Numer. Math. Theor. Meth. Appl. Vol. **13**, No. 2, pp. 296-319
doi: 10.4208/nmtma.OA-2019-0137 May 2020

Highly Accurate Numerical Schemes for Stochastic Optimal Control Via FBSDEs

Yu Fu[1], Weidong Zhao[2,*] and Tao Zhou[3]

[1] *College of Mathematics and Systems Science, Shandong University of Science and Technology, Qingdao 266590, China, and School of Mathematics & Institute of Finance, Shandong University, Jinan 250100, Shandong, China*
[2] *School of Mathematics & Institute of Finance, Shandong University, Jinan 250100, Shandong, China*
[3] *CMIS & LSEC, Institute of Computational Mathematics and Scientific/ Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China*

Received 1 September 2019; Accepted (in revised version) 18 November 2019

Abstract. This work is concerned with numerical schemes for stochastic optimal control problems (SOCPs) by means of forward backward stochastic differential equations (FBSDEs). We first convert the stochastic optimal control problem into an equivalent stochastic optimality system of FBSDEs. Then we design an efficient second order FBSDE solver and an quasi-Newton type optimization solver for the resulting system. It is noticed that our approach admits the second order rate of convergence even when the state equation is approximated by the Euler scheme. Several numerical examples are presented to illustrate the effectiveness and the accuracy of the proposed numerical schemes.

AMS subject classifications: 60H35, 93E20, 93E25, 49M29, 65C20, 65K15

Key words: Forward backward stochastic differential equations, stochastic optimal control, stochastic maximum principle, projected quasi-Newton methods.

1. Introduction

On a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with the natural filtration $\mathcal{F}_t = \sigma\{W_s; 0 \leq s \leq t\}$, we consider a controlled diffusion process X_t (also known as the controlled state), governed by the stochastic differential equation (SDE)

$$\begin{cases} \mathrm{d}X_t = b(t, X_t, \alpha_t)\mathrm{d}t + \sigma(t, X_t, \alpha_t)\mathrm{d}W_t, \qquad t \in (0, T], \\ X_0 = x_0, \end{cases} \tag{1.1}$$

*Corresponding author. *Email addresses:* `nielf_fu@sdust.edu.cn` (Y. Fu), `wdzhao@sdu.edu.cn` (W. D. Zhao), `tzhou@lsec.cc.ac.cn` (T. Zhou)

where $W_t = (W_{1,t}, \cdots, W_{m,t})^\top$ is a standard m-dimensional Brownian motion, α_t is the control process, X_t is called the trajectory with respect to α_t, and

$$b:\ [0,T] \times \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}^d \qquad \text{and} \qquad \sigma:\ [0,T] \times \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}^{d\times m}$$

are deterministic functions that are referred as the drift and diffusion coefficients, respectively.

The admissible control set $\mathscr{U}$ for the α_t is defined as

$$\mathscr{U} = \{\alpha_{\cdot} \in \mathscr{M}^2(\mathbb{R}^k) \big| \alpha_t \in \mathrm{U},\ t \in [0,T],\ \ \text{a.e., a.s.}\},$$

where $\mathscr{M}^2(\mathbb{R}^k)$ denote the space of all the $\mathcal{F}_t$-adapted processes valued in $\mathbb{R}^k$:

$$\mathbb{E}\left[\int_0^T |X_t|^2 \,\mathrm{d}t\right] < \infty.$$

Here U is a nonempty, convex and closed subset of $\mathbb{R}^k$. For $\alpha_{\cdot} \in \mathscr{U}$, it is well known that (1.1) admits a unique solution X_t under standard conditions on b and σ (e.g., Liptchiz continuous).

The cost functional $J(\alpha_{\cdot}) : \mathscr{U} \to \mathbb{R}$ under consideration is

$$J(\alpha_{\cdot}) = \mathbb{E}\left[\int_0^T f(t, X_t, \alpha_t)\mathrm{d}t + \varphi(X_T)\right], \tag{1.2}$$

where $f(t,x,\alpha) : [0,T] \times \mathbb{R}^d \times \mathbb{R}^k \to \mathbb{R}$ and $\varphi(x) : \mathbb{R}^d \to \mathbb{R}$ are deterministic functions that are called the running cost and the terminal cost, respectively.

The SOCP is to find $\alpha_{\cdot}^* \in \mathscr{U}$ such that $J(\alpha_{\cdot}^*)$ attains its minimum over the admissible control set, i.e.,

$$J(\alpha_{\cdot}^*) = \min_{\alpha_{\cdot} \in \mathscr{U}} J(\alpha_{\cdot}), \qquad \alpha_{\cdot}^* = \arg\min_{\alpha_{\cdot} \in \mathscr{U}} J(\alpha_{\cdot}), \tag{1.3}$$

where $\alpha_{\cdot}^*$ and $J(\alpha_{\cdot}^*)$ are called the optimal control and the optimal value, respectively.

In the theoretical point of view, one of the most popular approaches for studying the above SOCPs is the Pontryagin's maximum principle [5, 22, 24, 33]. For stochastic problems, essential difficulties arise when the diffusion coefficients depend on α_t. This was overcome by Peng [26] by considering the second-order term in the Taylor expansion of the spike variation, yielding a more general version of stochastic maximum principle. Another popular approach is the Bellman's dynamic programming that involves the Hamilton-Jacob-Bellman (HJB) equations [2–4, 6, 10, 18]. Based the backward stochastic differential equations (BSDEs) theory [25, 28], the generalized dynamic programming principle and HJB equations have also been developed in [29]. Moreover, the nonlinear Feynman-Kac formula [27] established a relationship between SOCPs and PDE constrained optimal control problems [8, 13, 17, 21].

Numerical methods for SOCPs have drawn more and more attention in recent years. While most attempts have been made by solving the associated HJB equations (e.g., [1, 16, 19, 20, 36, 37, 39, 40]), little attention has been paid to develop direct stochastic