2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

A Finite Step Scheme for General Near-Optimal Stochastic Control

Danchi Jiang

Abstract— **This paper reports a finite step scheme for the computation of near-optimal control of general stochastic systems with their diffusion terms affected by control. It is developed based on second order estimation of the stochastic system and the associated two adjoint variational equations. The search is an extension of the standard steepest descent method to the functional case with a random step size based on the variation of an auxiliary function $\mathcal{H}$. Convergence analysis are included to show this scheme does converge to a desired admissible control in finite step. Consistency of the approximation of the associated adjoint equations is also discussed. A linear quadratic control example is included for illustration purpose.**

I. Introduction

At present, there are roughly two major types of computational approaches that have comparatively solid theoretical basis. One is based on the Markov chain approximation method. To this method, the stochastic control system is first approximated using a Markov chain evolving in a finite state set, then an optimal control problem is also approximated on this Markov chain and finite state set. At last, an optimal control is computed or estimated for the Markov chain optimal control problem. In [1], [2], this techniques is discussed in detail. In [2], [3], it is clarified that the Markov chain approximation schema proposed is also suitable to the type of stochastic optimal control problems where control affects the diffusion term. Another approach is based on the numerical solution of Hamilton-Jacobi-Bellman (HJB) equation, which is obtained using dynamic programming technique. Theoretically, this HJB equation gives a partial differential equation that the value function must satisfy. For surveys of existing approaches, see [4], [5].

One of the key issues here is how to measure whether an admissible control is close to the optimal one. A mathematical way could be using certain type of metric in the space of all admissible controls. However, in a practical settings the "near-optimal" concept proposed in [6] makes more sense. This is not only because usually a smooth optimal control may not exist, but also because, by near-optimal the solution is measured by how close to the optimal value the cost function value is, which fits perfectly the objective to optimize the cost function.

Based on the concept of near-optimality, we develop a scheme for the computation of general optimal stochastic control problem. By this scheme, the cost function is guaranteed to minimized by at least a fixed amount until control is near-optimal in the sense that the cost function is close to the optimal one. A loose bound is also given for the number of iteration steps to achieve a near-optimal solution.

Danchi Jiang is with the School of Engineering, University of Tasmania.
This work is supported by National ICT Australia, which is funded through the Australian Government's *_Backing_Australia's_Ability_initiative*, in part through the Australian Research Council.

The scheme is developed using a point-wise gradient direction of an auxiliary cost function $\mathcal{H}$. This auxiliary function is computed using the solutions to two adjoint backward stochastic differential equations (BSDEs) that reflect the first-order and second-order approximations of the original systems. With many researches have been conducted towards to numerical solution of BSDEs, our research is not emphasized on this aspects. Interested reader can see [7], [8], [9], [10], [11], [12] for discussions of BSDEs and their numerical computation methods.
bbb

II. Problem Statements and Preliminaries

Let us first introduce notations that will be used in this paper:

$(\Omega, \mathcal{F}, \mathcal{P}, B)$	:	a four-tuple defined by the sample space $\Omega \subseteq \Re^n$, the σ-field $\mathcal{F}$, the probability measure $\mathcal{P}$, and the l-dimensional independent Brownian motion $B := B(t), t_0 \leq t \leq T$ with $B(t_0) = 0$.
$\mathcal{F}_t^{t_0}$	:	the natural filtration generated by $B(t)$ augmented by all P-null sets in $\mathcal{F}$.
$u(\cdot)$	:	the m-dimensional control input. It is a $\mathcal{F}_t^{t_0}$-adapted measurable random process taken value in a given compact set $\Gamma \subset \Re^m$. Let us also denote the bound of the compact set Γ as Γ_b.
$U_{ad}[t_0, T]$	:	the set of admissible controls. For given $(\Omega, \mathcal{F}, \mathcal{P}, B)$, each element of $U_{ad}[t_0, T]$ is a control process $u(\cdot)$.
$M^\top$	:	the transpose of matrix or vector M.
$\mid a \mid$	:	the norm of a vector or a matrix a. It is the sum of absolution value of its components.
$\frac{\partial \rho}{\partial z}$	:	the partial derivative of a vector function ρ with respect to a vector variable z. More specifically, $\left(\frac{\partial \rho}{\partial z}\right)_{i,j} = \frac{\partial \rho_i}{\partial z_j}$.
$\frac{\partial^2 \rho}{\partial z^\top \partial z}$	:	the second order partial derivative of a scaler function ρ with respect to a vector variable z. More specifically, $\left(\frac{\partial^2 \rho}{\partial z^\top \partial z}\right)_{i,j} = \frac{\partial^2 \rho}{\partial z_i \partial z_j}$.

a_i	:	the i-th row vector or a matrix a.
$a_{\cdot i}$	:	the i-th column vector or a matrix a.
(a_{ij})	:	A matrix that its (i,j)-th element is a_{ij}.

Consider the following stochastic system with control:

$$\begin{aligned} dx(t) &= f(t,x(t),u(t))dt + \sigma(t,x(t),u(t))dB(t), \\ x(t_0) &= x_0, \end{aligned} \quad (1)$$

where f, σ are measurable in (t,x,u). The objective of optimal control is to find the admissible control to minimize the cost function given by:

$$J(u(\cdot)) = E\{\int_{t_0}^{T} L(t,x(t),u(t))dt + h(x(T))\}, \quad (2)$$

where L is a measurable function in (t,x,u) and h is a function of $x(T)$. If the time interval and the initial state value of considering can change, J is also a function of t_0, x_0 and T. We also let the optimal cost, usually called value function, be denoted as V or $V(t_0,,x_0,T)$.

We say an admissible control u^ϵ is near optimal if the the value of the corresponding cost is near the value of V. More specifically, for given $\epsilon > 0$, u^ϵ is ϵ-optimal if $\mid J(u(\cdot)) - V \mid \leq \epsilon$. In some cases it is not easy to justify the real meaning of near-optimal for one single fixed ϵ. We consider a positive sequence converges to zero, say, $\mathcal{E} := \{\epsilon_n\} \to 0+$. A sequence of admissible controls $\{u_n(\cdot)\}$ is called $\mathcal{E}$-optimal if $\mid J(u_n) - V \mid \leq \epsilon_n$.

In this study, we need the following assumption:

Assumption 1: We assume:

A1. f, σ, and L are measurable in (t,x,u) and twice continuously differentiable in x and u.

A2. h is twice continuously differentiable.

A3. There is a constant C such that the following Lipschitz type conditions are satisfied:

$$\begin{aligned} &\mid \rho(t,x,u) \mid \leq C(1+ \mid x \mid), \\ \mid \rho(t,x,u) - \rho(t,x',u') \mid + \mid &\frac{\partial \rho}{\partial x}(t,x,u) - \frac{\partial \rho}{\partial x}(t,x',u') \mid \\ &+ \mid \frac{\partial \rho}{\partial u}(t,x,u) - \frac{\partial \rho}{\partial u}(t,x',u') \mid \\ &\leq C(\mid x - x' \mid + \mid u - u' \mid), \\ &\mid h(x) \mid \leq C(1+ \mid x \mid), \\ \mid h(x) - h(x') \mid + \mid &\frac{\partial h}{\partial x}(x) - \frac{\partial h}{\partial x}(x') \mid \leq C \mid x - x' \mid, \end{aligned}$$

where $\rho = f, \sigma, L$.

In the proof of the proposed scheme, we also need the following result borrowed from [13]:

Lemma 1: Ekeland's Principle Let (S,d) be a complete metric space and $\rho(\cdot) : S \to \Re^1$ be lower-semicontinuous and bounded below. For $\epsilon \geq 0$, if $u^\epsilon \in S$ satisfies

$$\rho(u^\epsilon) \leq \inf_{u \in S} \rho(u) + \epsilon.$$

Then, for any $\lambda > 0$, there exists a $u^\lambda \in S$ such that $\rho(u^\lambda) \leq \rho(u^\epsilon)$, $d(u^\lambda, u^\epsilon) \leq \lambda$, $\rho(u^\lambda) \leq \rho(u) + \frac{\epsilon}{\lambda} d(u^\lambda, u^\epsilon)$ where $d(u^\lambda, u^\epsilon)$ is the distance in the metric space.

We also need a result for backward stochastic differential equations:

Lemma 2: Consider a backward stochastic differential equation (BSDE) in the following form:

$$x(t) + \int_t^T \xi(s,x(s),y(s),\gamma(s))ds + \int_t^T y(s)dB(s) = X, \quad (3)$$

where $B(t)$ is a Brownian motion or Wiener process defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, $\mathcal{F}_t$ is its natural filtration, X is a $\mathcal{F}_T$ measurable random vector, x, Y are n and $n \times m$ dimensional random vector and matrix variables, respecticely, ξ is a $\mathcal{P} \otimes \mathcal{B}_d \otimes \mathcal{B}_{n\times m} \otimes \mathcal{B}_l / \mathcal{B}_d$ measurable function, $\mathcal{P}$ is the σ-algebra of $\mathcal{F}_t$-progressively measurable subsets of $\Omega \times [t_0, T], and$ $\gamma(s)$ is a l-dimesional random vector process that represents uncertainty of the model. Assume that ξ is uniformly Lipschitz in the variables x and y with the Lipschitz constant C_b independent of γ. Then, the solution pair (x,y) uniquely exists for each given $\gamma(\cdot)$ and the solution is uniformly bounded. More specifically, there exists a constant C_5 such that

$$E \int_{t_0}^{T} \mid x(t) \mid^2 + \mid y(t) \mid^2 dt < C_5.$$

The uniqueness and existence are existing results. For more details, see [8], [11]. By constructing a Cauchy sequence $x_k(t), y_k(t), k = 0,1,2,\ldots$ as follows:

$$x_k(t) + \int_t^T \xi(s,x_{k-1}(s),y_{k-1}(s),\gamma(s))ds + \int_t^T y_k(s)dB(s),$$

starting from any square integrable processes pair$(x_0(t), y_0(t))$, and applying Proposition 2.2 in [11], we can prove the solution is uniformly bounded. Details are omitted because of space and the fact that these techniques are standard.

III. Adjoint Equations, Maximum Principle, and Conditions to Near-Optimality

Most results and techniques in this section can be found in [14], [6]. They are critical in understanding the computation scheme proposed in this paper.

For a given admissible control input $\bar{u}(t)$, let the trajectory of the equation (1) be denoted as $\bar{x}(t)$. Also let the first and second adjoint processes, denoted as $y_1(t)$ and $y_2(t)$ respectively, be the same as that defined in [14]. Let two costs be defined as:

$$J_1(u(\cdot)) = E[\int_{t_0}^{T} \frac{\partial L}{\partial x}(t,\bar{x}(t),\bar{u}(t))y_1(t)dt + \frac{\partial h}{\partial x}(\bar{x}(T))y_1(T)], \quad (4)$$

$$J_2(u(\cdot)) = E[\int_{t_0}^{T} \frac{\partial L}{\partial x}(t,\bar{x}(t),\bar{u}(t))y_2(t)dt + \frac{\partial h}{\partial x}(\bar{x}(T))y_2(T)]. \quad (5)$$

The associated adjoint equations are given by:

$$dp(t) = -\{\frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t))^\top p(t) + \sum_{i=1}^{m} \frac{\partial \sigma_{\cdot i}}{\partial x}(t, \bar{x}(t), \bar{u}(t))^\top K_{\cdot i}(t) + \frac{\partial L}{\partial x}(t, \bar{x}(t), \bar{u}(t))\}dt + K(t)dB(t),\ p(T) = \frac{\partial h}{\partial x}(\bar{x}(T)). \quad (6)$$

$$dQ(t) = -\{\frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t))^\top Q(t) + Q(t)\frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t)) + \sum_{i=1}^{m}[\frac{\partial \sigma_{\cdot i}}{\partial x}(t, \bar{x}(t), \bar{u}(t))^\top Q(t)\frac{\partial \sigma_{\cdot i}}{\partial x}(t, \bar{x}(t), \bar{u}(t)) + \frac{\partial \sigma_{\cdot i}}{\partial x}(t, \bar{x}(t), \bar{u}(t))^\top R_{\cdot i}(t) + R_{\cdot i}(t)\frac{\partial \sigma_{\cdot i}}{\partial x}(t, \bar{x}(t), \bar{u}(t))] + \Lambda(t)\}dt + R(t)dB(t),$$
$$Q(T) = \frac{\partial^2 h}{\partial x^\top \partial x}(\bar{x}(T)), \quad (7)$$

where $\Lambda(t) = \frac{\partial^2 L}{\partial x^\top \partial x}(t, \bar{x}(t), \bar{u}(t)) + \sum_{i=1}^{n}\{p^i(t)\frac{\partial^2 f^i}{\partial x^\top \partial x}(t, \bar{x}(t), \bar{u}(t)) + K^i(t)\frac{\partial^2 \sigma^i}{\partial x^\top \partial x}(t, \bar{x}(t), \bar{u}(t))\}$. Following the standard reasoning process in [6] and [14], the following relationships can be established:

$$J_1(u(\cdot)) = E\int_{t_0}^{T} p(t)^\top [f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))]dt + E\int_{t_0}^{T} \mathrm{tr}\{K(t)[\sigma(t, \bar{x}(t), u(t) - \sigma(t, \bar{x}(t), \bar{u}(t))]\}dt, \quad (8)$$

$$J_2(u(\cdot)) = E\int_{t_0}^{T} \frac{1}{2}\left[y_1^\top(t)\left(p(t)\frac{\partial^2 f}{\partial x^\top \partial x}(t, \bar{x}(t), \bar{u}(t)) + \sum_{i=1}^{m} K_i(t)\frac{\partial^2 \sigma^i}{\partial x^\top \partial x}(t, \bar{x}(t), \bar{u}(t))\right) y_1(t)\right] dt + E\int_{t_0}^{T} p^\top(t)[\frac{\partial f}{\partial x}(t, \bar{x}(t), u(t)) - \frac{\partial f}{\partial x}(t, \bar{x}(t), \bar{u}(t))]y_1(t)dt + E\int_{t_0}^{T}\sum_{i=1}^{m} K_i^\top(t)[\frac{\partial \sigma}{\partial x}(t, \bar{x}(t), u(t)) - \frac{\partial \sigma}{\partial x}(t, \bar{x}(t), \bar{u}(t))]y_1(t)dt. \quad (9)$$

Define the Hamiltonian function H as:

$$H(t, x, v, p, K) = L(t, x, v) + p^\top f(t, x, v) + \sum_{i=1}^{m} K_i^\top \sigma^i(t, x, v). \quad (10)$$

Notice that now we have $\Lambda(t) = \frac{\partial^2 H}{\partial x^\top \partial x}(t, \bar{x}(t), \bar{u}(t), p(t), K(t))$. We can define the third cost function:

$$J_3(u(\cdot)) = \frac{1}{2}E\left\{\int_{t_0}^{T} y_1^\top(t)\frac{\partial^2 H}{\partial x^\top \partial x}(t, \bar{x}(t), p(t), K(t))y_1(t)dt + y_1^\top(T)\frac{\partial^2 h}{\partial x^\top \partial x}(\bar{x}(T))y_1(T)\right\}. \quad (11)$$

And also following the reasoning process in [14] we can see that

$$J_3(u(\cdot)) = E\int_{t_0}^{T} \mathrm{tr}(Q(t)\Phi(t)) + \sum_{i=1}^{m} \mathrm{tr}(R_i(t)\Psi_i(t))dt, \quad (12)$$

where $\Phi(t), \Psi(t)$ are given by:

$$\Phi(t) = y_1(t)[f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))]^\top + [f(t, \bar{x}(t), u(t)) - f(t, \bar{x}(t), \bar{u}(t))]y_1^\top(t) + \left(\frac{\partial \sigma_{ij}}{\partial x}(t, \bar{x}(t), \bar{u}(t))y_1(t)\right) \cdot [\sigma(t, \bar{x}(t), u(t)) - \sigma(t, \bar{x}(t), \bar{u}(t))]^\top + [\sigma(t, \bar{x}(t), u(t)) - \sigma(t, \bar{x}(t), \bar{u}(t))] \cdot \left(\frac{\partial \sigma_{ij}}{\partial x}(t, \bar{x}(t), \bar{u}(t))y_1(t)\right)^\top + [\sigma(t, \bar{x}(t), u(t)) - \sigma(t, \bar{x}(t), \bar{u}(t))] \cdot [\sigma(t, \bar{x}(t), u(t)) - \sigma(t, \bar{x}(t), \bar{u}(t))]^\top, \quad (13)$$

$$\Psi(t) = y_1(t)[\sigma(t, \bar{x}(t), u(t)) - \sigma(t, \bar{x}(t), \bar{u}(t))]^\top + [\sigma(t, \bar{x}(t), u(t)) - \sigma(t, \bar{x}(t), \bar{u}(t))]y_1^\top(t). \quad (14)$$

Based on the solution of the two adjoint backward stochastic differential equations given by (6) and (7), we can also define a modified Hamiltonian function $\mathcal{H}$ as:

$$\mathcal{H}^{(\bar{x},\bar{u})}(t, x, u) = H(t, x, u, p(t), K(t) - Q(t)\sigma(t, \bar{x}(t), \bar{u}(t))) + \frac{1}{2}\sigma(t, x, u)^\top Q(t)\sigma(t, x, u). \quad (15)$$

Then, the so-called Maximum Principle in [6], [14], [15] and the major results in [6] can be summarized by the following theorem:

Theorem 1: Assume that conditions in Assumption 1 in previous section are all satisfied.

(1). The necessary condition for an admissible control $\bar{u}(t)$ and its corresponding solution $\bar{x}(t)$ minimize the cost function J if $\bar{u}(\cdot)$ minimizes the functional $\int_{t_0}^{T} \mathcal{H}^{(\bar{x},\bar{u})}(t, \bar{x}(t), u(t))dt$.

(2). There exists a constant $C_1 > 0$, such that an admissible control $\bar{u}(t)$ and its corresponding solution $\bar{x}(t)$ is ϵ-optimal if the following inequality is satisfied:

$$\int_{t_0}^{T} \mathcal{H}^{(\bar{x},\bar{u})}(t, \bar{x}(t), \bar{u}(t))dt \leq \inf_{u \in U_{ad}[t_0,T]} \int_{t_0}^{T} \mathcal{H}^{(\bar{x},\bar{u})}(t, \bar{x}(t), u(t))dt + C_1\epsilon. \quad (16)$$

(3). There exists a constant C_1, which is related to the Lipschitz constant C, the time interval $[t_0, T]$ and the value range Γ of admissible controls but independent of ϵ. Assume the Hamiltonian function $H(t,x(t),u(t),p,K)$ defined in (10) and the function $h(x)$ in the terminal cost of the cost function given by (2) are convex in with respect to t and x, respectively. If for some $\epsilon > 0$, there is an admissible control $\bar{u}$ such that the inequality (16) holds. Then,

$$J(\bar{u}(\cdot)) \leq \inf_{u \in U_{ad}[t_0,T]} J(u(\cdot)) + C_1 \epsilon^{\frac{1}{2}}.$$

In the proof of the result corresponding to (3) of Theorem 1 in [6], one can see that $\frac{\partial H}{\partial u}$ to be small is the key step. However, it is of more interests in this paper to evaluate whether $\int_{t_0}^{T} \mid \frac{\partial \mathcal{H}}{\partial u} \mid dt$ is small. In fact, it is the major criterion that our scheme based on. For this concern, we have the following theorem:

Theorem 2: Assume that conditions in Assumption 1 in previous section are all satisfied.

(1). There exists a constant $C_2 > 0$ independent of ϵ such that, if $u^{\epsilon}(\cdot)$ is an ϵ-optimal control, the following estimation holds:

$$\int_{t_0}^{T} \mid \frac{\partial \mathcal{H}^{x^{\epsilon},u^{\epsilon}}}{\partial u}(t, x^{\epsilon}(t), u^{\epsilon}(t)) \mid dt < C_2 \epsilon^{\frac{1}{2}}. \tag{17}$$

(2). Assume the Hamiltonian function H defined in (10) and the function $h(x(t))$ in the terminal cost of the cost function given by (2) are convex. Then, there is a constant $C_3 > 0$ independent of ϵ such that for any $\epsilon > 0$, as long as an admissible control $\bar{u}(\cdot)$ satisfies

$$\int_{t_0}^{T} \mid \frac{\partial \mathcal{H}^{x^{\epsilon},u^{\epsilon}}}{\partial u}(t, x^{\epsilon}(t), u^{\epsilon}(t)) \mid dt < C_3 \epsilon, \tag{18}$$

$\bar{u}$ is guaranteed to be an ϵ-optimal control.

The proof is omitted because of length limitation. Interested reader may request details from the author or see [16].

Algorithm 1: (Functional Stochastic Steepest Descent Scheme)

Initial step: Start with a randomly selected admissible control $\bar{u}(\cdot)$.

Updating rule:

- Calculate $\bar{x}_k(t)$ according to the stochastic equation (1) .
- Solve the first order adjoint equation (6) for $\bar{p}_k(t)$.
- Solve the second order adjoint equation (7) for $\bar{K}_k(t)$.
- Update $\bar{u}(t)$ with

$$\bar{u}_{k+1}(t) = \bar{u}_k(t) - \lambda(t) \frac{\partial \mathcal{H}^{\bar{x}_k,\bar{u}_k}}{\partial u}(t, \bar{x}_k(t), \bar{u}_k(t)), \tag{19}$$

where $\lambda(t)$ is a positive scaler random process bounded by a some constant $C_{\bar{u}(\cdot)}$.

Example:

Consider the case where the stochastic control system is linear and the cost function is quadratic. More specifically, consider the following problem:

$$\begin{aligned} dx(t) = \;& [A(t)x(t) + B(t)u(t) + b(t)]dt \\ & + [C(t)x(t) + D(t)u(t) + \sigma(t)]dW(t), \\ x(t_0) = \;& x_0, \end{aligned} \tag{20}$$

$$\begin{aligned} & J(u(\cdot)) = E \quad \frac{1}{2} \quad {}_{t_0}^{T} [< L(t)x(t), x(t) > \\ & +2 < M(t)x(t), u(t) > \\ & + < M(t)u(t), u(t) >]dt + \frac{1}{2} < Gx(T), x(T) > \end{aligned} \tag{21}$$

where $W(t)$ is scalar Brownian motion, L, M, N, G are matrices of appropriate dimensions, and $< \cdot, \cdot >$ denote the inner product in Euclidean space.

Now the adjoint equations become:

$$\begin{cases} dp(t) = & -\left\{ A^{\top}(t)p(t) + C^{\top}(t)K(t) \right. \\ & \left. + \frac{1}{2}L(t)x(t) + u^{\top}(t)M(t) \right\} dt \\ & + K(t)dW(t), \\ p(T) = & Gx(T), \end{cases} \tag{22}$$

$$\begin{cases} dQ(t) = & -\left\{ A^{\top}(t)Q(t) + Q(t)A(t) \right. \\ & + C^{\top}(t)Q(t)C(t) \\ & \left. + C^{\top}(t)N(t) + N(t)C(t) + L(t) \right\} dt \\ & + N(t)dW(t), \\ Q(T) = & G. \end{cases} \tag{23}$$

The Hamiltonian function H is the following:

$$\begin{aligned} & H(t, x_0, u, p, K) \\ = \; & \frac{1}{2} < L(t)x(t), x(t) > + < M(t)x(t), u(t) > \\ & + \frac{1}{2} < N(t)u(t), u(t) > \\ & + p^{\top}(t)[A(t)x(t) + B(t)u(t) + b(t)] \\ & + K^{\top}(t)[C(t)x(t) + D(t)u(t) + \sigma(t)]. \end{aligned}$$

It can be calculated that:

$$\begin{aligned} \frac{\partial \mathcal{H}^{\bar{x},\bar{u}}}{\partial u}(t, x, u) &= \frac{\partial H(t, x_0, u, \bar{p}, \bar{K})}{\partial u} \\ &= M(t)\bar{x}(t) + N(t)\bar{u}(t) \\ & \quad + B^{\top}(t)\bar{p}(t) + D^{\top}(t)\bar{K}(t). \end{aligned} \tag{24}$$

Apparently, the proposed scheme is a normal iterative method to solve the equation:

$$M(t)x(t) + N(t)u(t) + B^{\top}(t)p(t) + D^{\top}(t)K(t) = 0$$

for u. It is shown in [17, Page 309] that, together with positive semidefinitness of $R(t) - D^{\top}(t)Q(t)D(t)$, this equation is the linear quadratic version of Maximum Principle.

IV. Convergence Analysis

In this section, technical details of the convergence analysis will be given. Consider two consecutive controls $u_k(\cdot)$, and $u_{k+1}(\cdot)$ in the iterative computing process using scheme in Algorithm 1. We will analyze the convergence of the Algorithm 1 through evaluating the difference $J(u_{k+1}(\cdot)) - J(u_k(\cdot))$. First, we will see how well the first-order and second-order processes can be used to estimate the difference. To this end, the following lemma holds:

Lemma 3: For any two generic consecutive controls calculated based on the updating rule in Algorithm 1, assume the step size is chosen such that $E\int_{t_0}^{T} \mid \lambda_k(t) \mid^2 dt \leq 1$. The following estimation holds:

$$E\left|x_{k+1}(t)-x_k(t)-y_1(t)-y_2(t)\right|^2 \leq C_5 E(\int_{t_0}^{t} \mid \lambda_k(\tau) \mid^2 d\tau)^2,\ \forall t\in[t_0,T], \quad (25)$$

where $y_1(t), y_2(t)$ are first-order and second-order processes.
The proof is omitted here. Interested readers may see [16].

Now we are ready to present the convergence property of the Algorithm 1.

Theorem 3: For given an $\epsilon > 0$, assume that the admissible control value set Γ is big enough so that no ϵ-optimal control can reach the δ_0 neighborhood of the boundary of Γ, where δ_0 is a small positive constant. Then, there exist positive constants α, β and δ such that, if $\lambda_k(t)$ is selected in the interval $[\alpha, \beta]$ and satisfies $E\int_{t_0}^{T} \mid \lambda_k(t) \mid^2 dt \leq 1$, the cost value is guaranteed to decrease at least by the amount of δ, by using the proposed scheme starting from any admissible control. Therefore, if the cost function $J(u(\cdot))$ is bounded below, the proposed scheme achieves ϵ-optimal solution in finite steps.
The proof is omitted to save space. Interested reader may see [16] or ask the author for more details.

V. Consistent Approximation Solutions to Backward Stochastic Differential Equations

The convergence analysis conducted in the last section is based on the assumption that precise solutions to two adjoint equations can be available. Such an assumption is not practical. However, the analysis can still be valid with some modification. In this section, the approximation of backward stochastic differential equation is discussed. Rather than to develop an approximation algorithm, the objective is to consider the precision required for a consistent approximation algorithm so as to achieve desirable near-optimal solution to the original stochastic control problem.

Consider the case where two pairs of approximated solutions, $(\hat{p}_k(t), \hat{K}_k(t))$ and $(\hat{Q}_k(t), \hat{R}_k(t))$, are obtained to the adjoint equations (6) and (7), respectively, at the step k using the proposed iterative scheme. Let the adjoint processes corresponding to $\hat{u}_{k+1}$ be denoted as $\hat{y}_1(t), \hat{y}_2(t)$, respectively, where $\hat{u}_{k+1}$ is the next step control computed using the approximation solutions $\hat{p}_k, \hat{K}_k, \hat{Q}_k, \hat{R}_k$. For simplicity, we also denote all other variables associated with these approximation solution using a hat symbol $\hat{}$.

Carefully considering the convergence analysis in last section. We need the following lemma:

Lemma 4: Given an admissible control $\bar{u}(\cdot)$, let the corresponding state trajectory be denoted as $\bar{x}(\cdot)$, and the solution to the adjoint equations (6), (7) be denoted as $\bar{p}(\cdot), \bar{K}(\cdot), \bar{Q}(\cdot), \bar{R}(\cdot)$, respectively. Assume $\hat{p}(\cdot), \hat{K}(\cdot), \hat{Q}(\cdot), \hat{R}(\cdot)$ are approximated solutions to those adjoint equations such that

$$E\int_{t_0}^{T}[\mid \hat{p}(t)-\bar{p}(t) \mid^2 + \mid \hat{K}(t)-\bar{K}(t) \mid^2 + \mid \hat{Q}(t)-\bar{Q}(t) \mid^2 + \mid \hat{R}(t)-\bar{R}(t) \mid^2]dt < \eta, \quad (26)$$

where η is a given small positive real number. Then, the following results hold:
(1). $E\int_{t_0}^{t}\left|\frac{\partial H}{\partial u}(\tau, \hat{x}_k, \hat{u}_k, \hat{p}_k(\hat{u}_k), \hat{K}_k(\hat{u}+k))\right|^2 d\tau$ is bounded by a constant. This constant is depend on the Lipschitz constant C and the approximation error η.
(2). There is a constant $\hat{C}$ such that the auxiliary cost functions J_1, J_2, J_3 defined in (4), (5), and (11), respectively, satisfy the following estimation:

$$\mid J_i(u(\cdot)) - \hat{J}_i(u(\cdot)) \mid < \hat{C}\eta, i=1,2,3., \quad (27)$$

where $\hat{J}_i$ is the corresponding costs defined in (8), (9), and (12) with $p(t)$, $K(t)$, $Q(t)$, $R(t)$ replaced by $\hat{p}(t), \hat{K}(t), \hat{Q}(t), \hat{R}(t)$, respectively.
Proof is omitted. See [16] for details.

Applying the claim (1) in Lemma 4, one can see the following result corresponding to Lemma 3:

Corollary 1: Under the assumption in Lemma 4, for the second order approximation of the trajectory process $x(\cdot)$ at $(\hat{x}_k, \hat{u}_k)$ based on the approximation solution of the corresponding BSDEs, there exists a constant $\hat{C}_5$ such that:

$$E\left|\hat{x}_{k+1}(t)-\hat{x}_k(t)-\hat{y}_1(t)-\hat{y}_2(t)\right|^2 \leq \hat{C}_5 E(\int_{t_0}^{t} \mid \lambda_k(\tau) \mid^2 d\tau)^2, \forall t\in[t_0,T], \quad (28)$$

where $\hat{C}_5$ is only depend on the Lipschitz constant C and the approximation error bound η.

Remark 1: In fact, the assumption in (26) is not restrictive for numerical algorithms of BSDE. In [12], a schema has been proposed for the equation

$$Y_t + \int_t^T Z_s dB_s = \xi + \int_t^T f(s,\omega,Y_s,Y_s,Z_s)ds,$$

such that

$$E\mid Y(t)-Y_n(t)\mid^2 + E\int_t^T \mid Z(s)-Z_n(s)\mid^2 ds \leq C'''(\epsilon_n + \delta_n + 1/\lambda),$$

where C''' is a constant, ϵ_n and δ_n are discretization error in time and space, λ is a parameter for a related poison process that can be converge to positive infinity in the approximation process, (Y_t, Z_t) and (Y_n, Z_n) are the precise and approximated solutions, respectively. It can be seen that this convergence result guarantees (26).

Theorem 4: Assume all assumptions and conditions in Theorem 3 are satisfied. Choose $\eta_k \leq \sup_{t\in[t_0,T]} \lambda_k^2(t)$. Then, there exist positive constants α, β and δ such that, if $\lambda_k(t)$ is selected in the interval $[\alpha, \beta]$ and satisfies

$$E\int_{t_0}^{T} \mid \lambda_k(t) \mid^2 dt \leq 1,$$

the cost value is guaranteed to decrease at least by the amount of δ, by using the proposed scheme starting from any admissible control with the approximated solutions of the adjoint equation satisfying (26) where the η is replaced by η_k at step k. Therefore, if the cost function $J(u(\cdot))$ is bounded below, the proposed computation scheme achieves ϵ-optimal solution in finite steps.
Proof is omitted. See [16] for details.

VI. Concluding Remarks

In this paper, a scheme is proposed for the computation of optimal control for general stochastic system where the diffusion term is affected by control input. It is developed based on the concept of near-optimality and the generalization of the well-known Maximum Principle to first-order condition. An Hamiltonian type auxiliary function is constructed to capture the gradient information at optimal solution. Based on that, a gradient direction is calculated for the searching of optimal control.

The proposed scheme provides us with a framework for discretization for numerical computation. This can be seen in two aspects. On one hand, the BSDEs can be solved numerically as long as the error is consistent to the desired one. On the other hand, the control updating rule is automatically discretized if the initial admissible control is discrete and the BSDEs are solved numerically. In this sense, the proposed scheme differs from the Markov approximation methods in [2] in the way that we compute the control iteratively at the first stage, then this control is updated by numerical solution to a group of equations free of control while in their methods, the system is approximated using a normalized markov discretization techniques. Then, the resulted controlled Markov system is solved for optimal control.

The proposed scheme also provides some insight of the "gradient direction" in the optimization problems where the cost is a function while the constraints are stochastic equations. It is of interests in pursue further understanding of such type of concept in a more abstract sense so as to develop better optimization methods in functional spaces.

The computation scheme is based on the numerical solution to sequences of backward stochastic differential equations. The consistency of such numerical solution to the proposed scheme is discussed. However, numerical computation for BSDE is not studied in detail in this work. Heuristically, we believe that it is easier to solve than the inequality with adjoint equation constraints in the generalized Maximum Principle. There are some computation methods have been proposed for BSDEs. For examples, see the [12], [9] and references cited therein. It is still of interest to investigate the related consistent BSDE algorithm.

This scheme is only proposed here theoretically, even though we believe it can be implemented easily provided that a consistent algorithm to BSDEs is available. It is also an important issue in future to develop the corresponding numerical computation package so as to compare the proposed scheme and major existing algorithm by numerical simulation.

References

[1] Harold J. Kushner and Dupuis Paul G. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer-Verlag, New York, 1992.

[2] H. J. Kushner. Numerical methods for stochastic control problems in continuous time. *SIAM Journal on Control and Optimization*, pages 999–1048, 1990.

[3] Claus Munk. Numerical methods for continuous-time, continuous-state stochastic control problems. working paper. 1997.

[4] C. S. Tapiero and A Sulem. Computational aspects in applied stochastic control. *Computational Economics*.

[5] F. B. Hanson. Computational stochastic control: basic foundations, complexity and techniques. In *Proceedings of 42nd IEEE Conference on Decision and Control*, pages 3024–3029, Hawii, USA, 2003.

[6] Xun Yu Zhou. Stochastic near-optimal controls: Necessary and sufficient conditions for near-optimality. *SIAM Journal on Control and Optimization*, pages 929–947, 1998.

[7] Jin Ma and J. Yong. *Forwad-Backward Stochastic Differential Equations and Their Applications*. Springer-Verlag, Berlin Heidelberg, 1999.

[8] E Pardoux and Peng S.G. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, pages 55–61, 1990.

[9] Jin Ma, Philip Protter, Jaime San Matin, and Soledad Torres. Numerical method for backward stochastic differential equations. *Annals of Applied Probability*, pages 302–316, 2002.

[10] Bruno Bouchard and Touzi Nizar. Discrete time approximation and monte-carlo simulation of backward stochastic differential equations. *Preprint*, 2002.

[11] N. E. Karoui. Backward stochastic differential equations. In *Backward Stochastic Differential Equations*, pages 7–26. Addison Wesley Longman Ltd., Edinburgh Gate, Harlow Essex CM20 2JE, England, 1997.

[12] V. Bally. Approximation scheme for solutions of bsde. In *Backward Stochastic Differential Equations*, pages 177–191. Addison Wesley Longman Ltd., Edinburgh Gate, Harlow Essex CM20 2JE, England, 1997.

[13] I. Ekeland. On the variational principle. *Journal of Mathematical Analysis and Applications*, pages 324–253, 1974.

[14] Shige Peng. A general stochastic maximum principle for optimal control problems. *SIAM Journal on Control and Optimization*, pages 966–979, 1990.

[15] Xun Yu Zhou. A unified treatment of maximum principle and dynamic programming in stochastic controls. *Stochastics and Stochastics Reports*, pages 137–161, 1991.

[16] Danchi Jiang. A finite step scheme for general near-optimal stochastic control. *Working paper*, 2004.

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