INTERNATIONAL JOURNAL OF
NUMERICAL ANALYSIS AND MODELING
Volume 10, Number 4, Pages 757–774

AN EFFECTIVE GRADIENT PROJECTION METHOD FOR STOCHASTIC OPTIMAL CONTROL

NING DU, JINGTAO SHI, AND WENBIN LIU

Abstract. In this work, we propose a simple yet effective gradient projection algorithm for a class of stochastic optimal control problems. The basic iteration block is to compute gradient projection of the objective functional by solving the state and co-state equations via some Euler methods and by using the Monte Carlo simulations. Convergence properties are discussed and extensive numerical tests are carried out. Possibility of extending this algorithm to more general stochastic optimal control is also discussed.

Key words. stochastic optimal control, numerical method, gradient projection algorithm

1. Introduction

Stochastic optimal control is an essential tool for developing and analyzing models that have stochastic dynamics, and it has been fully developed both theoretically and practically in mathematics, physics and engineering. There has existed a very extensive body of literature in this area, and it is impossible to present an even very brief review on its development here. Some introductive accounts (more from mathematical points of view) can be found, for example, in [5, 6, 18, 30], and [11, 14]. Some of the research relevant to our work can be found in [5, 15, 18, 29], and [4, 10, 16, 31, 41, 42, 43]. Practical examples of stochastic optimal control include engineering systems [10, 27, 31, 43, 45], option pricing and portfolio optimization models from finance [25, 26, 37, 47, 50], analysis of climate change policies [1], and biological and medical applications [17].

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$ be a complete probability space with the natural filtration $\{\mathcal{F}_t\}_{t\geq 0}$, which is generated by a one-dimensional standard Brownian motion $\{W_t\}_{t\geq 0}$. Let $T > 0$ be a fixed real number that is called time horizon. We denote by $L^2(\Omega, \mathcal{F}_T; R)$ the space of real-valued square-integrable $\mathcal{F}_T$-measurable random variables, and by $L^2_{\mathcal{F}}([0,T]; R)$ the space of real-valued square-integrable $\mathcal{F}_t$-adapted processes such that

$$E\Big\{\int_0^T |y_t|^2 dt\Big\} < +\infty. \tag{1}$$

In this paper we consider numerical solutions to the following stochastic control problem. The objective functional

$$J(y,u) = \int_0^T E[h(y)]dt + \int_0^T j(u)dt, \tag{2}$$

where h and j are smooth functions with the continuous first order derivatives, $u \in U_{ad}$ is a deterministic control, where U_{ad} is a close convex set in the control space $L^2(0,T)$.

Received by the editors December 17, 2012 and, in revised form, January 30, 2013.
2000 *Mathematics Subject Classification.* 49K15, 65L12, 93E20.
This research was supported by the National Natural Science Foundation of China under the Grant Nos. 10801092, 91130010, 11201265, and 11201264.

An admissible control u^* is called optimal if it attains the minimum of $J(y(u), u)$, where the state $y(u) \in L^2_{\mathcal{F}}([0,T]; R)$ is a stochastic process which is generated by

$$dy = f(t, y, u)dt + g(t, y)dW_t, \quad y(0) = y_0. \tag{3}$$

In this paper we assume that f, g are continuously differentiable with respect to (t, y, u) and (t, y), respectively, and that their derivatives are bounded.

Under the above assumptions, we know equation (3) admits a unique solution $y(\cdot) \in L^2_{\mathcal{F}}([0,T]; R)$ for the given $(y_0, u(\cdot)) \in R \times U_{ad}$ (see [22]). We call such a $y(\cdot)$ the corresponding trajectory. Let us note that here the control does not appear in the diffusion term for ease of exposition. For the general case, one would need to have more theoretical preparations on backward stochastic differential equations for a rigorous treatment of the adjoint state equations (see [6, 41, 42]), although our methods are still applicable.

In general most realistic models do not admit closed form solutions and thus effective numerical methods play a key role for practical applications of stochastic optimal control. In the literature numerous numbers of numerical methods have been proposed for stochastic optimal control and the related problems. Numerical methods used to solve stochastic optimal control have at least four broad classes: Those transferring the control problem into finite dimensional stochastic programming, see, e.g., [12, 16, 19, 20, 31, 43, 46, 48]; those based on Dynamic Programming Principle (DPP), see e.g. [7, 30], in particular those solving HJB equations for the feedback solutions - there are many references in this area, see [2, 5, 6, 13, 21] for some early work; the third class based on martingale methods, see e.g. [25, 26, 44]; and those based on the Stochastic Maximum Principle (SMP) [18]. The method proposed in this paper is based on an iterative algorithm for the solution of the SMP. There exists extensive research on the first three classes methods. Although the SMP is widely used in solving the stochastic optimal control, see, e.g., [47] and [50], it is not often used in numerical algorithms yet. The likely reasons are that it will not directly produce the feedback control as explained below, and the computation of the adjoins requires the solution of a backwards stochastic differential equation (BSDE), which is computationally expensive.

Compared with the deterministic optimal control, stochastic optimal control is much more complicated from the perspective of obtaining numerical solutions that are realizable to real applications. One of the reasons is that often the value of optimal control $u(t)$ at a time t will depend on ω (so $u(t, \omega)$) so that it is not very useful to only compute and then apply the numerical solutions of the optimal control like in the deterministic case. To be practically useful, some forms of feedback relationship between the optimal state and optimal control need to be computed numerically as well (as in the approach of Bellman Equation), as otherwise the optimal control is difficult to realize. Therefore the existing numerical methods in the literature are rather complex. In this paper we study a useful case where the control is deterministic (but the state is still stochastic) as the first step towards developing fast numerical algorithms for general stochastic optimal control. In this case the optimal control does not directly depend on ω (but depends on $y(t)$) so that it is meaningful to just compute the optimal control and apply it without the feedback laws. This is quite desirable in some business and engineering decision making where the stochastic effect is not overwhelming and thus deterministic decision rules are desirable and sufficient. A deterministic solution is also useful for future planning. Such examples can be found e.g. in [10, 31] (Engineering Control), [12] (Financial) and [43] (Stochastic Hybrid Systems). In this work we are then able to derive simple yet effective numerical algorithms with convergence analysis. More importantly these algorithms can be extended to the case where the optimal control is stochastic as to be seen later, although further study like feedback law regression will then be needed, which will be investigated in the next step. Thus this study will pave the way to developing effective methods for the general case.

The method proposed in this paper utilizes SMP and orbit Monte Carlo simulation. We utilize the adjoint state to compute the gradient efficiently and then search the next descending direction using the negative gradient. For computation of the adjoint state, we do not directly use algorithms for a BSDE but utilize the special constructs of the adjoint equation so that we can compute its numerical solutions quickly. Furthermore this method does not need to use any PDE or DPP, and thus is quite simple to implement. More importantly it has the potential capability of handling high dimensional problems and is easy to parallelize for large scale problems. We will present extensive numerical tests to show its features.

The plan of this paper is as follows: In Section 2, we introduce the algorithm and examine its convergence. In Section 3, we present extensive numerical tests to verify this algorithm. Conclusions are summarized in Section 4. Finally, in Section 5, we describe the details of designing the numerical examples.

2. A Gradient Algorithm for Stochastic Optimal Control

Let $J(u) = J(y(u), u)$, where $y(u)$ is the solution of (3), to be referred to as the reduced objective functional. We first state this projection method for the reduced control problem in a general space. Let U be a real Hilbert space with $U' = U$, and K be a closed convex subset of U. Consider

$$\min_{u \in K} J(u) \tag{4}$$

where $J(u)$ is a convex functional on U. The widely used necessary and sufficient optimality condition of (4) reads (see [32])

$$(J'(u), v - u) \geq 0, \qquad \forall v \in K, \tag{5}$$

where $J' \in U$ is the G-differential of J.

Let $b(\cdot,\cdot)$ be a symmetric and positive definite bilinear form such that there exist constants c_0 and c_1 satisfying

$$|b(u, v)| \leq c_1 \|u\|_U \|v\|_U, \ \ \forall u, v \in U, \tag{6}$$

$$b(u, u) \geq c_0 \|u\|_U^2. \tag{7}$$

Define operator $\mathbf{b} : U \to U$ by

$$(\mathbf{b}u, v) = b(u, v), \quad \forall u, v \in U. \tag{8}$$

It is clear that the norm $\|\cdot\|_b = \sqrt{b(\cdot,\cdot)}$ is equivalent to the norm $\|\cdot\|_U$ by the assumptions.

Now define the projection operator P_K^b $U \to K$: For a given $w \in U$, find $P_K^b w \in K$ such that

$$b(P_K^b w - w, P_K^b w - w) = \min_{u \in K} b(u - w, u - w), \tag{9}$$

which is equivalent to

$$b(P_K^b w - w, v - P_K^b w) \geq 0, \qquad \forall v \in K. \tag{10}$$

It is clear that P_K^b is well-defined for any closed convex subset in U.

It follows from (5) that for any $\rho > 0$ the solution u of (4) reads:

$$u = P_K^b(u - \rho \mathbf{b}^{-1} J'(u)). \tag{11}$$

For the reduced control problem (4), define an iterative scheme ($n = 0, 1, 2, \ldots$):

$$\left\{\begin{array}{l} b(u_{n+\frac{1}{2}}, v) = b(u_n, v) - \rho_n (J_n'(u_n), v), \quad \forall v \in U, \\ u_{n+1} = P_K^b\left(u_{n+\frac{1}{2}}\right). \end{array}\right. \tag{12}$$

where J_n' is the approximate functional of J' in the nth iteration. One can extend this scheme to the case where b also depends on n. Then the Newton method is included as a special case where $\rho_n = 1$ and b is defined by the Hessian. In [33] this pre-conditional algorithm has been widely used to solve deterministic optimal control problems governed by various PDEs, and is shown quite efficient even for large scale problems.

For the above scheme we have the following convergence result:

Theorem 1. *Assume that J' is Lipschitz and uniformly monotone in the sense that there are positive constants c, C such that*

$$|J'(u) - J'(v)| \le C\|u - v\|_U, \ \forall u, v \in U, \tag{13}$$

$$(J'(u) - J'(v), u - v) \ge c\|u - v\|_U^2, \ \forall u, v \in U. \tag{14}$$

Furthermore, if J_n' is uniformly convergent to J' in the sense that

$$\|J_n'(u) - J'(u)\|_b \le \varepsilon_n \to 0, \tag{15}$$

then there exists an $\epsilon > 0$ such that

$$\lim_{n\to\infty} \|u - u_n\|_b = 0, \tag{16}$$

provided $\rho_n < \epsilon$.

Proof. For the iteration scheme (12) we have

$$u_{n+1} = P_K^b\left(u_n - \rho_n \mathbf{b}^{-1} J_n'(u_n)\right). \tag{17}$$

Using (11), we deduce

$$u_{n+1} - u = P_K^b\left(u_n - \rho_n \mathbf{b}^{-1} J_n'(u_n)\right) - P_K^b\left(u - \rho_n \mathbf{b}^{-1} J'(u)\right) \tag{18}$$

By (5) we have

$$\begin{array}{rcl} \|u_{n+1} - u\|_b^2 & = & \|P_K^b\left(u_n - \rho_n \mathbf{b}^{-1} J_n'(u_n)\right) - P_K^b\left(u - \rho_n \mathbf{b}^{-1} J'(u)\right)\|_b^2 \\ & \le & \|u_n - u - \rho_n \mathbf{b}^{-1}\left(J_n'(u_n) - J'(u)\right)\|_b^2 \\ & = & \|u_n - u\|_b^2 + \|\rho_n \mathbf{b}^{-1}\left(J_n'(u_n) - J'(u)\right)\|_b^2 \\ & & -2b(\rho_n \mathbf{b}^{-1}\left(J_n'(u_n) - J'(u)\right), u_n - u). \end{array} \tag{19}$$

Noting that

$$\begin{array}{rl} & \|\rho_n \mathbf{b}^{-1}\left(J_n'(u_n) - J'(u)\right)\|_b^2 \\ = & \|\rho_n \mathbf{b}^{-1}\left(J_n'(u_n) - J'(u_n) + J'(u_n) - J'(u)\right)\|_b^2 \\ \le & C\rho_n^2(\|u_n - u\|_b^2 + \varepsilon_n^2), \end{array} \tag{20}$$

$$\begin{array}{rl} & -2b(\rho_n \mathbf{b}^{-1}\left(J_n'(u_n) - J'(u)\right), u_n - u) \\ = & -2b(\rho_n \mathbf{b}^{-1}\left(J_n'(u_n) - J'(u_n) + J'(u_n) - J'(u)\right), u_n - u) \\ \le & -c\rho_n\|u_n - u\|_b^2 + 2\varepsilon_n\rho_n\|u_n - u\|_b \\ \le & -c\rho_n\|u_n - u\|_b^2 + \rho_n^2\|u_n - u\|_b^2 + \varepsilon_n^2, \end{array} \tag{21}$$

we have

$$\|u_{n+1}-u\|_b^2 \leq \left[1-c\rho_n\left(1-\frac{C+1}{c}\rho_n\right)\right]\|u_n-u\|_b^2+(C\rho_n^2+1)\varepsilon_n^2. \tag{22}$$

Choose $\rho_n < \epsilon$ to be sufficiently small such that

$$0 \leq 1-c\rho_n\left(1-\frac{C+1}{c}\rho_n\right) \leq \delta < 1, \tag{23}$$

and denote $\hat{\varepsilon}_n = (C\rho_n^2+1)\varepsilon_n^2$. Then it follows:

$$\|u_{n+1}-u\|_b^2 \leq \delta\|u_n-u\|_b^2+\hat{\varepsilon}_n. \tag{24}$$

It then can be derived that

$$\|u_{n+1}-u\|_b^2 \leq \delta^{n+1}\|u_0-u\|_b^2+\sum_{i=0}^{n}\hat{\varepsilon}_i\delta^{n-i}. \tag{25}$$

Let $S_n = \sum\limits_{i=0}^{n}\hat{\varepsilon}_i\delta^{n-i}$. We only need to show that S_n converges to zero when $n \to \infty$.

Since $\lim\limits_{n\to\infty}\hat{\varepsilon}_n=0$, then there exists a constant $K > 0$ such that $|\hat{\varepsilon}_n| \leq K$. Furthermore, $\forall\varepsilon > 0$, there exists a positive integer $m(m < n)$ satisfies

$$\hat{\varepsilon}_i < \frac{1-\delta}{2}\varepsilon, \tag{26}$$

so that

$$\begin{aligned} S_n &= \sum_{i=0}^{n}\hat{\varepsilon}_i\delta^{n-i} \\ &= \sum_{i=0}^{m}\hat{\varepsilon}_i\delta^{n-i}+\sum_{i=m+1}^{n}\hat{\varepsilon}_i\delta^{n-i} \\ &\leq K\frac{1-\delta^{m+1}}{1-\delta}\delta^{n-m}+\frac{1-\delta}{2}\varepsilon\frac{1-\delta^{n-m}}{1-\delta} \\ &\leq \frac{K}{1-\delta}\delta^{n-m}+\frac{\varepsilon}{2}. \end{aligned} \tag{27}$$

Since $0 < \delta < 1$, then for sufficiently large n, $\dfrac{K}{1-\delta}\delta^{n-m} \leq \dfrac{\varepsilon}{2}$ and $S_n \leq \varepsilon$. Then the convergence result holds.

□

In the actual computational process, the discrete functional J_n' often remains unchanged for a fixed partition, i.e., $J_n' = J_h'$, and the corresponding iterative scheme is

$$\begin{cases} b(u_{n+\frac{1}{2}},v) = b(u_n,v)-\rho_n(J_h'(u_n),v), \quad \forall v \in U, \\ u_{n+1} = P_K^b\left(u_{n+\frac{1}{2}}\right), \end{cases} \tag{28}$$

where h is the maximum step size of the partition. Therefore, we need to consider the stability of the current scheme:

Corollary 1. *Suppose all the conditions of Theorem 1 hold except that the convergence of J_n' to J' is replaced by*

$$\|J_h'(u)-J'(u)\|_b \leq \varepsilon, \tag{29}$$

where ε is a positive number, then there exists a constant $\hat{C} > 0$ such that

$$\|u-u_n\|_b < \delta^n\|u_0-u\|_b^2+\hat{C}\varepsilon^2. \tag{30}$$

Proof. The proof is almost the same as that of the above theorem, except that we replace J'_n by J'_h, and ε_n by ε. Thus, we have

$$\|u_{n+1} - u\|_b^2 \le \delta \|u_n - u\|_b^2 + (C\rho_n^2 + 1)\varepsilon^2 \le \delta \|u_n - u\|_b^2 + \bar{C}\varepsilon^2, \tag{31}$$

where $\bar{C}$ is an upper bound of $C\rho_n^2 + 1$. Then

$$\|u_{n+1} - u\|_b^2 \le \delta^{n+1} \|u_0 - u\|_b^2 + \bar{C}\varepsilon^2 \sum_{i=0}^{n} \delta^{n-i} \le \delta^{n+1} \|u_0 - u\|_b^2 + \frac{\bar{C}}{1-\delta}\varepsilon^2. \tag{32}$$

Let $\hat{C} = \dfrac{\bar{C}}{1-\delta}$, then the proof is finished. □

The above results can be extended to the case where J' is only Lipschitz and monotone locally. It follows from the proof that we still have convergence and the error estimate if u_0 and u are close enough.

It is possible to compute the gradient directly like in [10]. However it is in general inefficient. In order to apply this abstract algorithm to our stochastic control problem, we first compute the directional derivative of the objective functional by using the adjoint equation:

Let $u(\cdot)$ be an optimal control for (2)-(3), and $y(\cdot)$ the corresponding optimal trajectory. Let $v(\cdot) \in L^2(0,T)$ be given such that $v(\cdot) \in U_{ad}$. We take $v^\rho = u(\cdot) + \rho v(\cdot)$, $0 \le \rho \le 1$. Since U_{ad} is convex, $v^\rho(\cdot) \in U_{ad}$. Then we have

$$\begin{aligned} J'(u)(v) &= \lim_{\rho \to 0} \frac{J(u+\rho v) - J(u)}{\rho} \\ &= E\Big[\int_0^T h'(y) D(y)(v) dt\Big] + \int_0^T j'(u) v dt. \end{aligned} \tag{33}$$

where $v \in L^2(0,T)$ and

$$D(y)(v) = \lim_{\rho \to 0} \frac{1}{\rho}[y(u+\rho v) - y(u)]. \tag{34}$$

It follows that this derivative exists (see [15, 41, 47]). Noting that

$$dy = f(t,y,u)dt + g(t,y)dW_t, \quad y(0) = y_0, \tag{35}$$

and then

$$y = y_0 + \int_0^t f(s,y,u)ds + \int_0^t g(s,y)dW_s, \tag{36}$$

we find that $D(y)(v)$ is determined by

$$D(y)(v) = \int_0^t [f'_y(s,y,u)D(y)(v) + f'_u(s,y,u)v]ds + \int_0^t g'_y(s,y)D(y)(v)dW_s, \tag{37}$$

and then

$$d(D(y)(v)) = [f'_y(t,y,u)D(y)(v) + f'_u(t,y,u)v]dt + g'_y(t,y)D(y)(v)dW_t. \tag{38}$$

Define an adjoint state p backwards adapted such that $E\Big\{\int_0^T |p_t|^2 dt\Big\} < +\infty$

$$-dp = [h'(y) + pf'_y(t,y,u) - p(g'_y(t,y))^2]dt + pg'_y(t,y)dW_t, \quad p(T) = 0. \tag{39}$$

Using stochastic integration by parts formula, we have

$$
\begin{aligned}
J'(u)(v) &= E\Big[\int_0^T h'(y)D(y)(v)dt\Big] + \int_0^T j'(u)vdt \\
&= E\Big[\int_0^T \big(-dp - pf_y'(t,y,u)dt + p(g_y'(t,y))^2dt \\
&\qquad -pg_y'(t,y)dW_t\big)D(y)(v)\Big] + \int_0^T j'(u)vdt \\
&= E\Big[\int_0^T pd(D(y)(v)) + \int_0^T dp\cdot d(D(y)(v)) \\
&\qquad -\int_0^T pf_y'(t,y,u)D(y)(v)dt + \int_0^T p(g_y'(t,y))^2D(y)(v)dt \\
&\qquad -\int_0^T pg_y'(t,y)D(y)(v)dW_t\Big] + \int_0^T j'(u)vdt \\
&= E\Big[\int_0^T p\big[(f_y'(t,y,u)D(y)(v) + f_u'(t,y,u)v)dt \\
&\qquad +g_y'(t,y)D(y)(v)dW_t\big] - \int_0^T p(g_y'(t,y))^2D(y)(v)dt \\
&\qquad -\int_0^T pf_y'(t,y,u)D(y)(v)dt + \int_0^T p(g_y'(t,y))^2D(y)(v)dt \\
&\qquad -\int_0^T pg_y'(t,y)D(y)(v)dW_t\Big] + \int_0^T j'(u)vdt \\
&= \int_0^T E[p(f_u'(t,y,u)) + j'(u)]vdt.
\end{aligned} \tag{40}
$$

Remark 2.1. Our purpose of introducing p is to obtain a more convenient way to compute the gradient of the objective in applying the project gradient algorithm. There exist different ways to introduce an adjoint state aiming to obtain the SMP for stochastic optimal control, which has been a theme subject to intensively studies in the literature (see a summary in [51]). In the above derivation of $J'(u)(v)$ we have not assumed that p is a $\mathcal{F}_t$-adapted process, as it is well-known that equation (39) may not have a $\mathcal{F}_t$-adapted solution. However we did not utilize any martingale properties in the derivation, and the final formulation of $J'(u)(v)$ makes sense without the $\mathcal{F}_t$-adaptivity for p, see [6]. More importantly, our extensive numerical tests show that the algorithm described below is relatively simple and yet quite fast. More rigorous treatment is quite involved but definitely possible, and relevant technicalities can be found in, for example, [3], [23], [28], [35], [39], and [40]. On the other hand, if imposing $\mathcal{F}_t$-adaptivity for p, then we have to use a BSDE to redefine an adjoint process to simplify $J'(u)(v)$ as many researchers do (see, e.g. [35] and [42]). However a fast numerical algorithm for a BSDE is yet to be developed, particulary for the high dimensional case, see [35] and [49]. For the case where u appears in the diffusion term, it is still not clear if one has to use a BSDE to develop fast algorithms, and this is still under active investigation.

We then apply the above projection algorithm to the optimal control problem. Let $U = L^2(0,T)$ and $b(\cdot,\cdot) = (\cdot,\cdot)_U$. From the above computation we know:

$$J'(u)(v) = (E[p(f_u'(t,y,u))] + j'(u), v), \tag{41}$$

where y, p are the solutions of the state and co-state equations (35) and (39). In order to numerically compute the derivative, we use the Euler scheme for computing y, p, and then adopt the gradient method for numerically solving the optimal control problem. This yields the following algorithm (PPGA).

$$\left\{\begin{array}{l} b(u_{n+\frac{1}{2}},v) = b(u_n,v) - \rho_n(E[p_n(f'_u(t,y_n,u_n))] + j'(u_n),v), \\ \qquad\qquad u_{n+\frac{1}{2}}, u_n \in U, \quad \forall v \in U, \\ u_{n+1} = P_K^b\left(u_{n+\frac{1}{2}}\right), \end{array}\right. \tag{42}$$

where y_n, p_n, u_n represent the step functions constructed via (we have omitted the subscript h) the Euler scheme of the state and the co-state equations as follows:

(i) Choose the initial control u_0 arbitrarily.

For $n = 0, 1, \cdots$, let $u = u_n$, do the following iteration loops (ii)-(v).

(ii) Use Euler explicit scheme to compute the state equation:

$$\left\{\begin{array}{l} y^0 = y_0, \\ y^{m+1} = y^m + f(t^m, y^m, u^m)\Delta t_m + g(t^m, y^m)\Delta W_m, \\ \qquad m = 0, 1, \cdots, M-1. \end{array}\right. \tag{43}$$

where M is the total number of time steps. $\Delta t_m = t^{m+1} - t^m$ represents the mth time-step size, $\Delta W_m = W_{t^{m+1}} - W_{t^m}$ is the $N(0, \Delta t_m)$ increment of the Brownian motion W_t on $[t^m, t^{m+1}]$.

(iii) Use Euler implicit scheme to compute the adjoint equation:

$$\left\{\begin{array}{l} p^M = 0, \\ p^m = p^{m+1} + [h'(y^m) + p^m f'_y(t^m, y^m, u^m) \\ \qquad -p^m(g'_y(t^m, y^m))^2]\Delta t_m + p^m g'_y(t^m, y^m)\Delta W_m, \\ \qquad m = M-1, M-2, \cdots, 0. \end{array}\right. \tag{44}$$

The steps (ii) and (iii) are simulated for several orbits and then $E[p^m(f'_u(t^m, y^m, u^m))]$ can be obtained approximately for $m = 0, 1, \cdots, M$.

(iv) Use the gradient method to update the control:

$$\left\{\begin{array}{l} u^m_{n+\frac{1}{2}} = u^m - \rho_n\left(E[p^m(f'_u(t^m, y^m, u^m))] + j'(u^m)\right), \quad m = 0, 1, \cdots, M. \\ \qquad u^m_{n+1} = P_K^b\left(u_{n+\frac{1}{2}}\right). \end{array}\right. \tag{45}$$

(v) Compute $\epsilon_n = \|u_n - u_{n+1}\|_\infty$. If ϵ_n is small enough, exit. Otherwise, Let $u = u_{n+1}$, repeat the iteration loops (ii)-(v).

Next we use Theorem 1 and its corollary to discuss the convergence of the above algorithm by showing convergent property of J'_n to J', i.e., $\|J'(u_n) - J'_n(u_n)\| \leq \varepsilon_n \to 0$ if h is kept reduced towards to zero in iterations, or $\|J'(u_n) - J'_n(u_n)\| \leq \varepsilon$ if the size h is fixed during the iterations, assuming that the above Euler Schemes are convergent (see, e.g., [24]).

To this end, we first notice that

$$J'(u_n)(v) = (E[\hat{p}_n(f'_u(t, \hat{y}_n, u_n))] + j'(u_n), v), \tag{46}$$

where u_n is the control of the current iterative step, $\hat{y}_n$ and $\hat{p}_n$ are the corresponding solution of the following SDEs, respectively:

$$d\hat{y}_n = f(t, \hat{y}_n, u_n)dt + g(t, \hat{y}_n)dW_t, \quad \hat{y}_n(0) = y_0, \tag{47}$$

(48)

$$-d\hat{p}_n = [h'(\hat{y}_n) + \hat{p}_n f'_y(t, \hat{y}_n, u_n) - \hat{p}_n(g'_y(t, \hat{y}_n))^2]dt + \hat{p}_n g'_y(t, \hat{y}_n)dW_t, \quad \hat{p}_n(T) = 0.$$

On the other hand, we have

$$J'_n(u_n)(v) = (E[p_n(f'_u(t, y_n, u_n))] + j'(u_n), v), \tag{49}$$

where y_n and p_n are the approximate solutions of $\hat{y}_n$ and $\hat{p}_n$, respectively, obtained by using the Euler schemes we mentioned above. Thus,

$$
\begin{aligned}
& |J'(u_n)(v) - J'_n(u_n)(v)| \\
= \;& (E[\hat{p}_n(f'_u(t,\hat{y}_n,u_n))] - E[p_n(f'_u(t,y_n,u_n))], v) \\
= \;& (E[\hat{p}_n(f'_u(t,\hat{y}_n,u_n))] - E[\hat{p}_n(f'_u(t,y_n,u_n))], v) \\
& \quad + (E[\hat{p}_n(f'_u(t,y_n,u_n))] - E[p_n(f'_u(t,y_n,u_n))], v) \\
\le \;& (E[|\hat{p}_n| \cdot |\hat{y}_n - y_n|], |v|) + (E[|\hat{p}_n - p_n| \cdot |(f'_u(t,y_n,u_n))|], |v|) \\
\le \;& C(\|E[|\hat{y}_n - y_n|]\| + \|E[|\hat{p}_n - p_n|]\|) \cdot \|v\|,
\end{aligned} \tag{50}
$$

Let $\varepsilon_n = C(\|E[|\hat{y}_n - y_n|]\| + \|E[|\hat{p}_n - p_n|]\|)$, then $\|J'(u_n) - J'_n(u_n)\| \le \varepsilon_n$, and we know that $\lim\limits_{n\to\infty} \varepsilon_n = 0$ from the assumptions on convergence of the above Euler schemes.

If the size h is fixed, then we only can have $\|J'(u_n) - J'_n(u_n)\| \le \varepsilon$. Then we can only infer u_n is close to the optimal control u.

The above approach is applicable to the case where the control is a stochastic process. However the convergence analysis may be much involved as can be seen in [4]. Furthermore with the simulated sample data of optimal orbits and optimal control, it is possible to reconstruct the feedback laws via suitable stochastic regressions (see [16]), which is to be studied in the next stage.

3. Numerical Experiments

In this section, we present some numerical experiments to demonstrate our discretisation schemes and the methods developed in the above section.

Our first and second numerical examples are the Black-Scholes type of optimal control problems:

$$
\begin{aligned}
& \min_{u \in L^2(0,T)} J(u) = \frac{1}{2}\int_0^T E[(y - y^\star)^2]dt + \frac{1}{2}\int_0^T u^2 dt, \\
& \text{s.t.} \quad dy(t) = u(t)y(t)dt + \sigma y(t)dW_t, \quad y(0) = y_0,
\end{aligned} \tag{51}
$$

where σ is a constant.

According to the procedure in Section 2, the optimal control problem can be solved by

$$
\begin{cases}
dy = uydt + \sigma y dW_t, \quad y(0) = y_0, \\
-dp = (y - y^\star + up - p\sigma^2)dt + p\sigma dW_t, \quad p(T) = 0, \\
u = -E(py).
\end{cases} \tag{52}
$$

For comparison, we design two examples which have the exact expression of the optimal control as follows (see the Appendix). The first is

$$
u = \frac{T - t}{\dfrac{1}{y_0} - Tt + \dfrac{t^2}{2}}, \qquad y^\star = \frac{e^{\sigma^2 t} - (T-t)^2}{\dfrac{1}{y_0} - Tt + \dfrac{t^2}{2}} + 1. \tag{53}
$$

We choose $T = 1, y_0 = 1$ in the numerical computation.

Figures 1a, 1b show the results of the first numerical experiment with σ =0.0, 0.1, 0.3, 0.5, respectively. The time step is set $h = 0.02$, and 2000 orbits are used in the Monte Carlo simulation. The exit condition for the iteration is $\epsilon_n < 0.01$.

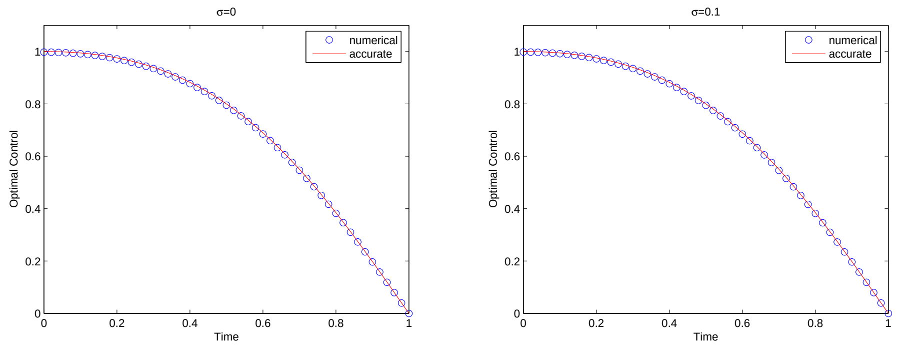

Figure 1a: Black-Scholes experiment I with $\sigma = 0.0, 0.1$

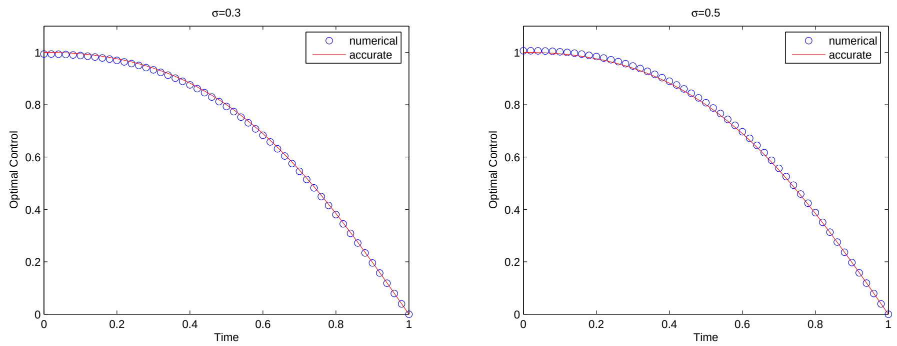

Figure 1b: Black-Scholes experiment I with $\sigma = 0.3, 0.5$

The second example is

$$u = \frac{e^{-T} - e^{-t}}{\dfrac{1}{y_0} + 1 - e^{-t} - e^{-T}t}, \qquad y^{\star} = \frac{e^{\sigma^2 t} - (e^{-T} - e^{-t})^2}{\dfrac{1}{y_0} + 1 - e^{-t} - e^{-T}t} - e^{-t}, \tag{54}$$

we also choose $T = 1, y_0 = 1$ in the numerical computation.

Figures 2a, 2b show the results of the second numerical experiment with σ =0.2, 0.5, 0.7, 1.0, respectively. The same numerical settings for the Euler Scheme and Monte Carlo simulation are used.

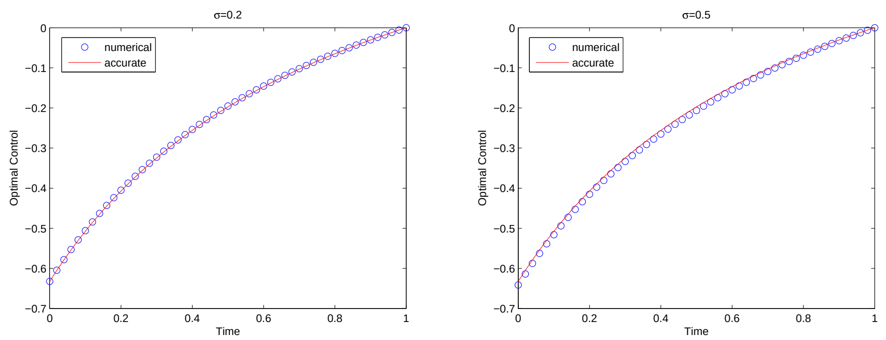

Figure 2a: Black-Scholes experiment II with $\sigma = 0.2, 0.5$

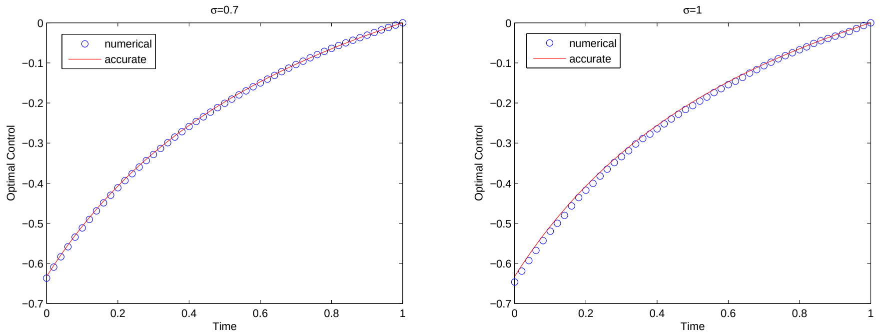

Figure 2b: Black-Scholes experiment II with $\sigma = 0.7, 1.0$

The third example is the inventory control problem with deterministic control (see [52]): Let $y(t)$ be the inventory level at time t for a goods, and let $u(t)$ and $r(t)$ be the production and market demand rates of this goods. But the inventory can be easily damaged, the amount of the damage is a stochastic process: Let σ be the damage rate of the inventory, then

$$dy = (u - r)dt + \sigma dW_t \tag{55}$$

The objective is to minimize the total production and storage cost:

$$\min_{u \geq 0} J = 0.5[c_1 \int_0^T E[(y - y^\star)^2] + c_2 \int_0^T u^2]. \tag{56}$$

Based on the procedure in Section 2, we have

$$\begin{cases} dy = (u - r)dt + \sigma dW_t, \quad y(0) = y_0, \\ -dp = (y - y^\star)dt, \quad p(T) = 0, \\ u = \max\big(0, -\dfrac{c_1}{c_2}E(p)\big). \end{cases} \tag{57}$$

We design the following example with exact control (see the Appendix):

$$c_1 = c_2 = 1,\ y_0 = 0,\ y^\star = 0.5Tt - 0.25t^2 + 1,\ r = 0.5(T - t),\ u = T - t. \tag{58}$$

Figures 3a, 3b show the results of the third numerical experiment with σ =0.0, 1.0, 3.0, 5.0, respectively. The same numerical settings for the Euler Scheme and Monte Carlo simulation are used and, the exit condition for the iteration is $\epsilon_n < 0.005$.

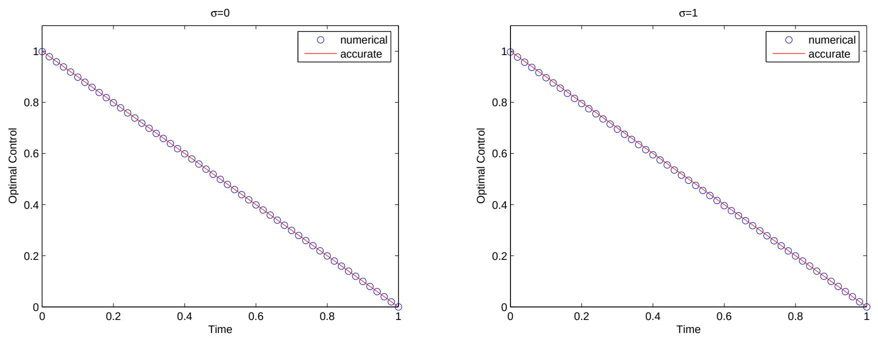

Figure 3a: Inventory experiment with $\sigma = 0.0, 1.0$

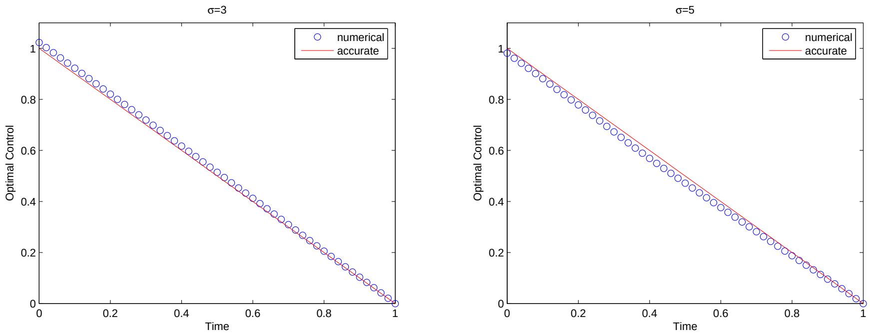

Figure 3b: Inventory experiment with $\sigma = 3.0, 5.0$

The fourth example is the following type of optimal control problem:

$$
\begin{aligned}
&\min_{u\in L^2(0,T)} J(u) = \frac{1}{2}\int_0^T E[(y-y^\star)^2]dt + \frac{1}{2}\int_0^T (u-u^\star)^2 dt, \\
&\text{s.t.} \quad dy(t) = \frac{1}{2}u(t)(u(t)-u^\star(t))y(t)dt + \sigma y(t)dW_t, \quad y(0) = y_0,
\end{aligned} \tag{59}
$$

where σ is a constant.

The optimal control problem can be solved by

$$
\begin{cases}
dy = \frac{1}{2}u(u-u^\star)ydt + \sigma y dW_t, \quad y(0) = y_0, \\
-dp = (y - y^\star - p\sigma^2)dt + p\sigma dW_t, \quad p(T) = 0, \\
u - u^\star + E[p(u - \frac{u^\star}{2})] = 0.
\end{cases} \tag{60}
$$

It is clear that we have the exact solutions:.

$$
u = u^\star = 6\sin\pi t, \qquad y = y^\star = y_0 e^{-\frac{\sigma^2}{2}t + \sigma W_t}. \tag{61}
$$

We choose $T = 1, y_0 = 1$ in the numerical computation.

Figures 4 show the results of the fourth numerical experiment with σ =0.3, 0.5, respectively. The same numerical settings for the Euler Scheme and Monte Carlo simulation are used and, the exit condition for the iteration is $\epsilon_n < 0.05$.

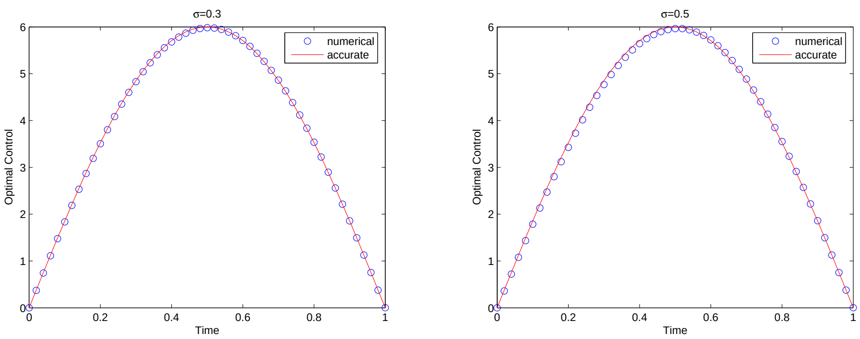

Figure 4: The fourth experiment with $\sigma = 0.3, 0.5$

The fifth example is the following type of optimal control problem:

$$
\begin{aligned}
&\min_{u\in L^2(0,T)} J(u) = \frac{1}{2}\int_0^T E[(y-1)^2]dt + \frac{1}{2}\int_0^T u^2 dt, \\
&\text{s.t.} \quad dy(t) = u(t)y(t)dt + \sigma\sqrt{1+y(t)^2}dW_t, \quad y(0) = y_0,
\end{aligned} \tag{62}
$$

where σ is a constant.

The optimal control problem can be solved by

$$\begin{cases} dy = uydt + \sigma\sqrt{1+y^2}dW_t, \quad y(0) = y_0, \\ -dp = \big(y - 1 + pu - p\dfrac{\sigma^2 y^2}{1+y^2}\big)dt + p\dfrac{\sigma y}{\sqrt{1+y^2}}dW_t, \quad p(T) = 0, \\ u = -E(py). \end{cases} \tag{63}$$

However this time we do not know the exact solution. In the numerical tests below, we choose $T = 1, y_0 = 1$ in the computation.

Figures 5 show the results of the fifth numerical experiment with σ =0.5, 0.7, respectively. The same numerical settings for the Euler Scheme and Monte Carlo simulation are used and, the exit condition for the iteration is $\epsilon_n < 0.005$.

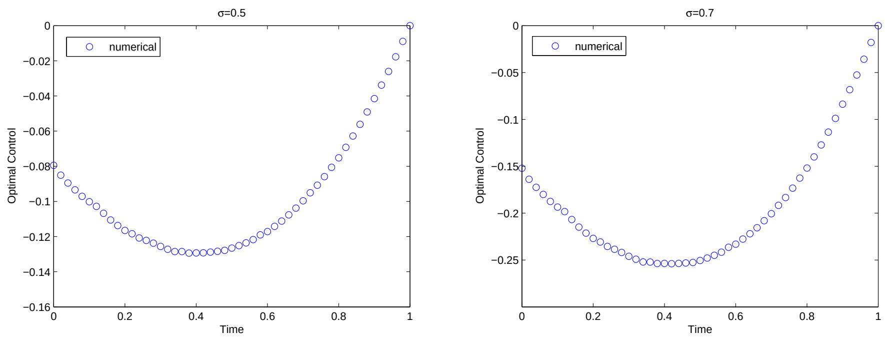

Figure 5: The fifth experiment with $\sigma = 0.5, 0.7$

Since we do not know the exact solution this time, we carry out several numerical simulations of $J(u)$ for some selected control $u(t)$, and compare them with the optimal numerical value $J(u^*)$. First, we choose the constant control $u(t) = c$, where c ranges from -1 to 1 with the step 0.01. The corresponding values of $J(c)$ are shown in Figure 6 for $\sigma = 0.7$. The minimum value of $J(c) = 0.2751$, is reached at $c = -0.2$, which is close to (but still larger than) the value of $J(u^*)$. The corresponding u^* is shown in the right figure of Figure 5.

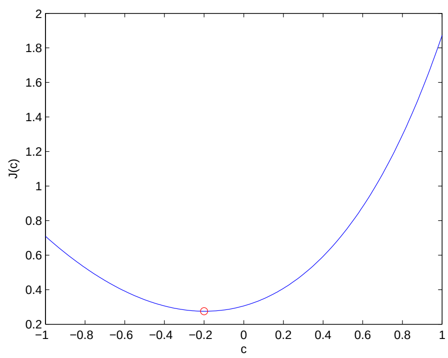

Figure 6: Illustration of $J(c)$ as a function of c for $\sigma = 0.7$

We also tested other randomly selected controls, e.g., $u = t^2 - 1$, $u = \sin(t) - 1$, etc. All these simulations produce much larger values of functional $J(u)$ than $J(u^*)$.

4. Conclusions

In this work we propose a numerical algorithm for a class of stochastic optimal control, where the control is deterministic. This algorithm is relatively simple for implementation, but quite effective as shown in our extensive numerical tests. It is applicable for the high dimensional case where there are many risky assets to be managed. The purpose of selecting this simpler control class in this paper is that we wish to separate the study of numerically computing the optimal control, and that of constructing feedback laws, which is to be studied in the next stage.

5. Appendix

In this section, we describe the details of designing the numerical experiments in Section 3. The exact solutions are based on a different optimal condition, which can be used to solve the exact optimal control.

5.1. The Black-Scholes cases. We first examine the Black-Scholes control problem with deterministic control $u(t) \in L^2(0,T)$:

$$J(u) = \frac{1}{2}\int_0^T E[(y-y^\star)^2]dt + \frac{1}{2}\int_0^T u^2 dt, \tag{64}$$

where $y^\star$ is deterministic. We need to determine $\min\limits_{u \in L^2(0,T)} J(u)$ with

$$dy(t) = u(t)y(t)dt + \sigma y(t)dW_t, \quad y(0) = y_0, \tag{65}$$

where σ is a constant.

In the follows, we derive a different optimality condition, which can be used to design the close form optimal control. It follows from Itô formula that

$$y(t) = y_0 e^{\int_0^t u(s)ds - \frac{\sigma^2}{2}t + \sigma W_t}, \tag{66}$$

and then

$$E(y) = y_0 e^{\int_0^t u(s)ds - \frac{\sigma^2}{2}t} E(e^{\sigma W_t}) = y_0 e^{\int_0^t u(s)ds}, \tag{67}$$

so we have

$$dE(y) = u(t)E(y)dt \tag{68}$$

and

$$E[(y-y^\star)^2] = [E(y)]^2 e^{\sigma^2 t} - 2y^\star E(y) + [y^\star]^2. \tag{69}$$

Noting that

$$J'(u)(v) = \int_0^T [E(y)e^{\sigma^2 t} - y^\star][E(y)'(u)(v)]dt + \int_0^T uvdt, \tag{70}$$

if we define p satisfies

$$-dp(t) = (u(t)p(t) + E(y)e^{\sigma^2 t} - y^\star)dt, \quad p(T) = 0, \tag{71}$$

then

$$\begin{aligned} J'(u)(v) &= \int_0^T \big(-\frac{dp}{dt} - up\big)[E(y)'(u)(v)]dt + \int_0^T uvdt \\ &= \int_0^T p\frac{d[E(y)'(u)(v)]}{dt}dt - \int_0^T up[E(y)'(u)(v)]dt + \int_0^T uvdt \\ &= \int_0^T (u + pE(y))vdt = 0, \quad \forall v \in L^2(0,T), \end{aligned} \tag{72}$$

which means

$$u = -pE(y). \tag{73}$$

Based on the above deduction, we construct the following two examples:

I. Let $p = t - T$, from

$$dE(y) = u(t)E(y)dt = -p(E(y))^2, \quad E(y)(0) = y_0 \tag{74}$$

we have

$$E(y) = \frac{1}{\frac{1}{y_0} - Tt + \frac{t^2}{2}}. \tag{75}$$

Then the optimal control u can be derived as

$$u = \frac{T - t}{\frac{1}{y_0} - Tt + \frac{t^2}{2}}. \tag{76}$$

Finally, $y^\star$ can be obtained as follows:

$$y^\star = \frac{dp}{dt} + u(t)p(t) + E(y)e^{\sigma^2 t} = \frac{e^{\sigma^2 t} - (T-t)^2}{\frac{1}{y_0} - Tt + \frac{t^2}{2}} + 1. \tag{77}$$

II. Similarly, let $p = e^{-t} - e^{-T}$, we have

$$E(y) = \frac{1}{\frac{1}{y_0} + 1 - e^{-t} - e^{-T}t}. \tag{78}$$

Then

$$u = \frac{e^{-T} - e^{-t}}{\frac{1}{y_0} + 1 - e^{-t} - e^{-T}t} \tag{79}$$

and

$$y^\star = \frac{e^{\sigma^2 t} - (e^{-T} - e^{-t})^2}{\frac{1}{y_0} + 1 - e^{-t} - e^{-T}t} - e^{-t}. \tag{80}$$

5.2. The Inventory case. The Inventory control problem is to minimize the total production and storage cost

$$\min_{u \geq 0} J = 0.5[c_1 \int_0^T E[(y - y^\star)^2] + c_2 \int_0^T u^2]. \tag{81}$$

s.t.

$$dy = (u - r)dt + \sigma dW_t, \quad y(0) = y_0. \tag{82}$$

We will derive a similar optimality condition to design the numerical example. It follows that

$$y = y_0 + \int_0^t (u - r)dt + \sigma W_t = E(y) + \sigma W_t, \tag{83}$$

so that

$$E(y) = y_0 + \int_0^t (u(s) - r(s))ds \tag{84}$$

and

$$dE(y) = (u - r)dt. \tag{85}$$

Noting that

$$E[(y - y^\star)^2] = E[(E(y) - y^\star + \sigma W_t)^2] = [E(y) - y^\star]^2 + \sigma^2 t, \tag{86}$$

if we define

$$-dp = [E(y) - y^\star]dt, \quad p(T) = 0, \tag{87}$$

then

$$\begin{aligned} J'(u)(v) &= c_1 \int_0^T \big(-\frac{dp}{dt}\big)[E(y)'(u)(v)]dt + c_2 \int_0^T uvdt \\ &= c_1 \int_0^T p \frac{d[E(y)'(u)(v)]}{dt} dt + c_2 \int_0^T uvdt \\ &= \int_0^T (c_1 p + c_2 u)vdt. \end{aligned} \tag{88}$$

The optimal condition reads

$$\int_0^T (c_1 p + c_2 u)(v - u)dt \geq 0, \quad \forall v \geq 0, \tag{89}$$

which means

$$u(t) = \max(0, -\frac{c_1}{c_2}p). \tag{90}$$

Now we design the corresponding numerical experiment. Let $c_1 = c_2 = 1, \quad y_0 = 0$, and $p = t - T$, we first have

$$u(t) = \max(0, -\frac{c_1}{c_2}p) = T - t. \tag{91}$$

Let $r(t) = 0.5(T - t)$, then

$$E(y) = \int_0^T (u - r)dt = 0.5Tt - 0.25t^2, \tag{92}$$

and finally

$$y^\star = \frac{dp}{dt} + E(y) = 0.5Tt - 0.25t^2 + 1. \tag{93}$$

References

[1] O. Bahn, A. Haurie, and R. Malhame, *A stochastic control model for optimal timing of climate policies*, Automatica, 44(6)(2008), 1545-1558.

[2] M. Bardi and I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*, Birkhäuser, Boston, 1997.

[3] M. Bismut, Mecinique Aleatoire, *Lecture Notes in Mathematics No.866*, Springer, 1981.

[4] K. Barty, J.-S. Roy, and C. Strugarek, *A stochastic gradient type algorithm for closed-loop problems*, Math. Program., 119(1)(2009), 51-78.

[5] A. Bensoussan, *Lecture on Stochastic Control*, in Nonlinear Filtering and Stochastic Control, Lecture Notes in Mathematics 972, Proc. Cortona, Springer-Verlag, Berlin, New York, 1981.

[6] A. Bensoussan, *Stochastic Control by Functional Analysis Methods*, North-Holland Publishing Company, New York, 1982.

[7] D.P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. II. Athena Scientific, 2007.

[8] F. Black and M. Scholes, *The pricing of options and corporate Liabilities*, Journal of Political Economy, 81(1973), 637-654.

[9] J.F. Bonnans, J.-C. Gilbert, C. Lemaréchal, and C.A. Sagastizábal, *Numerical Optimization: Theoretical and Practical Aspects*, 2nd ed., Springer, 2006.

[10] 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, 1990.

[11] P. Carpentier, G. Cohen, and A. Dallagi, *Particle methods for stochastic optimal control problems*, arXiv:0907.4663v1, 2009.

[12] J. Chen, Some Problems of Optimal Control and the Application in Financial Mathematics, Master's Thesis, Tongji University, China, 2007.

[13] M. Crandall and P. L. Lions, *Viscosity solutions of Hamilton-Jacobi equations*, Trans. Amer. Math. Soc., 277 (1983), 1-42.

[14] A. Dallagi, *Méthodes particulaires en commande optimale stochastique*, Thèse de doctorat, Université Paris 1, Panthéon-Sorbonne, 2007.

[15] R.J. Elliott, *The optimal control of diffusions*, Appl. Math. Opt., 22(1990), 229-240.

[16] P. Girardeau, *A comparison of sample-based stochastic optimal control methods*, arXiv:1002.1812, 2010.

[17] F.B. Hanson, *Applied stochastic processes and control for Jump-diffusions: modeling, analysis, and computation.* SIAM Books: Advances in Design and Control Series, 2007.

[18] U.G. Haussmann, *Some examples of optimal stochastic controls or: the stochastic maximum principle at work*, SIAM Rev., 23(3)(1981), 292-307.

[19] H. Heitsch, and W. Romisch, *Scenario reduction algorithms in stochastic programming*, Comput. Optim. Appl., 24(2003), 187-206.

[20] H. Heitsch, W. Romisch, and C. Strugarek, *Stability of multistage stochastic programs.* SIAM J. Optim., 17(2006), 511-525.

[21] C.S. Huang, S. Wang, and K.L. Teo, *Solving Hamilton-Jacobi-Bellman equations by a modified method of characteristics*, Nonlinear Anal-Theor., 40(2000), 279-293.

[22] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North-Holland, Kodansha, 1989.

[23] K. Ito, *Extension of stochastic integral*, Proceedings of Intern. Symp. on SDE, J. Wiley, 1978.

[24] P.E. Kloeden, and E. Platen, *Numerical Solution Of Stochastic Differential Equations*, Springer, Berlin, 1995.

[25] R. Korn, *Optimal Portfolios: Stochastic Models for Optimal Investment and Risk Management in Continuous Time*, World ScientificSingapore, 1997.

[26] R. Korn, and H. Kraft, *A stochastic control approach to portfolio problems with stochastic interest rates*, SIAM J. Control Optim., 40(4)(2001), 1250-1269.

[27] D. Kuhn, P. Parpas, and B. Rustem, *Stochastic optimization of investment planning problems in the electric power industry*, in Process Systems Engineering Volume 5: Energy Systems Engineering, 2008.

[28] H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge Press, 1997.

[29] H.J. Kushner, *Necessary conditions for continuous parameter stochastic optimization problems*, SIAM J. Control, 10(1972), 550-565.

[30] H.J. Kushner and P. Dupuis, *Numerical methods for stochastic control problems in continuous time*, 2nd ed., Springer-Verlag, New York, 2001.

[31] 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(2)(2008), 419-430.

[32] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, Berlin, 1971.

[33] W.B. Liu and N.N. Yan, *Adaptive Finite Element Methods for Optimal Control Governed by PDEs*, Science Press, Beijing, 2008.

[34] A. Ludwig, *Random Dynamical Systems*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 1998.

[35] J. Ma and Jiongmin Yong, *Forward-Backward Stochastic Differential Equations and their Applications*, Springer, 2007.

[36] J. Ma, P. Protter, J. San Martin, and S. Torres, *Numerical method for backward stochastic differential equations*, Ann. Appl. Probab., 12(1)(2002), 302-316.

[37] R.C. Merton, *Continuous-Time Finance*, Blackwell, Cambridge, Massachusetts, 1992.

[38] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, 6th ed., Springer, 2003.

[39] F.Russo and P. Vallois, Forward, *Backward and symmetric stochastic integration*, Probab. Theory Rel., 97(3)(1993), 403-421.

[40] E. Pardoux and P. Protter, *Two side stochastic integral with its caluclus*, Technical Report No.86-3, Purdue University, 1986.

[41] S.G. Peng, *A general stochastic maximum principle for optimal control problems*, SIAM J. Control Optim., 28(4)(1990), 966-979.
[42] S.G. Peng, *Backward stochastic differential equations and applications to optimal control*, Appl. Math. Optim., 27(1993), 125-144.
[43] 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 2005, LNCS 3414 (Eds. M. Morari, L. Thiele), p.557-572, Springer-Verlag, 2005.
[44] L.C.G. Rogers, *Pathwise Stochastic Optimal Control*, SIAM J. Control Optim., 46(3)(2007), 1116-1132.
[45] S.P. Sethi and Q. Zhang, *Hierarchical decision making in stochastic manufacturing systems*, Birkhauser Verlag, Basel, Switzerland, 1994.
[46] A. Shapiro and A. Ruszczynski, *Stochastic Programming.* Elsevier, Amsterdam, 2003.
[47] J.T. Shi and Z. Wu, *A stochastic maximum principle for optimal control of jump diffusions and applications to finance*, Chinese Journal of Applied Probability and Statistics, 27(2011), 127-137.
[48] J.C. Spall, *Stochastic Optimization II.6*, in Handbook of Computational Statistics (Eds. J. Gentle, W. Härdle, and Y. Mori), p.169-197, Springer-Verlag, New York, 2004.
[49] J.L, Wang, *Numerical Methods and Their Error Estimates for Backward Stochastic Differential Equations*, Ph.D thesis, Shandong Universithy, 2009.
[50] W.S. Xu, *Maximum principle for a stochastic optimal control problem and application to portfolio/consumption choice*, J. Optimiz. Theory App., 98(3)(1998), 719-731.
[51] T. Yoneyama, *Adjoint processes in stochastic optimal control*, Ph.D Thesis, Imperical College, London, 1983.
[52] Z.Q Zhao, *Distributed optimal control problems in quantatitive economics and management*, PhD. Thesis, Northeastern University, China, 2000.

School of Mathematics, Shandong University, Jinan 250100, China
E-mail: duning@sdu.edu.cn and shijingtao@sdu.edu.cn

KBS, University of Kent, CT2 7PE, UK
E-mail: W.B.Liu@kent.ac.uk