

RESEARCH

Open Access

# Stochastic optimal control to a nonlinear differential game

Othusitse Basimanebotlhe<sup>1,2\*</sup> and Xiaoping Xue<sup>1</sup>

\*Correspondence:

ot\_bas@yahoo.co.uk

<sup>1</sup>Department of Mathematics,  
Harbin Institute of Science and  
Technology, Nangang District,  
Harbin, 150001, P.R. China

<sup>2</sup>Department of Mathematics,  
University of Botswana, 4775  
Notwane Rd., Gaborone, Botswana

## Abstract

The paper studies the optimal control of a nonlinear stochastic differential game of two persons subjected to noisy measurements. The logarithmic transformation to the value function is used in trying to find the solution of the problem. The conversion of a quasilinear partial differential equation to an ordinary linear differential equation is considered. Lastly, the iterative optimal control path estimates for the minimization maximization differential game are attained.

**Keywords:** nonlinear stochastic differential equation; Brownian motion; stochastic optimal control; Ito's lemma

## 1 Introduction

Control theory is a field of mathematics and engineering used in a wide range of fields and their applications, such as architecture, communications, queueing theory, robotics and in economics as evidenced in [1–3] just to mention a few. Control theory is a subject of much interest in today's real world. As stated in [4], optimizing a sequence of actions to attain some future goal is the general topic of control theory. Therefore, the objective of optimal control theory is to attain an optimal regulation of the system evolution [5]. Without the indulgence of the noise, the continuous time control problems can be solved in two ways: using Pontryagin Minimum Principle (PMP), which is a pair of ordinary differential equations, or the Hamilton-Jacobi-Bellman (HJB), which is a partial differential equation as in [4]. The addition of differential equations as constraints in the optimization problem leads to the property that in optimal control theory the minimum is no longer represented by one point  $x^*$  in the state space but by a path or trajectory  $x^* = (x_i^*)_{i=1,\dots,N}$ , which is known as the optimal trajectory.

In the presence of noise, the PMP formalism has no obvious generalization as mentioned in [6]. However, the inclusion of the noise in the HJB framework is mathematically quite straightforward, while the numerical solution of either deterministic or stochastic HJB equation is difficult due to the curse of dimensionality. A control problem is said to be stochastic when it is subjected to some disturbances or noise terms and time dependent, that is being uncertain of its future state. Under control theory, there lies a topic of interest to this paper which is game theory.

Game theory deals with strategic interactions among several decision makers, known as players. These players have objectives that may be contradicting or non-contradicting. In situations whereby players' aims are not in contradiction, they are said to be in coop-

eration, which is enough to regard their aims as of one player, as seen in [7] and [8]. If the players' decisions are in contradiction, then it is widely known as non-cooperative game theory that is even more interesting as each player may have to be analyzed individually or in distinct groups basing on the similarity of the objectives. Non-cooperative game theory has been discussed by several authors [9–12].

The nonlinear stochastic optimal control theory is one of the optimization fundamentals with plethora of applications in several domains, see [13] and [14]. The main difficulty associated with stochastic minimax dynamic games studied here is the presence of the noise in the dynamical constraints and the solution to a nonlinear second order Hamilton Jacobi Bellman equation (HJB) as mentioned in [4].

In this paper a nonlinear stochastic problem is considered as a two-person zero sum game. One player tries to attain the reward at minimal costs, while the other player tries to maximize the costs of the other player, knowing that it would be advantageous to him. The problem is modeled as a dynamical system with nonlinear stochastic constraints, and our idea is to find the saddle point of the game after a certain time of play. The two controls act as opponents, that is, the control  $u$  is a stabilizer or minimizer, while  $v$  is the destabilizing or maximizing control variable. A different approach to finding the solution of the problem deployed through certain conditions assumed in this paper is similar to that of [15].

## 2 Problem formulation

Let  $(\Omega, \mathcal{F}, \mathbb{P}, \mathcal{F}_t)$  be a complete probability space, and let time interval  $[0, T]$  with  $0 < T < \infty$  be given. Assume that on this space an  $n$ -dimensional Brownian motion  $\{W(t), \mathcal{F}_t\}_{t \in [0, T]}$  is defined with  $\{\mathcal{F}_t\}_{t \in [0, T]}$  as the Brownian filtration, Guo *et al.* [16]. The expectation under the probability measure  $\mathbb{P}$  will be denoted by  $E^t$ , while  $\mathcal{F}$  is a  $\sigma$ -algebra of the subsets of  $\Omega$ .

**Definition 2.1** (Brownian motion) Brownian motion  $\{W(t)\}$  is the stochastic process with the following properties [17].

- (i) At any time  $s < t$ , the increment  $W(t) - W(s)$  is Gaussian with mean zero and variance  $E^t[(W(t) - W(s))^2] = t - s$ ; moreover, the increments associated with disjoint intervals are independent.
- (ii) Its sample path is continuous, that is, the function  $t \rightarrow W(t)$  is almost surely continuous.
- (iii) Its initial state and time is zero, that is,  $W(0) = 0$ .

Let  $\{x(t, \omega)\}_{t \in [0, T]}$  denote a stochastic process, then  $\{x(t, \omega)\}_{t \in [0, T]}$  is said to be  $\mathcal{F}_t$  adapted whenever  $\mathcal{F}_t$  is known at time  $t$ .

Consider the evolution of the system given by a nonlinear stochastic differential game of two persons,

$$\begin{aligned} dx(t, \omega) &= g(x(t, \omega), u(t, \omega), v(t, \omega), t) dt + \sigma(x(t, \omega)) dW(t) \quad \text{for } t \in [0, T], \\ x(0, \omega) &= x_0(\omega). \end{aligned} \tag{1}$$

Here, the volatility coefficient of the noise term  $\sigma$  is only dependent on the state trajectory  $x$ , similarly as addressed by [18]. The function  $g$  is dependent on the controls or policies, the state trajectory and the time  $t$ . The value  $x_0(\omega)$  denotes the initial random state. The

problem is formulated as a differential game with two opponents as in Theodorou [15] with

$$\begin{aligned} & \min_u \max_v J(t, x(t, \omega), u(x(t, \omega)), v(x(t, \omega))) \\ & = \min_u \max_v E^t \left\{ \phi(x(T, \omega)) + \int_0^T e^{-\beta t} \mathcal{L}(t, x(t, \omega), u(x(t, \omega)), v(x(t, \omega))) dt \right\} \end{aligned} \quad (2)$$

with

$$\begin{aligned} & \mathcal{L}(t, x(t, \omega), u(x(t, \omega)), v(x(t, \omega))) \\ & = q(x(t, \omega)) + \frac{1}{2} u^T(x(t, \omega)) R(t, \omega) u(x(t, \omega)) \\ & \quad - v^T(x(t, \omega)) S(t, \omega) v(x(t, \omega)), \end{aligned} \quad (3)$$

under stochastic nonlinear dynamical constraints,

$$\begin{aligned} dx(t, \omega) &= (f(x(t, \omega)) + G(\omega)u(x(t, \omega)) + H(\omega)v(x(t, \omega))) dt \\ & \quad + \sigma(t, x(t, \omega)) dW(t), \end{aligned} \quad (4)$$

where

- (i) the function  $\mathcal{L} : [0, T] \times \mathbb{R}^n \times E^p \times E^q \rightarrow \mathbb{R}$  is known as the immediate cost to go, while  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$  is the terminal cost;
- (ii)  $x(t, \omega) \in \mathbb{L}_2(\Omega, \mathcal{F}, \mathbb{P})$  is the  $n$ -dimensional random state vector for  $t \geq 0$  with  $\omega \in \Omega$  as the supporting set of a complete probability measure space  $\mathbb{P}$ ;
- (iii)  $u(x(t, \omega)) \in E^p$  and  $v(x(t, \omega)) \in E^q$  are the  $p$ -dimensional and  $q$ -dimensional random control variables of player one and player two defined respectively in the two metric spaces;
- (iv) the symmetric and positive definite random time varying matrices  $R(t, \omega)$  and  $S(t, \omega)$  are matrices associated with the controls of  $u(x(t, \omega))$  and  $v(x(t, \omega))$ , respectively;
- (v)  $\beta > 0$  is the discounting factor of the value function.

**Definition 2.2** (Admissible control) From [18] a control  $\{u(t), \mathcal{F}_t\}_{t \in [0, T]}$  is said to be admissible if

- (i) for every  $(t, x)$  the system of SDEs in (1) with initial condition  $x(0, \omega) = x_0(\omega)$  admits a pathwise unique strong solution;
- (ii) there exists some function  $\phi : \mathbb{R}^n \rightarrow U$  of class  $C^{1,2}$  such that  $u$  is in relative feedback to  $\phi$ , that is,  $u(t) = \phi(x(t))$  for every  $t \in [0, T]$ .

The two controls  $u(x(t, \omega))$  and  $v(x(t, \omega))$  are referred to as the stabilizing and the destabilizing controllers, respectively. The stabilizing controller minimizes the cost function, while the destabilizing controller tries to maximize the cost function.

Considering the nonlinear stochastic dynamical constraints, let

- (i)  $f : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}$  and  $\sigma : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times m}$  be bounded and continuous functions;
- (ii)  $G(\omega)$  and  $H(\omega)$  be  $n \times p$  and  $n \times q$  random matrices.

### 3 Conditions

The following standard regularity conditions are made as in [15] and [19] throughout the paper.

- (i) The functions  $f(x)$ ,  $L(t, x, u, v)$  are continuously differentiable in  $(t, x, u, v) \in [0, T] \times \mathbb{R}^n \times E^p \times E^q$  and  $\phi$  is twice differentiable in  $x \in \mathbb{R}^n$ .
- (ii)  $J$  and  $\mathcal{L}$  are nonnegative functions.
- (iii)  $V(x)$ ,  $\nabla_x V$ ,  $\mathcal{L}$ , and  $\phi$  are bounded, where  $V$  is defined as the value of the game that each control tries to optimize from.
- (iv) The controls are bounded in  $\mathbb{R}^m$  and given in the spaces

$$L_{\mathcal{F}}^2 := \{u : \|u(x(t, \omega))\|_2\} < \infty,$$

$$L_{\mathcal{F}}^\infty := \{u : \|u(x(t, \omega))\|_\infty\} < \infty,$$

thus

$$\|u(x(t, \omega))\|_2 = E^t \left( \int_0^T |u(x(t, \omega))|^2 dt \right),$$

$$\|v(x(t, \omega))\|_\infty = E^t \{ \sup |v(x(t, \omega))| : t \in [0, T] \}$$

for  $v$  continuous. The same conditions for  $u$  under (iv) are applicable to  $v$  as well.

### 4 Approach to the solution of stochastic optimal controls

Our approach in finding the optimal control is based on the definition of the saddle point given below as in [20] with slight changes to suit our problem.

#### Definition 4.1

- (i) If the pair  $(u^*(x(t, \omega)), v^*(x(t, \omega))) \in U_1 \times U_2$  is optimal, then there exists a saddle point of the game over the interval  $[0, T]$  with respect to  $x(t, \omega) \in \mathbb{R}^n$ , if

$$\begin{aligned} J(t, x(t, \omega), u^*(x(t, \omega)), v(x(t, \omega))) \\ \leq J(t, x(t, \omega), u^*(x(t, \omega)), v^*(x(t, \omega))) \\ \leq J(t, x(t, \omega), u(x(t, \omega)), v^*(x(t, \omega))) \end{aligned}$$

for all  $u(x(t, \omega)) \in U_1$  and  $v(x(t, \omega)) \in U_2$ , where  $U_1$  and  $U_2$  are nonempty sets of admissible controls.

- (ii) The upper value of the game at any path  $x(t, \omega)$  and time  $t \in [0, T]$  is defined by

$$V^*(x(t, \omega)) = \inf_{u \in U_1} \sup_{v \in U_2} J(t, x(t, \omega), u(x(t, \omega)), v(x(t, \omega))),$$

and the lower value of the game is

$$V_*(x(t, \omega)) = \sup_{v \in U_2} \inf_{u \in U_1} J(t, x(t, \omega), u(x(t, \omega)), v(x(t, \omega)))$$

and if

$$V^*(x(t, \omega)) = V_*(x(t, \omega)) \equiv V(x(t, \omega)).$$

The objective is to find the optimal admissible controls,  $u^*(x(t, \omega)) \in U_1$  and  $v^*(x(t, \omega)) \in U_2$ , such that  $V(x(t, \omega), t)$  satisfies Definition 4.1 for  $U_1 \subset E^p$  and  $U_2 \subset E^q$ .

**Theorem 4.1** (Bellman principle of optimality) *If  $u^*(x(t, \omega))$  is optimal over the interval  $[0, T]$  starting at an initial state  $x_0(\omega)$ , then  $u^*(x(t, \omega))$  is necessarily optimal over the subinterval  $[t, t + dt]$  for any  $dt$  such that  $T - t \geq dt > 0$ .*

For the proof of the above theorem, refer to [15].

Applying Theorem 4.1 and Definition 4.1 to the value of the game  $V(x(t, \omega))$ , we have that

$$\begin{aligned} V(x(t, \omega)) &= \min_u \max_v E^t \left\{ \phi(x(T, \omega)) \right. \\ &\quad \left. + \int_t^{t+dt} e^{-\beta\tau} \mathcal{L}(\tau, x(\tau, \omega), u(x(\tau, \omega)), v(x(\tau, \omega))) d\tau \right\} \\ &= E^t \left\{ \mathcal{L}(t, x(t, \omega), u^*(x(t, \omega)), v^*(x(t, \omega))) dt + e^{-\beta dt} V(x(t + dt, \omega)) \right\} \\ &= \mathcal{L}(t, x(t, \omega), u^*(x(t, \omega)), v^*(x(t, \omega))) dt \\ &\quad + t(1 - \beta dt) E^t [V(x(t + dt, \omega))]. \end{aligned} \quad (5)$$

We need to calculate the expectation of the function  $V(x(t + dt, \omega))$ . Approximating the function  $V(x(\cdot, \omega))$  using Taylor's formula, we have

$$\begin{aligned} V(x(t + dt, \omega)) &= V(x(t, \omega)) + V'(x(t, \omega)) [x(t + dt, \omega) - x(t, \omega)] \\ &\quad + \frac{1}{2} V''(x(t, \omega)) [x(t + dt, \omega) - x(t, \omega)]^2 + \dots \end{aligned}$$

Ignoring the terms of higher powers and letting  $dx(t, \omega) = x(t + dt, \omega) - x(t, \omega)$ , we get

$$V(x(t + dt, \omega)) = V(x(t, \omega)) + V'(x(t, \omega)) [dx(t, \omega)] + \frac{1}{2} V''(x(t, \omega)) [dx(t, \omega)]^2. \quad (6)$$

Substituting the stochastic equation (4) into equation (6) and using the properties of Ito's lemma, we give the function  $V(x(t + dt, \omega))$  by

$$\begin{aligned} V(x(t + dt, \omega)) &= V(x(t, \omega)) + \left[ (f(x(t, \omega)) + G(\omega)u(x(t, \omega)) \right. \\ &\quad \left. + H(\omega)v(x(t, \omega))) V'(x(t, \omega)) \right. \\ &\quad \left. + \frac{1}{2} \sigma^2(x(t, \omega), t) V''(x(t, \omega), t) \right] dt \\ &\quad + \sigma(x(t, \omega), t) V'(x(t, \omega)) dW(t). \end{aligned} \quad (7)$$

Taking the expectation of equation (7), we have

$$\begin{aligned} E^t [V(x(t + dt, \omega))] &= V(x(t, \omega)) + \left[ (f(x(t, \omega)) + G(\omega)u(x(t, \omega)) \right. \\ &\quad \left. + H(\omega)v(x(t, \omega))) V'(x(t, \omega)) \right. \\ &\quad \left. + \frac{1}{2} \sigma^2(x(t, \omega), t) V''(x(t, \omega), t) \right] dt. \end{aligned} \quad (8)$$

Substituting equation (8) to equation (5) yields

$$\begin{aligned}\beta V(x(t, \omega)) &= \mathcal{L}(t, x(t, \omega), u^*(x(t, \omega)), v^*(x(t, \omega))) + [f(x(t, \omega)) \\ &\quad + G(\omega)u(x(t, \omega)) + H(\omega)v(x(t, \omega))]V'(x(t, \omega)) \\ &\quad + \frac{1}{2} \text{Tr}[\sigma \sigma^T(x(t, \omega), t) V''(x(t, \omega))].\end{aligned}\quad (9)$$

The above equation is the Bellman equation similar to the one in [21] which is a parabolic differential equation that has simple solutions for some simple processes and utility functions. In this paper we will adopt the idea of [22] instead of solving the Bellman equation, which is not always easy. From the Bellman equation we can solve for the optimum values  $u(x(t, \omega)) \in U_1$  and  $v(x(t, \omega)) \in U_2$ , by taking the derivative with respect to  $u(x(t, \omega))$  and  $v(x(t, \omega))$ ,

$$\begin{aligned}0 &= \mathcal{L}_u + G^T(\omega) V_x(x(t, \omega), t), \\ R(t, \omega) u(x(t, \omega)) &= -G^T(\omega) V_x(x(t, \omega), t), \\ u^*(x(t, \omega)) &= -R^{-1}(t, \omega) G^T(\omega) V_x(x(t, \omega), t).\end{aligned}\quad (10)$$

As for the maximizer  $v(x(t, \omega))$ , we have

$$\begin{aligned}0 &= \mathcal{L}_v + H^T(\omega) V(x(t, \omega), t), \\ S(t, \omega) v(x(t, \omega)) &= H^T(\omega) V_x(x(t, \omega), t), \\ v^*(x(t, \omega)) &= S^{-1}(t, \omega) H^T(\omega) V_x(x(t, \omega), t).\end{aligned}\quad (11)$$

Substituting the values of  $u^*(x(t, \omega))$  and  $v^*(x(t, \omega))$  onto  $\mathcal{L}$  in equation (9) and collecting the like terms yields the expression

$$\begin{aligned}\beta V(x(t, \omega)) &= q(x(t, \omega)) \\ &\quad - \frac{1}{2} V_x^T(x(t, \omega), t) G(\omega) R^{-1}(t, \omega) G^T(\omega) V_x(x(t, \omega)) \\ &\quad + \frac{1}{2} V_x^T(x(t, \omega), t) H(\omega) S^{-1}(t, \omega) H^T(\omega) V_x(x(t, \omega)) \\ &\quad + V_x^T(x(t, \omega), t) f(x(t, \omega)) \\ &\quad + \frac{1}{2} \text{Tr}[V_{xx}(x(t, \omega)) \sigma(x(t, \omega)) \sigma^T(x(t, \omega))].\end{aligned}\quad (12)$$

Equation (12) is a nonlinear second order partial differential equation (PDE), and its solution is a bit challenging as it is nonlinear and in high dimensions. As assumed in [13], there is a connection between the controls and the variance of the Brownian noise. Considering the difference in our control weights, we have the following cases:

- (i)  $HS^{-1}H^T - GR^{-1}G^T < 0$  implies that more weight is on the minimizing control than on the maximizing control variable.
- (ii)  $HS^{-1}H^T - GR^{-1}G^T > 0$  implies more weight on the maximizing control than on the minimizing control variable.

(iii)  $HS^{-1}H^T - GR^{-1}G^T = 0$ , the weights of the controls are equivalent, hence it is an ideal situation for a minimax optimal control.

The intuition we get from [13] is that the higher the variance, the lower the weight of the controls, hence 'cheap' controls and *vice versa*. In our case we want to strike a deal such that both players attain their optimums. The variance of the Brownian noise here is given by  $\sigma\sigma^T > 0$ , therefore we want to attain a situation whereby  $\lambda(t)[GR^{-1}G^T - HS^{-1}H^T] = \sigma\sigma^T$  for all  $x \in \mathbb{R}^n$  and  $t \in [0, T]$ , where the difference of the control coefficients will be the same as the variance of the noise. Our assumption on the balancing parameter is different from the one suggested by other authors, as in [13] and [15], where the balancing term is just a constant parameter. In our case, the balancing variable  $\lambda(t)$  is dependent on  $t$  such that at any time instant the equality sign is attained as the variance terms differing with time.

Suppose that

$$V(x(t, \omega)) = -\lambda(t) \log \Phi(x(t, \omega)). \quad (13)$$

We determine all the partial derivatives of the new value function given in equation (13),

$$V_x(x(t, \omega)) = -\lambda(t) \frac{1}{\Phi(x(t, \omega))} \Phi_x(x(t, \omega)) \quad (14)$$

and

$$V_{xx}(x(t, \omega)) = -\lambda(t) \frac{\Phi_{xx}(x(t, \omega))\Phi(x(t, \omega)) - \Phi_x(x(t, \omega))\Phi_x(x(t, \omega))}{\Phi(x(t, \omega))\Phi^T(x(t, \omega))}. \quad (15)$$

Therefore substituting (13), (14), (15) and taking into consideration the assumption that  $\lambda(t)[GR^{-1}G^T - HS^{-1}H^T] = \sigma\sigma^T$  for all  $t \in [0, T]$  to the nonlinear PDE given in (12), we have

$$\begin{aligned} \beta \Phi(x(t, \omega)) \log \Phi(x(t, \omega)) &= -\frac{1}{\lambda(t)} \Phi(x(t, \omega)) q(x(t, \omega)) \\ &\quad + \Phi_x^T(x(t, \omega)) f(x(t, \omega)) \\ &\quad + \frac{1}{2} \text{Tr}[\Phi_{xx}(x(t, \omega)) \sigma(x(t, \omega)) \sigma^T(x(t, \omega))], \end{aligned} \quad (16)$$

which yields a second order quasilinear PDE with the boundary condition given as

$$\Phi(x(T, \omega)) = \exp\left(-\frac{1}{\lambda(T)} \phi(x(T, \omega))\right). \quad (17)$$

If the solution  $\Phi(x(t, \omega))$  is found to exist for equation (16), then we have the results given below.

**Theorem 4.2** *If  $\Phi(x(t, \omega))$  satisfies equation (16), then the transformed control optimums are given as*

$$u^*(x(t, \omega)) = \lambda(t) R^{-1}(t, \omega) G^T(\omega) \frac{\Phi_x(x(t, \omega))}{\Phi(x(t, \omega))}$$

and

$$v^*(x(t, \omega)) = -\lambda(t)S^{-1}(t, \omega)H^T(\omega) \frac{\Phi_x(x(t, \omega))}{\Phi(x(t, \omega))}$$

for the value

$$V(x(t, \omega)) = \lambda(t) \log \Phi(x(t, \omega)),$$

where  $\lambda(t)$  satisfies

$$\lambda(t)[GR^{-1}G^T - HS^{-1}H^T] = \sigma\sigma^T, \quad \forall x \in \mathbb{R}^n, t \in [0, T].$$

One would observe that  $u^*(x(t, \omega))$  is now positive while  $v^*(x(t, \omega))$ , this is so because the problem has been transformed from minimax to maxmin problem. The PDE in (16) is found to be a bit difficult to solve in terms of dependence variables  $x$  and  $t$ , therefore in this paper we resort to transforming the above PDE to an ODE for which, in most cases, a solution can be obtained. Consider a one-dimensional problem for this case, thus  $n = 1$  and fix  $t$ , then the equation becomes more dependent on  $x$ . This leads to a nonlinear ODE, and before solving the nonlinear ODE, we have the following assumptions.

- (A:1) (i)  $\Phi(x(t, \omega), t)$ ,  $f(x(t, \omega))$  and  $q(x(t, \omega))$  are nonnegative functions.  
(ii)  $\Phi(x(t, \omega), t)$  is Lipschitz continuous for all  $(t, x) \in ([0, T] \times \mathbb{R})$  and  $\omega \in \Omega$ .  
(iii)  $f(x(t, \omega))$  and  $q(x(t, \omega))$  are also continuous functions and bounded functions for all  $x \in \mathbb{R}$ .

Let

$$\sigma(x(t, \omega))\sigma^T(x(t, \omega)) = \theta(x(t, \omega)) > 0.$$

Multiplying throughout by  $\theta^{-1}(x(t, \omega))$ , we have

$$\left[ \frac{d^2 \Phi(x(t, \omega), t)}{dx^2} + \tilde{f}(x) \frac{d\Phi(x(t, \omega), t)}{dx} \right] = \frac{2}{\lambda} \tilde{q}(x) \Phi(x(t, \omega), t) + r(x) \Phi(x(t, \omega), t) \log \Phi(x(t, \omega), t), \quad (18)$$

where

$$\begin{aligned} \tilde{f}(x(t, \omega)) &= 2f(x(t, \omega))\theta^{-1}(x(t, \omega)), \\ \tilde{q}(x(t, \omega)) &= 2q(x(t, \omega))\theta^{-1}(x(t, \omega)) \end{aligned}$$

and

$$r(x(t, \omega)) = \beta\theta^{-1}(x(t, \omega)).$$

For transformation and simplicity purposes, we would represent the following functions as  $\mathcal{U} = \Phi(x(t, \omega), t)$  and  $\mathcal{V} = \frac{d\Phi(x(t, \omega), t)}{dx}$ .

This yields the following first order ODE:

$$\begin{cases} \dot{\mathcal{U}} = \mathcal{V}, \\ \dot{\mathcal{V}} = -\tilde{f}(x)\mathcal{V} + \frac{2}{\lambda}\tilde{q}(x)\mathcal{U} + r(x)\mathcal{U} \log \mathcal{U}, \end{cases} \quad (19)$$



which gives the equation

$$F(x, \dot{\mathcal{U}}, \dot{\mathcal{V}}) = \begin{pmatrix} \mathcal{V} \\ -\tilde{f}(x)\mathcal{V} + \frac{2}{\lambda}\tilde{q}(x)\mathcal{U} + r(x)\mathcal{U} \log \mathcal{U} \end{pmatrix}. \quad (20)$$

Given the following conditions:

- (A:2) (i)  $\mathcal{U} = \mathcal{U}_1 \times \mathcal{U}_2 \in \mathbb{R}^m$  is a compact and bounded set.  
(ii)  $I \in \mathbb{R}$  is bounded.  
(iii)  $H = [a, b]$ ,  $a > 0$  and  $b > 0$

$$F : \mathcal{U} \times I \times H \rightarrow \mathbb{R}^2 (\|\cdot\|_{\ell_1} \text{ norm}).$$

By the Lipschitz condition in (A:1), we have

$$\begin{aligned} |F(x, \mathcal{U}, \mathcal{V}) - F(x, \tilde{\mathcal{U}}, \tilde{\mathcal{V}})| &\leq |\mathcal{U} - \tilde{\mathcal{U}}| + |\tilde{f}(x)| |\mathcal{V} - \tilde{\mathcal{V}}| \\ &\quad + \frac{2}{|\lambda|} |\tilde{q}(x)| |\mathcal{U} - \tilde{\mathcal{U}}| + |\tilde{r}(x)| |\mathcal{U} \log \mathcal{U} - \tilde{\mathcal{U}} \log \tilde{\mathcal{U}}|, \end{aligned} \quad (21)$$

we know that

$$\begin{aligned} |\mathcal{U} \log \mathcal{U} - \tilde{\mathcal{U}} \log \tilde{\mathcal{U}}| &= (\log(\xi) + 1) |\mathcal{U} - \tilde{\mathcal{U}}| \quad \text{for } \xi \in [a, b] \\ &\leq \max_{\xi \in [a, b]} (1 + \log \xi) |\mathcal{U} - \tilde{\mathcal{U}}|. \end{aligned} \quad (22)$$

For the equation

$$X = \begin{pmatrix} \mathcal{U} \\ \mathcal{V} \end{pmatrix}.$$

Therefore,

$$\begin{cases} \dot{X} = F(x, X), \\ X(x_0(\omega)) = X_0(\omega) \in I_0 \times H_0. \end{cases} \quad (23)$$

Hence the solution has been found to exist, with the terminal condition given by

$$\Phi(x_0(\omega), t) = \exp\left(-\frac{1}{\lambda(t)} \phi(x_0(\omega))\right). \quad (24)$$

In summary we have the following results.

**Theorem 4.3** Consider a special case for the equation

$$\begin{aligned} \beta \Phi(x(t, \omega)) \log \Phi(x(t, \omega)) &= -\frac{1}{\lambda(t)} \Phi(x(t, \omega)) q(x(t, \omega)) \\ &\quad + \Phi_x^T(x(t, \omega)) f(x(t, \omega)) \\ &\quad + \frac{1}{2} \text{Tr}[\Phi_{xx}(x(t, \omega)) \sigma(x(t, \omega)) \sigma^T(x(t, \omega))] \end{aligned}$$

for a one-dimensional problem and for

$$\Phi(x(t, \omega), t) = \Phi(x(t, \omega)).$$

Then, assuming that (A:1) and (A:2) hold, at least one solution has been found to exist.

The solution in (23) is not necessarily unique, and to attain uniqueness, more boundary conditions to the ODE must be given. For a one-dimensional problem at least one solution has been found to exist, and for  $n \geq 2$  the equation is a PDE which is difficult to solve.

#### 4.1 Iterative optimal control estimates

From Theorem 4.3, consider the estimated value function to be given as

$$\begin{aligned} \Phi(x(t_{j+1}, \omega)) &= \int_{\partial \Upsilon} \rho(\Upsilon | x_j) \exp\left(-\frac{1}{\lambda(T)} \phi(x(T, \omega))\right) \\ &\quad \times \exp\left(-\int_{t_j}^{t_{N-1}} (|\tilde{f}_j| + |\tilde{Q}_j|) dt\right) d\Upsilon, \end{aligned} \quad (25)$$

where

$$\Upsilon = (x(t_j, \omega), x(t_{j+1}, \omega), \dots, x(t_{N-1}, \omega))$$

and

$$d\Upsilon = (dx(t_j, \omega) dx(t_{j+1}, \omega) \cdots dx(t_{N-1}, \omega)).$$

The expectation of the value function is driven by stochastic differential equation (19). The function  $\rho(\Upsilon | x_j)$  in equation (25) is the probability density function of the transitions, and the function  $\tilde{Q}_j$  will be defined later.

Certainly, we cannot surely know future paths and the future control values due to the presence of the noise to the problem. This does not mean we have to give up since future paths cannot be certainly known, therefore we may estimate future paths, hence future control values, in order to attain optimums as the controls are dependent on the path control.

The continuous time interval is divided into small time intervals to attain small equal discrete paths assuming we are not distorting the trajectory in any way, that is, let

$$x_{j+1}(\omega) - x_j(\omega) = x(t_{j+1}, \omega) - x(t_j, \omega) \quad \text{for all } t_j \in [\epsilon, T - \epsilon] \text{ for } \epsilon \rightarrow 0.$$

Suppose the transition between the paths is given by

$$\begin{aligned} \rho(\Upsilon | x_j) &= \rho(x_{N-1}, \dots, x_{j+1} | x_j) \\ &= \prod_{j=0}^{N-1} \rho(x_{j+1} | x_j), \\ &\quad x_j \text{'s are identically independent and } j = 0 \text{ is the initial state.} \end{aligned} \quad (26)$$

The above equation is the cumulative probability density function for the sample path from  $x_j$  to  $x_{N-1}$ . The transitions of the sample paths are Markovian as they are solely dependent on the current path ( $x_j$ ) at time  $t_j$ . Following the work of [15] to the latter, we take the noise term to be Gaussian distributed with mean zero and variance  $\theta(x) = \sigma \sigma^T$  as given earlier. Therefore,

$$\rho(x_{j+1}|x_j) = \frac{\exp(-\frac{1}{2} \sum_{j=0}^{N-1} \|\frac{x_{j+1}(\omega) - x_j(\omega)}{\delta_j} - g_j(\omega)\|^2 \delta_j \theta_j^{-1}(\omega))}{\prod_{j=0}^{N-1} ((2\pi)|\theta_j(\omega)|)^{\frac{1}{2}}} \quad (27)$$

for  $\delta_j = t_{j+1} - t_j$ , which is the change in time  $t$ .

Hence we have the following results given as a lemma.

**Lemma 4.1** *From both Theorem 4.2 and Theorem 4.3, and assuming that the transitions are given by equation (27), we give the iterative optimal controls as*

$$u^*(\omega) = -R_j^{-1}(\omega)G^T(\omega) \frac{\exp(A_j(\omega) + B_j(\omega))}{\int_{\partial\Upsilon} \exp(A_j(\omega) + B_j(\omega)) d\Upsilon}$$

and

$$v^*(\omega) = -S_j^{-1}(\omega)H^T(\omega) \frac{\exp(A_j(\omega) + B_j(\omega))}{\int_{\partial\Upsilon} \exp(A_j(\omega) + B_j(\omega)) d\Upsilon}$$

for the estimated value function

$$\begin{aligned} \Phi(x(t_{j+1}, \omega)) &= \int_{\partial\Upsilon} \rho(\Upsilon|x_j) \exp\left(-\frac{1}{\lambda(T)}\phi(x(T, \omega))\right) \\ &\quad \times \exp\left(-\sum_{j=0}^{N-1} (|\tilde{f}_j| + |\tilde{Q}_j|) dt\right) d\Upsilon, \end{aligned}$$

where

$$A_j(\omega) = -\frac{1}{2} \sum_{j=0}^{N-1} \left\| \frac{x_{j+1}(\omega) - x_j(\omega)}{\delta_j} - g_j(\omega) \right\|^2 \delta_j \theta_j^{-1}(\omega)$$

and

$$B_j(\omega) = \sum_{j=0}^{N-1} (|\tilde{f}_j| + |\tilde{Q}_j|) \theta_j^{-1}(\omega) \delta_j.$$

*Proof* From Theorem 4.2, suppose that the solution is given as an estimated iterative value function in equation (25). Consider the discrete paths of the optimal trajectory given as

$$\begin{aligned} \Phi(x(t_{j+1}, \omega)) &= \int_{\partial\Upsilon} \rho(\Upsilon|x_j) \exp\left(-\frac{1}{\lambda(T)}\phi(x(T, \omega))\right) \\ &\quad \times \exp\left(-\int_{t_j}^{t_{N-1}} (|\tilde{f}_j| + |\tilde{Q}_j|) dt\right) d\Upsilon. \end{aligned} \quad (28)$$

Now, substituting equation (27) to equation (25), we have

$$\begin{aligned}\Phi(x(t_j, \omega), t) &= \int_{\partial \Upsilon} \frac{\exp(-\frac{1}{\lambda(T)}\phi(x(T, \omega)))}{\prod_{j=0}^{N-1} ((2\pi)|\theta_j|)^{\frac{1}{2}}} \\ &\quad \times \exp\left(-\frac{1}{2} \sum_{j=0}^{N-1} \frac{1}{\lambda_j} \left\| \frac{x_{j+1} - x_j}{\delta_j} - g_j \right\|^2 \delta_j \theta_j^{-1}\right) \\ &\quad \times \exp\left(\sum_{j=0}^{N-1} (|\tilde{f}_j| + |\tilde{Q}_j|) \delta_j\right) d\Upsilon\end{aligned}\quad (29)$$

for

$$|\tilde{Q}(x)| = A|r(x)| + \left(1 + \frac{2}{\lambda} |\tilde{q}(x)|\right),$$

where

$$C = \max_{\xi \in (a, b)} (1 + \log \xi).$$

We know that

$$\tilde{f}(x) = 2f(x)\theta^{-1}(x), \quad \tilde{q}(x) = q(x)\theta^{-1}(x)$$

and

$$r(x) = \beta\theta^{-1}(x).$$

Therefore, we have

$$\begin{aligned}\Phi(x(t_j, \omega)) &= \int_{\partial \Upsilon} \frac{\exp(-\frac{1}{\lambda(T)}\phi(x(T, \omega)))}{\prod_{j=0}^{N-1} ((2\pi)|\theta_j|)^{\frac{1}{2}}} \\ &\quad \times \left[ \exp\left(-\frac{1}{2} \sum_{j=0}^{N-1} \frac{1}{\lambda_j} \left\| \frac{x_{j+1} - x_j}{\delta_j} - g_j \right\|^2 \delta_j \theta_j^{-1}\right) \right] \\ &\quad \times \left[ \exp\left(\sum_{j=0}^{N-1} (|\tilde{f}_j| + |\tilde{Q}_j|) \delta_j\right) \right] d\Upsilon.\end{aligned}\quad (30)$$

Let

$$\begin{aligned}A_j(\omega) &= -\frac{1}{2} \sum_{j=0}^{N-1} \left\| \frac{x_{j+1}(\omega) - x_j(\omega)}{\delta_j} - g_j(\omega) \right\|^2 \delta_j \theta_j^{-1}(\omega), \\ B_j(\omega) &= \sum_{j=0}^{N-1} (|\tilde{f}_j| + |\tilde{Q}_j|) \theta_j^{-1}(\omega) \delta_j\end{aligned}$$

and

$$f_j(\omega) = f(x(t_j, \omega)).$$

Similarly, extending that representation to other functions, we may express the iterative value of the game as

$$\Phi_j(\omega) = \int_{\partial\Upsilon} \frac{\exp(-\frac{1}{\lambda(T)}\phi(x(T, \omega)))}{\prod_{j=0}^{N-1} ((2\pi)|\theta_j|)^{\frac{1}{2}}} [\exp(A_j(\omega) + B_j(\omega))] d\Upsilon. \quad (31)$$

Applying Theorem 4.2, we obtain the optimal iterative control estimates, which completes the proof.  $\square$

## 5 Conclusion and future work

The paper studied the nonlinear stochastic control problem of a zero sum differential game. The problem studied here is similar to the one in [15], the difference is that ours is a minimax problem and the state function is dependent on an additional random variable  $\omega$ , which was suggested as future work. Equating the control weights to the variance of the Brownian noise played an important role towards the solution. A quasilinear PDE was attained due to the logarithmic transformation of the value function, though the equation was difficult to solve. However, the quasilinear differential equation was converted to an ODE by fixing the time variable  $t$ , which showed that at least a solution (saddle point) does exist. The uniqueness of the solution can be attained if more boundary conditions are set to the problem. The problem is path-dependent and Gaussian distributed with mean zero and variance  $\theta$ . Therefore, the Gaussian distribution has played a vital role in obtaining the control path estimates as given in Lemma 4.1. We are of the view that nonlinear Feynman-Kac formula can be of great help in attaining the solution for the nonlinear PDE [23]. Lastly, note that for  $n \geq 2$ , the problem is a nonlinear PDE, which still remains an open problem according to our understanding.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

The mathematical derivation, development of theorems and the writing up of the article is done by the first author. The other author helped in mathematical derivations and cross checking of the errors within the paper. All authors read and approved the final manuscript.

### Acknowledgements

We would like to take this opportunity to thank the unknown referees and the Chinese Scholarship Council for giving us an opportunity to further our studies on this research area. Our sincere acknowledgements are also extended to our supervisor, families, colleagues for moral support and guidance.

Received: 9 April 2014 Accepted: 24 September 2014 Published: 14 Oct 2014

### References

1. Hellerstein, JL, Morrison, V, Eilebrecht, E: Applying control theory in the real world: experience with building a controller for the .NET thread pool. *ACM SIGMETRICS Perform. Eval. Rev.* **37** 38-42 (2010)
2. Kendrick, D: Applications of control theory to macroeconomics. *Ann. Econ. Soc. Meas.* **5**, 171-190 (1976)
3. Kilian, C: *Modern Control Technology*. Thomson Delmar Learning, Clifton Park (2005)
4. Kappen, HJ: *Stochastic optimal control theory*. ICML, Helsinki, Radboud University, Nijmegen, Netherlands (2008)
5. Mazliak, L: An introduction to probabilistic methods In *stochastic control*. Laboratory Probabilities, University of Paris, France (1996)
6. Yong, J, Zhou, X: *Stochastic Controls. Hamiltonian Systems and HJB Equations*. Springer, Berlin (1999)
7. Vorob'ev, NH: *Game Theory*. Springer, Berlin (1977)
8. Tirole, J, Fudenberg, D: *Game Theory*. MIT Press, Cambridge (1991)
9. Basar, T: *Lecture notes on non-cooperative game theory*. University of Illinois, Urbana (2010)
10. Ho, YC, Bryson, AE, Baron, S: Differential games and optimal pursuit-evasion strategies. *IEEE Trans. Autom. Control* **10**, 385-389 (1965)
11. Yeung, DWK: A feedback Nash equilibrium solution for noncooperative innovations in a stochastic differential game framework. *Stoch. Anal. Appl.* **9**, 195-213 (1991)

12. Browne, S: Stochastic differential portfolio games. *J. Appl. Probab.* **37**, 126-147 (2000)
13. Theodorou, EA, Todorov, E: Stochastic optimal control for nonlinear Markov jump diffusion processes. PhD thesis, University of Southern California (2011)
14. Fleming, WH, Soner, HM: *Controlled Markov Processes and Viscosity Solutions: Applications of Mathematics*, 2nd edn. Springer, New York (2006)
15. Theodorou, EA: Iterative path integral stochastic optimal theory: theory and applications to motor control. Thesis, University of Southern California (2011)
16. Guo, X, Liu, J, Zhou, XY: A constrained non-linear regular-singular stochastic control problem, with applications. *Stoch. Process. Appl.* **109**, 167-187 (2004)
17. Oksendal, BK: *Stochastic Differential Equations: An Introduction with Applications*, 6th edn. Springer, Berlin (2003)
18. Josa-Fombellida, R, Rincon-Zapatero, JP: New approach to stochastic optimal control. *J. Optim. Theory Appl.* **135**, 163-177 (2007)
19. Silva, F: Interior penalty approximation for optimal control problems. Optimality conditions in stochastic optimal control theory. CMAP, Ecole Polytechnique, and INRIA Saclay (2010)
20. Mou, L, Yong, J: Two person zero sum linear quadratic stochastic differential games by Hilbert space method. *J. Ind. Manag. Optim.* **2**(1), 93-115 (2006)
21. Rigobon, R: *Brownian Motion and Stochastic Calculus Introductory Notes*. MIT Press, Cambridge (2009)
22. Chow, GC: Optimal control without solving the Bellman equation. *Econometric Research Program. Research Memo No. 364*, Princeton University (1992)
23. Peng, S, Wang, F: BSDE, path-dependent PDE and nonlinear Feynman-Kac formula. School of Mathematics, Shandong University, Jinan, China (2011)

10.1186/1687-1847-2014-266

**Cite this article as:** Basimanebotlhe and Xue: Stochastic optimal control to a nonlinear differential game. *Advances in Difference Equations* 2014, **2014**:266

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)