbb{E}}[\phi_n]\big)^2\\ &\leqslant C\Delta t \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}}\!\int_{t_n}^{t} \Big(\mathbb{E}[\mathcal{L}^0 \bar{\phi}(s, x_s)]\Big)^2 ds\, dt\\ &\quad\ + C\Delta t \sum_{n=0}^{N-1} \Big(\big(\mathbb{E}[\phi_{t_n}] - \hat{\mathbb{E}}[\phi_{t_n}]\big)^2 + \big(\hat{\mathbb{E}}[\phi_{t_n}] - \hat{\mathbb{E}}[\phi_n]\big)^2\Big)\\ &\leqslant C(\Delta t)^2 + C(\Delta x)^4/(\Delta t)^2 + C\Delta t \sum_{n=0}^{N-1} \hat{\mathbb{E}}\big[(\mu_n)^2 + (\nu_n)^2\big]\\ &= \mathcal{O}\big((\Delta t)^2\big) + \mathcal{O}\big((\Delta x)^4/(\Delta t)^2\big).\end{aligned}$$

Notice that the above estimations are valid since under Assumption 2.1 each estimate hold true uniformly in u. We complete the proof. ∎

5. Numerical experiments. In this section, we present several numerical examples to verify the efficiency of our numerical approach. In all our computations, we need to choose a reasonable parameter ρ. Motivated by the error estimates in the last section, we notice that the scheme admits a good convergence property with sufficiently small ρ. However, extremely small ρ would decrease the convergence rate of the iteration. In our examples, we shall simply choose $\rho_i = 1/\sqrt{i}$. And in what follows, we shall denote by "CR" the convergence rate.

Example 1. Our first example has been used in [12]. The optimal control problem is stated as

$$J(u^*) = \min_{u \in K} J(u),$$

with the cost functional

$$J(u) = \frac{1}{2}\int_0^T \mathbb{E}\big[\big(x_t - x^*(t)\big)^2\big]dt + \frac{1}{2}\int_0^T u^2(t)dt, \qquad K = U,$$

and the controlled state equation

$$dx_t = u(t)x_t\, dt + \sigma x_t\, dW_t.$$

Here σ is a constant. The deterministic function x^* and the corresponding exact solution u^* are given by

$$u^*(t) = \frac{T-t}{\dfrac{1}{x_0} - Tt + \dfrac{t^2}{2}}, \qquad x^*(t) = \frac{e^{\sigma^2 t} - (T-t)^2}{\dfrac{1}{x_0} - Tt + \dfrac{t^2}{2}} + 1. \tag{56}$$

We set $x_0 = 1$, $T = 1$, and $\sigma = 0.1$, and the number of samples for approximating the expectation is chosen as $M = 10^5$, and we set the tolerance as $\epsilon_0 = 10^{-5}$. Numerical results by our gradient projection method are presented in Figure 1.

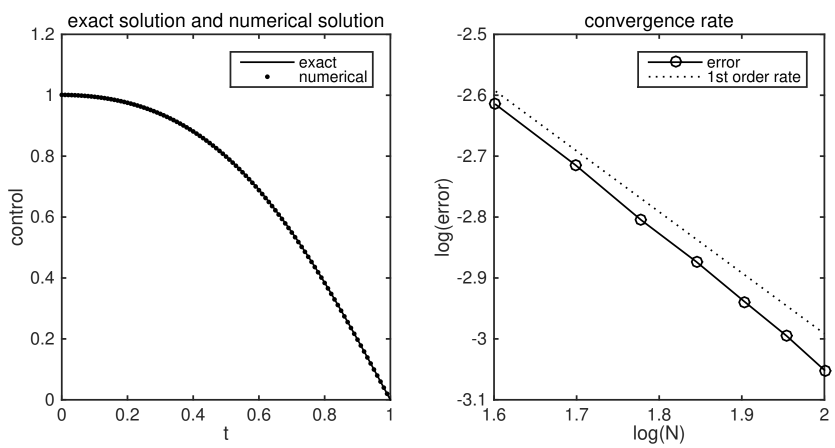

FIG. 1. *Numerical results for Example* 1 *with solution* (56).

The left plot shows that the numerical solution matches the exact solution very well when $N = 100$. In the right plot, we have tested the error decays with $N = 40, 50, \dots, 100$, and it is clearly shown that the method admits a first order rate of convergence.

Next, we test a different pair (x^*, u^*) which is given by

$$u^*(t) = \frac{e^{-T} - e^{-t}}{\dfrac{1}{x_0} + 1 - e^{-t} - te^{-T}}, \qquad x^*(t) = \frac{e^{\sigma^2 t} - (e^{-T} - e^{-t})^2}{\dfrac{1}{x_0} + 1 - e^{-t} - te^{-T}} - e^{-t}. \tag{57}$$

We set $\sigma = 0.1$, $M = 10^5$, $\epsilon_0 = 10^{-5}$, and $N = 40, 50, \dots, 100$. The numerical results are given in Figure 2. Again, the numerical solution matches the exact solution very well, and a first order convergence rate is observed.

Example 2. Our second example is also from [12]. More precisely, we consider

$$J(u^*) = \min_{u \in K} J(u),$$

with

$$\begin{aligned} &J(u) = \frac{1}{2}\int_0^T \mathbb{E}\big[\big(x_t - x^*(t)\big)^2\big]dt + \frac{1}{2}\int_0^T u^2(t)dt, \qquad K = U, \\ &dx_t = \big(u(t) - r(t)\big)dt + \sigma u(t) dW_t. \end{aligned}$$

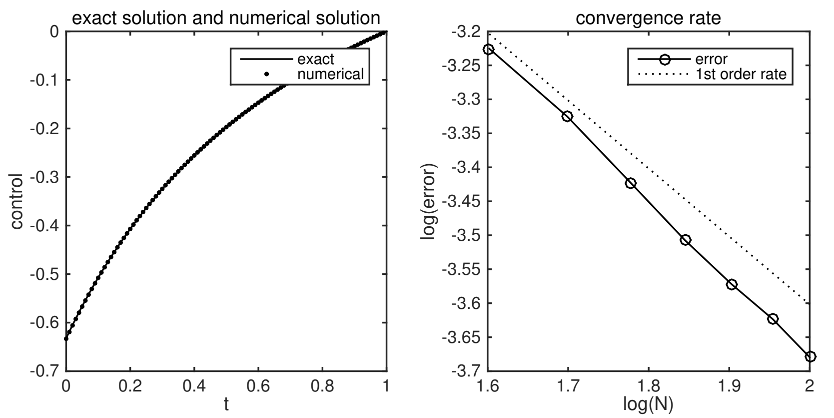

FIG. 2. *Numerical results for Example* 1 *with solution* (57).

Here we set $r(t) = u^*(t)/2$, $x_0 = 0$, and $T = 1$, and σ is a constant. The deterministic function x^* and the corresponding exact solution u^* are chosen as

$$u^*(t) = \frac{T-t}{\sigma^2(T-t)+1}, \qquad x^*(t) = \frac{t}{2\sigma^2} - \frac{1}{2\sigma^4}\ln\frac{\sigma^2 T+1}{\sigma^2(T-t)+1} + 1.$$

In our computations, we choose $\sigma = 0.1$, $M = 10^5$, $\epsilon_0 = 10^{-5}$, and $N = 40, 50, \ldots, 100$. The numerical results are shown in Figure 3. Similar conclusions can be made as for Example 1. The method converges with the first order accuracy.

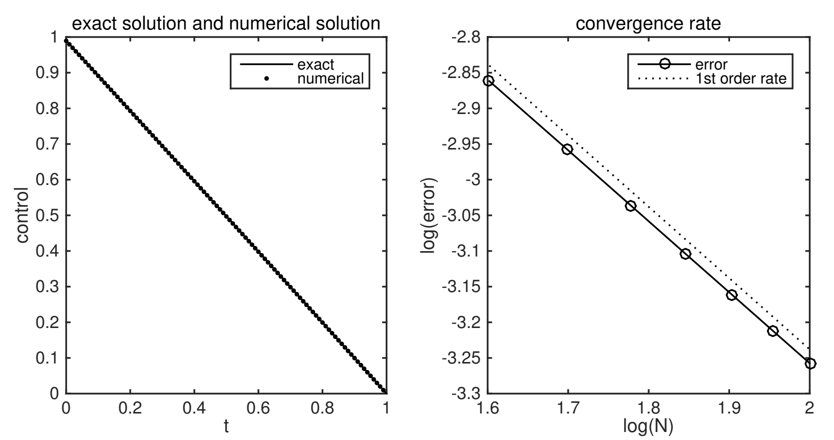

FIG. 3. *Numerical results for Example* 2.

Example 3. The previous discussions have focused on the deterministic control, that is, $u \in U$. In this example, we will show that our method can also be used to solve SOCPs with feedback control.

This example is set to be the same as in (1)–(2), except that the control constraint set is now a set of stochastic controls:

$$K_{\mathcal{F}} = \{u \in U_{\mathcal{F}} \,\big|\, u_t(\omega) \in \mathcal{C} \text{ a.e. a.s.}\}. \tag{58}$$

It follows from stochastic optimal control theory that the optimal control is actually a feedback control; more precisely, there exists a function $\bar{u}^*$ such that $u_t^* = \bar{u}^*(t, x_t)$ (see, e.g., [46, 14]). Given a feedback control u with $u_t = \bar{u}(t, x_t)$, by introducing the adjoint processes (p, q) in the same way as in the deterministic case, and by applying the Itô formula, we can show that

$$J'(u)_t = p_t b'_u(x_t, u_t) + q_t \sigma'_u(x_t, u_t) + j'(u_t). \tag{59}$$

Notice that u_t is a function of t and x_t; then by (59) we know that $J'(u)_t$ is also a function of t and x_t. Therefore, due to the feedback property of the control, we can write $J'(u)$ pointwisely in time-space grids, namely,

$$J'(u)_{t_n}^x = p_{t_n}^x b'_u\big(x, \bar{u}(t_n, x)\big) + q_{t_n}^x \sigma'_u\big(x, \bar{u}(t_n, x)\big) + j'\big(\bar{u}(t_n, x)\big), \tag{60}$$

where $x \in D_h$ and $J'(u)_t^x$ denotes $J'(u)_t$ valued at $x_t = x$. In the above equation, by introducing our numerical solutions p_n and q_n, we get the approximated $J'_N(\cdot)$ of $J'(\cdot)$:

$$J'_N(u)_n^k = p_n^k b'_u\big(x_k, \bar{u}(t_n, x_k)\big) + q_n^k \sigma'_u\big(x_k, \bar{u}(t_n, x_k)\big) + j'\big(\bar{u}(t_n, x_k)\big). \tag{61}$$

Since the constraint K (58) is also pointwise in time and space, the projection problem at the grid point (t_n, x), $x \in D_h$, can be written as

$$\bar{u}^*(t_n, x) = P_{\mathcal{C}}\big(\bar{u}^*(t_n, x) - \rho J'(u^*)_{t_n}^x\big).$$

Here we shall not compute the feedback law explicitly; however, we do compute the values of the control at the grid point. Then u^* is updated in the following way:

$$\bar{u}^{i+1}(t_n, x_k) = P_{\mathcal{C}}\big(\bar{u}^i(t_n, x_k) - \rho_i J'_N(u^i)_n^k\big). \tag{62}$$

Notice that due to the change of the space of control, we get rid of the expectation in the computation of $J'(u)$, meaning that we no longer need the history information before time t to compute $J'(u)_t$, but only the information at time instance t. Consequently, if a proper space partition $\{x_k\}_k$ is obtained, and the constraint K is pointwise in time, then we can run the algorithm in a backward manner as described in Algorithm 2. Compared to Algorithm 1, we notice that under the same spatial partition, Algorithm 2 can save a lot of restoration.

Algorithm 2 Gradient projection method

Set the initial guess of the control $\{\bar{u}(t_n, x_k)\}_{n,k}$ and the error tolerance ϵ_0;
1. Set the terminal condition: $p_N^k = g(x_k)$, $x_k \in \mathbb{R}_h$;
2. For $n = N-1, \dots, 1, 0$, do
   a. solve (p_n, q_n) by (41)–(42);
   b. compute $J'_N(u)_n^k$ by (61);
   c. update u by (62);

Repeat a–c until $\sup_k |\bar{u}^{i+1}(t_n, x_k) - \bar{u}^i(t_n, x_k)| \le \epsilon_0$.

We now test Algorithm 2 for Example 3 with K defined in (58), and compare the results using feedback control with the results obtained by using the deterministic control. For the feedback control, we shall use the rectangular rule and the Monte Carlo method to compute the integral and the expectation of the objective functional, respectively. The numerical results are listed in Table 1. It is shown that the use of feedback control can indeed improve the results (it produces a smaller value of objective functional), and this is reasonable as we are minimizing the objective functional within a larger control set.

TABLE 1
Numerical results for Example 3.

N	$J(u)$ with Algorithm 1	$J(u)$ with Algorithm 2
100	0.84833	0.62535
200	0.84797	0.64507
400	0.84777	0.65509
800	0.84770	0.66013

Example 4. Our last example is a portfolio problem. We consider the following example, which was used in [9]:

$$J(u^*) = \min_{u \in K} J(u),$$

with

$$J(u) = \frac{1}{2}\mathbb{E}\left[(x_T - \kappa)^2\right], \qquad K = \left\{u \in U_{\mathcal{F}};\ -1 \leqslant u_t \leqslant 1,\ \text{a.e. a.s.}\right\},$$
$$dx_t = (\zeta\sigma u_t + r)x_t\, dt + \sigma u_t x_t\, dW_t.$$

The parameters are chosen as

$$T = 50,\ \kappa = 1000,\ x_0 = 300,\ r = 0.02,\ \sigma = 0.1,\ \zeta = 0.05.$$

We set $\epsilon_0 = 10^{-4}$, $L = 4$, and $\rho_i = 0.01/i$, and the space region is given by $[-100, 900]$. The optimal value of $J(u)$ given in [9] is 15023. To show the convergence rate, we perform experiments with $N = 1000, 2000, 4000, 8000$, and we choose $M = N^2/10$. The corresponding numerical solutions for $J(u)$ are listed in Table 2. It is clear that the method admits a first order rate of convergence. This example shows that Algorithm 2 is capable of solving some optimal control problems involving feedback control.

TABLE 2
Numerical results for Example 4.

N	1000	2000	4000	8000	Optimal
$J(u)$	15196	15107	15069	15045	15023
CR	-	1.0423	0.8688	1.0641	

6. Conclusion.

In this work, we propose a gradient projection method for solving stochastic optimal control problems. The scheme contains a fixed-point iteration of the control and an Euler scheme for solving the adjoint equation that is given by BSDEs. The Euler method is used to solve the adjoint BSDEs. We rigorously prove that our numerical method admits a first order rate of convergence. Several numerical tests are presented to support our theoretical finding.

REFERENCES

[1] O. Bahn, A. Haurie, and R. Malhame, *A stochastic control model for optimal timing of climate policies*, Automatica J. IFAC, 44 (2008), pp. 1545–1558.
[2] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser Boston, Boston, 1997.
[3] K. Barty, J.-S. Roy, and C. Strugarek, *A stochastic gradient type algorithm for closed-loop problems*, Math. Program., 119 (2009), pp. 51–78.

[4] A. Bensoussan, *Lecture on stochastic control*, in Nonlinear Filtering and Stochastic Control, Lecture Notes in Math. 972, Springer-Verlag, Berlin, New York, 1982, pp. 1–62.
[5] A. Bensoussan, *Stochastic Control by Functional Analysis Methods*, North-Holland, New York, 1982.
[6] D.P. Bertsekas, *Dynamic Programming and Optimal Control*, Vol. II., Athena Scientific, Belmont, MA, 2007.
[7] F. Campillo, J. Nekkachi, and E. Pardoux, *Numerical methods in ergodic stochastic control: Application to semi-active vehicle suspensions*, in Proceedings of the 29th IEEE Conference on Decision and Control (Honolulu, HI), IEEE, Washington, DC, 1990.
[8] P. Carpentier, G. Cohen, and A. Dallagi, *Particle methods for stochastic optimal control problems*, Comput. Optim. Appl., 56 (2013), pp. 635–674.
[9] J. Chen, *Some Problems of Optimal Control and the Application in Financial Mathematics*, Master's thesis, Tongji University, Shanghai, China, 2007.
[10] M. Crandall and P. L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277 (1983), pp. 1–42.
[11] A. Dallagi, *Méthodes particulaires en commande optimale stochastique*, Thèse de doctorat, Université Paris 1, Panthéon-Sorbonne, Paris, France, 2007.
[12] N. Du, J. T. Shi, and W. B. Liu, *An effective gradient projection method for stochastic optimal control*, Int. J. Numer. Anal. Model., 4 (2013), pp. 757–774.
[13] R. J. Elliott, *The optimal control of diffusions*, Appl. Math. Optim., 22 (1990), pp. 229–240.
[14] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer, New York, 2012.
[15] P. Girardeau, *A comparison of sample-based stochastic optimal control methods*, Stochastic Program. E-Print Ser., 2010 (2010), https://doi.org/10.18452/8408.
[16] 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.
[17] M. Gunzburger, *Perspectives in Flow Control and Optimization*, SIAM, Philadelphia, 2003, https://doi.org/10.1137/1.9780898718720.
[18] M. Gunzburger, H. Lee, and J. Lee, *Error estimates of stochastic optimal Neumann boundary control problems*, SIAM J. Numer. Anal., 49 (2011), pp. 1532–1552, https://doi.org/10.1137/100801731.
[19] F. B. Hanson, *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation*, SIAM, Philadelphia, 2007.
[20] U. G. Haussmann, *Some examples of optimal stochastic controls or: The stochastic maximum principle at work*, SIAM Rev., 23 (1981), pp. 292–307, https://doi.org/10.1137/1023062.
[21] H. Heitsch and W. Römisch, *Scenario Reduction Algorithms in Stochastic Programming*, Comput. Optim. Appl., 24 (2003), pp. 187–206.
[22] H. Heitsch, W. Römisch, and C. Strugarek, *Stability of multistage stochastic programs*, SIAM J. Optim., 17 (2006), pp. 511–525, https://doi.org/10.1137/050632865.
[23] C. S. Huang, S. Wang, and K. L. Teo, *Solving Hamilton-Jacobi-Bellman equations by a modified method of characteristics*, Nonlinear Anal., 40 (2000), pp. 279–293.
[24] R. Korn, *Optimal Portfolios: Stochastic Models for Optimal Investment and Risk Management in Continuous Time*, World Scientific, Singapore, 1997.
[25] R. Korn and H. Kraft, *A stochastic control approach to portfolio problems with stochastic interest rates*, SIAM J. Control Optim., 40 (2001), pp. 1250–1269, https://doi.org/10.1137/S0363012900377791.
[26] D. Kuhn, P. Parpas, and B. Rustem, *Stochastic optimization of investment planning problems in the electric power industry*, in Energy Systems Engineering, Process Systems Engrg. 5, Wiley-VCH Verlag, Weinheim, Germany, 2008, pp. 215–230.
[27] H. J. Kushner, *Necessary conditions for continuous parameter stochastic optimization problems*, SIAM J. Control, 10 (1972), pp. 550–565, https://doi.org/10.1137/0310041.
[28] H. J. Kushner and P. Dupuis, *Numerical Methods for Stochastic Control Problems in Continuous Time*, 2nd ed., Springer-Verlag, New York, 2001.
[29] A. Li, E. M. Feng, and X. L. Sun, *Stochastic optimal control and algorithm of the trajectory of horizontal wells*, J. Comput. Appl. Math., 212 (2008), pp. 419–430.
[30] J. Ma and J. Yong, *Forward-Backward Stochastic Differential Equations and Their Applications*, Springer, New York, 2007.
[31] G. Mastroianni and G. Monegato, *Error estimates for Gauss-Laguerre and Gauss-Hermite quadrature formulas*, in Approximation and Computation: A Festschrift in Honor of Walter Gautschi, Birkhäuser Boston, Boston, 1994, pp. 421–434.
[32] R. C. Merton, *Continuous-Time Finance*, Blackwell, Cambridge, MA, 1992.

[33] G. N. Milstein and M. V. Tretyakov, *Numerical algorithms for forward-backward stochastic differential equations*, SIAM J. Sci. Comput., 28 (2006), pp. 561–582, https://doi.org/10.1137/040614426.

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

[35] S. G. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim., 28 (1990), pp. 966–979, https://doi.org/10.1137/0328054.

[36] S. G. Peng, *Probabilistic interpretation for systems of quasilinear parabolic partial differential equations*, Stochastics Stochastics Rep., 37 (1991), pp. 61–74.

[37] S. G. Peng, *Backward stochastic differential equations and applications to optimal control*, Appl. Math. Optim., 27 (1993), pp. 125–144.

[38] R. Raffard, J. Hu, and C. Tomlin, *Adjoint-based optimal control of the expected exit time for stochastic hybrid systems*, in Hybrid Systems: Computation and Control, M. Morari and L. Thiele, eds., Lecture Notes in Comput. Sci. 3414, Springer-Verlag, Berlin, 2005, pp. 557–572.

[39] L. C. G. Rogers, *Pathwise stochastic optimal control*, SIAM J. Control Optim., 46 (2007), pp. 1116–1132, https://doi.org/10.1137/050642885.

[40] S. P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser Verlag, Basel, 1994.

[41] A. Shapiro and A. Ruszczynski, eds., *Stochastic Programming*, Elsevier, Amsterdam, 2003.

[42] J. T. Shi and Z. Wu, *A stochastic maximum principle for optimal control of jump diffusions and applications to finance*, Chinese J. Appl. Probab. Statist., 27 (2011), pp. 127–137.

[43] J. C. Spall, *Stochastic optimization* II.6, in Handbook of Computational Statistics, J. Gentle, W. Härdle, and Y. Mori, eds., Springer-Verlag, New York, 2004, pp. 169–197.

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

[45] W. S. Xu, *Maximum principle for a stochastic optimal control problem and application to portfolio/consumption choice*, J. Optim. Theory Appl., 98 (1998), pp. 719–731.

[46] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer, New York, 1999.

[47] 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, https://doi.org/10.1137/05063341X.

[48] 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, https://doi.org/10.1137/130941274.

[49] 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.