# Approximate Infinite-Horizon Optimal Control for Stochastic Systems

Giordano Scarciotti and Thulasi Mylvaganam

***Abstract*— The policy of an optimal control problem for nonlinear stochastic systems can be characterized by a second-order partial differential equation for which solutions are not readily available. In this paper we provide a systematic method for obtaining approximate solutions for the infinite-horizon optimal control problem in the stochastic framework. The method is demonstrated on an illustrative numerical example in which the control effort is not weighted, showing that the technique is able to deal with one of the most striking features of stochastic optimal control.**

## I. Introduction

Given a dynamical system it is often of interest to design control laws to achieve a certain objective *in an optimal manner*: optimal control concerns precisely these scenarios. The objectives are described by a cost functional which is to be minimized via the control inputs subject to the system dynamics. The problem has been studied extensively and relies upon two families of techniques, namely minimum principle methods and dynamic programming methods, see *e.g.* [1], [2], [3], [4], [5], [6], [7]. Optimal control problems can be formulated for different classes of systems, such as deterministic systems and stochastic systems. In particular, optimal control in the stochastic framework has received a large amount of attention because of the several economical and financial applications for which it can be used. Successful applications include the production planning problem [8], the investment versus consumption problem [9], [10], the technology diffusion problem [11], [12] and the optimal stopping problem [13]. Yet, several open problems remain, see *e.g.* [14]. It is well-known that, using the dynamic programming approach, the solution of a general nonlinear optimal control problem relies upon the solution of the Hamilton-Jacobi-Bellman (HJB) partial differential equation which is not easy to obtain, see *e.g.* [14]. In particular, in the stochastic framework the HJB is a second-order partial differential equation which is even more complicated to solve.

In this paper, building upon the concepts introduced in [15], [16] and extended to differential games [17], [18], [19], collision avoidance [20], [21], [22] and constrained optimal control [23], we propose a systematic method for constructing approximate solutions for the infinite-horizon optimal control problem for nonlinear stochastic systems. The method requires the solution of an *algebraic* equation in place of the HJB *partial differential* equation. Moreover, the level of approximation can be interpreted as an *additional cost* which is exactly quantifiable. Differently from [15] in which deterministic systems are studied, a stochastic framework is considered herein. Note that there are several important differences between the deterministic and the stochastic settings. One of the most interesting features of the stochastic framework is that it is possible to consider problems in which the control is not weighted or even carries a negative weight (*i.e.* increasing the control decreases the cost) an occurrence which does not make sense in the deterministic case, see [24], [14]. The proposed method is clearly able to deal with this type of problems, a capability which we demonstrate with a numerical example in which the control is not weighted.

The remainder of the paper is organized as follows. In Section II we define the optimal control problem for stochastic nonlinear systems, we present its exact solution and we formulate the approximate problem which we want to solve. In Section III we introduce a series of tools which facilitate the exposition of the main result. In Section IV we solve the approximate problem and we provide a policy which approximates the optimal control law of the original problem. In Section V we demonstrate the results of the paper with an example. Finally, Section VI contains some concluding remarks.

**Notation.** We use standard notation. $\mathbb{R}_{\geq 0}$ ($\mathbb{R}_{>0}$) denotes the set of non-negative (positive) real numbers. Similarly, $\mathbb{R}^{n\times n}_{>0}$ denotes the set of positive definite matrices of dimension $n$. The superscript $\top$ denotes the transposition operator. The symbol $E$ denotes the expected value operator. $(\nabla, \mathcal{A}, \mathcal{P})$ indicates a probability space with a given set $\nabla$, a $\sigma$-algebra $\mathcal{A}$ on $\nabla$ and a probability measure $\mathcal{P}$ on the measurable space $(\nabla, \mathcal{A})$.

## II. Problem formulation

In this section we formulate the full and approximate optimal control problem for stochastic nonlinear systems.

Let $\mathcal{W}_t$ be a standard Wiener process defined on a probability space $(\nabla, \mathcal{A}, \mathcal{P})$. A stochastic process $x_t$ is a function of two variables such that for each $t \in \mathbb{R}$, $x(t, \cdot)$ is a random variable and for each $w \in \nabla$, $x(\cdot, w)$ is called path of $x$. For ease of notation, we indicate the paths as just functions of $t$, *e.g.* the path of $x_t$ as $x : t \mapsto x(t)$ and the path of $\mathcal{W}_t$ as $\mathcal{W} : t \mapsto \mathcal{W}(t)$ (this is common in the literature, see *e.g.* [25]). Consider a stochastic nonlinear input-affine system described by the equations

$$\begin{aligned} dx &= [f(x) + g(x)u]dt + [h(x) + l(x)u]d\mathcal{W}, \\ x(0) &= x_0, \end{aligned} \tag{1}$$

with $x(t) \in \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$. Assume that $f : \mathbb{R}^n \to \mathbb{R}^n$, $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$, $h : \mathbb{R}^n \to \mathbb{R}^n$ and

G. Scarciotti is with the Dept. of Electrical and Electronic Engineering, Imperial College London, London, SW7 2AZ, UK, E-mail: g.scarciotti@ic.ac.uk.

T. Mylvaganam is with the Dept. of Aeronautics, Imperial College London, London, SW7 2AZ, UK, E-mail: t.mylvaganam@ic.ac.uk.

$l : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are smooth mappings and that the initial condition $x_0$ is independent of $\mathcal{W}(t)$ for all $t > 0$. Under these assumptions the initial value problem associated with (1) has a unique solution, see *e.g.* [25]. Moreover, assume that the origin is an equilibrium point of (1), *i.e.* $f(0) = 0$ and $h(0) = 0$. Since the functions are smooth this implies that there exist continuous, possibly non-unique, functions $F : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ and $H : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ such that $f(x) = F(x)x$ and $h(x) = H(x)x$ for all $x \in \mathbb{R}^n$. The full infinite-horizon optimal control problem can be formulated as follows.

**Problem 1.** The *infinite-horizon optimal control problem* consists in finding a control $u$ which minimizes the cost functional

$$J(x(0), u) = E\left\{\int_0^\infty \frac{1}{2}\left[q(x(t)) + r(x(t), u(t))\right] dt\right\}, \quad (2)$$

where $q = x^\top Q(x)x$, with $Q = Q^\top$ such that $Q : \mathbb{R}^n \to \mathbb{R}^{n\times n}_{>0}$, and $r = u^\top R(x)u$, with $R = R^\top$ such that $R : \mathbb{R}^n \to \mathbb{R}^{m\times m}$, subject to the constraint (1). Moreover, $u$ needs to be such that the zero equilibrium of the closed-loop system is asymptotically stable almost surely.

The solution to this problem is well known, see *e.g.* [14]. The optimal control is given by

$$u^* = -\left(R + l^\top V_{xx} l\right)^{-1}\left(g^\top V_x^\top + l^\top V_{xx} h\right). \quad (3)$$

where $V$, which is called value function, satisfies the equation

$$\begin{aligned}0 &= h^\top V_{xx} h + 2V_x f + q \\ &\quad - \left(V_x g + h^\top V_{xx} l\right)\left(R + l^\top V_{xx} l\right)^{-1}\left(g^\top V_x^\top + l^\top V_{xx} h\right).\end{aligned} \quad (4)$$

If we are able to find a function $V$ satisfying the second-order partial differential equation (4), then the control (3) solves Problem 1. A necessary condition for the solution of the infinite-horizon optimal control problem is that $R(x) + l(x)^\top V_{xx}(x) l(x)$ is invertible at least for every point of the optimal trajectory of (1) generated under $u^*$. While this is guaranteed if $R$ is positive definite (similarly to the deterministic case), as already pointed out the positive definiteness of $R$ is not required in general [14].

The main drawback of this approach is that solving the partial differential equation (4) is a difficult problem and very often closed-form solutions are not readily available. Hence, in this paper we provide a systematic technique to approximate the solution of the infinite-horizon optimal control problem. In the following we precisely quantify the entity of this approximation.

**Definition 1.** Consider system (1). The function $u$ in (1) is a *dynamic control law* if for some $\nu > 0$ it is described by the equations

$$\begin{aligned}d\xi &= \alpha(x,\xi)dt + \beta(x,\xi)d\mathcal{W}, \\ u &= \gamma(x,\xi),\end{aligned} \quad (5)$$

where $\xi \in \mathbb{R}^\nu$ is the state of the *dynamic extension*, $\alpha : \mathbb{R}^n \times \mathbb{R}^\nu \to \mathbb{R}^\nu$, $\beta : \mathbb{R}^n \times \mathbb{R}^\nu \to \mathbb{R}^\nu$ and $\gamma : \mathbb{R}^n \times \mathbb{R}^\nu \to \mathbb{R}^m$ are smooth mappings, with $\alpha(0,0) = 0$, $\beta(0,0) = 0$ and $\gamma(0,0) = 0$.

**Problem 2.** The *approximate regional dynamic infinite-horizon optimal control* problem consists in determining a stochastic *dynamic* control law (5) and a *region* $\Omega \subset \mathbb{R}^n \times \mathbb{R}^\nu$ containing the origin of $\mathbb{R}^n \times \mathbb{R}^\nu$ such that the cost functional

$$\begin{aligned}&\bar{J}(x(0), \xi(0), u) = \\ &= E\left\{\int_0^\infty \left[\frac{1}{2}(q(x(t)) + r(x(t), u(t))) + c(x(t), \xi(t))\right] dt\right\},\end{aligned} \quad (6)$$

with $c : \mathbb{R}^n \times \mathbb{R}^\nu \to \mathbb{R}_{\geq 0}$, is minimized in the sense of

$$\bar{J}(x(0), \xi(0), \gamma) \leq \bar{J}(x(0), \xi(0), \bar{u}),$$

for all $\bar{u}$ and for all $(x(0), \xi(0)) \in \Omega$ such that the closed-loop trajectories remain in $\Omega$ and converge to zero almost surely, *i.e.* the zero equilibrium is asymptotically stable almost surely.

Note that a solution of Problem 2 is a *local* solution of Problem 1 with respect to a modified cost functional. In particular, the original cost is modified by the *additional running cost* $c$. Clearly Problem 2 recovers Problem 1 if $\Omega = \mathbb{R}^n \times \mathbb{R}^\nu$ and $c = 0$.

## III. Useful tools: linear solution, algebraic solution and extended value function

To provide a systematic method for solving Problem 2, we first recall the solution of the linearized problem. System (1) is linearized as

$$dx = [Ax + Bu]dt + [Cx + Du]d\mathcal{W}, \quad (7)$$

with

$$A = \left.\frac{\partial f}{\partial x}\right|_{x=0}, \quad B = g(0), \quad C = \left.\frac{\partial h}{\partial x}\right|_{x=0}, \quad D = l(0),$$

and $\bar{Q} = Q(0)$ and $\bar{R} = R(0)$. The optimal control associated to the linearized problem is given by (see *e.g.* [14, Theorem 6.1])

$$u^* = -\left(\bar{R} + D^\top \bar{P} D\right)^{-1}\left(B^\top \bar{P}^\top + D^\top \bar{P} C\right)x, \quad (8)$$

where $\bar{P}$ is the solution of the stochastic algebraic Riccati equation

$$\begin{aligned}0 &= A^\top \bar{P} + \bar{P}A + C^\top \bar{P} C + \bar{Q} \\ &\quad -(\bar{P}B + C^\top \bar{P} D)(\bar{R} + D^\top \bar{P} D)^{-1}(B^\top \bar{P} + D^\top \bar{P} C), \\ &\bar{R} + D^\top \bar{P} D > 0.\end{aligned} \quad (9)$$

The stochastic algebraic Riccati equation can be solved numerically as shown in, *e.g.*, [26].

In what follows the structure of the linearized problem is modified to yield a solution of Problem 2. This is done through the definition of algebraic solution (see [15] for the deterministic notion). This forms the basis of the design of a dynamic control law which solves Problem 2.

**Definition 2.** Consider system (1) and the cost functional (2). Let $\Sigma : \mathbb{R}^n \to \mathbb{R}^{n\times n}_{\geq 0}$, with $\Sigma(x) = \Sigma(x)^\top$ and $\bar{\Sigma} = \Sigma(0)$. Assume that the system[1]

$$\begin{aligned} 0 &= F^\top P + PF + H^\top PH + Q + \Sigma \\ &\quad -(Pg + H^\top Pl)(R + l^\top Pl)^{-1}(g^\top P + l^\top PH), \qquad (10) \\ R &+ l^\top Pl > 0, \end{aligned}$$

has a global solution $P : \mathbb{R}^n \to \mathbb{R}^{n\times n}$, with $P(x) = P(x)^\top$, such that $P(0) = \bar{P}$, where $\bar{P}$ is the solution of the stochastic algebraic Riccati equation (9) with $\bar{Q}$ replaced by $\bar{Q} + \bar{\Sigma}$. The solution $P$ is said to be an *algebraic $\bar{P}$ solution*.

In the following let $\nu = n$ and assume the existence of an algebraic $\bar{P}$ solution. Exploiting the algebraic $\bar{P}$ solution we can define the function

$$\mathcal{V}(x,\xi) = \frac{1}{2}x^\top P(\xi)x + \frac{1}{2}||x-\xi||^2_W, \qquad (11)$$

with $\xi(t) \in \mathbb{R}^n$ and $W = W^\top \in \mathbb{R}^{n\times n}_{>0}$.

In the next section we show constructively under which conditions it is possible to design the dynamics of $\xi$ such that function (11), which we call the *extended value function*, is the value function corresponding to Problem 2.

## IV. Design of the dynamic extension

To streamline the presentation of the result of this section, we introduce some preliminary definitions. Let $\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be a continuous mapping such that

$$x^\top P(x) - x^\top P(\xi) = (x-\xi)^\top \Phi(x,\xi)^\top,$$

and $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be the Jacobian matrix defined as

$$\Psi(x,\xi) = \frac{1}{2}\frac{\partial}{\partial \xi}\left(P(\xi)x\right).$$

Define

$$\Delta(x,\xi) = (W - \Phi(x,\xi))W^{-1}\Psi(x,\xi)^\top.$$

Let $\Pi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be a continuous mapping such that

$$\Pi(x,\xi) = P(x) - P(\xi) - W,$$

and let $\mathcal{R} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be defined as

$$\mathcal{R}(x,\xi) = R + l^\top \mathcal{V}_{xx}\, l.$$

Finally, define

$$\begin{aligned} \mathcal{HJB} &= q + h^\top \mathcal{V}_{xx}\, h + \beta^\top \mathcal{V}_{\xi\xi}\, \beta + 2\,\mathcal{V}_x\, f + 2\,\mathcal{V}_\xi\, \alpha \\ &\quad - \left(\mathcal{V}_x\, g + h^\top \mathcal{V}_{xx}\, l\right)\left(R + l^\top \mathcal{V}_{xx}\, l\right)^{-1}\left(g^\top \mathcal{V}_x^\top + l^\top \mathcal{V}_{xx}\, h\right), \end{aligned} \qquad (12)$$

and

$$F_{cl} = F - g\mathcal{R}^{-1}(g^\top P(x) + l^\top \mathcal{V}_{xx}\, H).$$

[1]For brevity, the arguments of the mappings, including $F(x)$, $H(x)$, $g(x)$, $l(x)$, $P(x)$, $Q(x)$ and $R(X)$, are omitted when the arguments are clear from the context.

We are now ready to solve Problem 2. Recall that the arguments of mappings are omitted when they are clear from the context, for instance $P$ denotes $P(x)$.

**Proposition 1.** Consider the cost functional (6) and the interconnection of system (1) with the control law (5). Assume that $P$ is an algebraic $\bar{P}$ solution of (10) with $\Sigma$ selected such that

$$\Delta^\top F_{cl} + F_{cl}\Delta + Y_2 - H^\top \Pi H < \Sigma + Y_1 + \Delta^\top g\mathcal{R}^{-1}g^\top \Delta, \qquad (13)$$

where

$$\begin{aligned} Y_1(x) &= (Pg + H^\top \mathcal{V}_{xx}\, l)\mathcal{R}^{-1}(g^\top P + l^\top \mathcal{V}_{xx}\, H) \geq 0, \\ Y_2(x) &= (Pg + H^\top Pl)\left(R + l^\top Pl\right)^{-1}(g^\top P + l^\top PH) \geq 0, \end{aligned} \qquad (14)$$

for all $(x,\xi) \in \Omega$. Then there exists $\bar{k} \geq 0$ such that for all $k > \bar{k}$ the function $\mathcal{V}$ defined in (11) satisfies $\mathcal{HJB} \leq 0$ for all $(x,\xi) \in \Omega$ with

$$\begin{aligned} \dot{\xi} &= -k\,\mathcal{V}_\xi^\top, \\ u &= -\mathcal{R}^{-1}\left(g^\top \mathcal{V}_x^\top + l^\top \mathcal{V}_{xx}\, h\right). \end{aligned} \qquad (15)$$

Moreover, Problem 2 is solved with the additional running cost $c = -\,\mathcal{HJB} \geq 0$.

## V. Example

In this section we illustrate the results of the paper with a numerical example. Consider the nonlinear system

$$\begin{aligned} dx_1 &= x_2 dt, \\ dx_2 &= (\eta_1 x_1 x_2 + u)dt + \eta_2 u\, d\mathcal{W}, \end{aligned} \qquad (16)$$

with $x(t) = \begin{bmatrix} x_1 & x_2 \end{bmatrix}^\top \in \mathbb{R}^2$, $u(t) \in \mathbb{R}$, $\eta_1 \in \mathbb{R}\setminus\{0\}$, $\eta_2 \in (0,1)$ and the cost

$$J(x(0),u) = E\left\{\int_0^\infty \frac{1}{2}x(t)^\top x(t)dt\right\}. \qquad (17)$$

Note that by varying the parameter $\eta_1$ we can change the entity of the nonlinearity and by varying the parameter $\eta_2$ we can change the entity of the noise. Note that the control $u$ is not weighted in the cost functional (17) but, as long as $\eta_2 \neq 0$, the control is bounded [14]. Let

$$P(x) = \begin{bmatrix} p_{11}(x) & p_{12}(x) \\ p_{21}(x) & p_{22}(x) \end{bmatrix}$$

be an algebraic $\bar{P}$ solution of (10). Note that multiple choices for $F(x)$ are available. The selection

$$F(x) = \begin{bmatrix} 0 & 1 \\ \eta_1 x_2 & 0 \end{bmatrix}, \quad \Sigma(x) = \bar{\Sigma} = \begin{bmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{bmatrix},$$

yields[2]

$$p_{12} = \frac{-\eta_2^4(\eta_1 x_2 s_{22} + s_{11}) \pm \eta_2^4\sqrt{(\eta_1 s_{22} x_2 - s_{11})^2 + \dfrac{s_{11}s_{22}}{\eta_2^4}}}{4\eta_1\eta_2^4 x_2 - 1},$$

[2]In this example the algebraic solution $P(x)$ was found analytically using Maple.

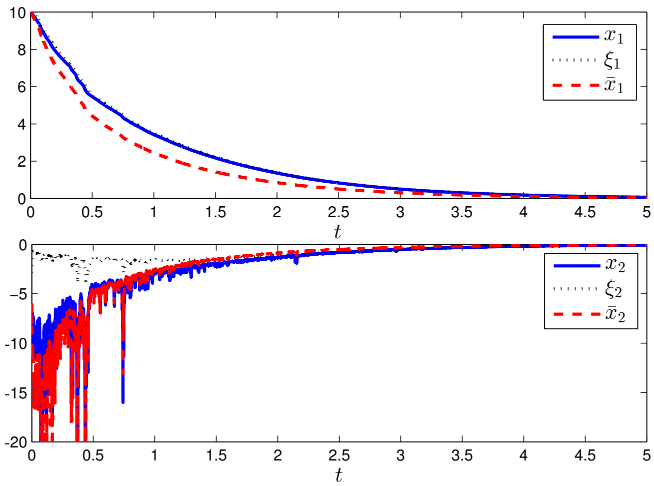


Fig. 1. Time histories for one specific path of $x$ (solid/blue line), of $\xi$ (dotted/black line) and of $\bar{x}$ (dashed/red line). First component in the top graph and second component in the bottom graph.

$$p_{22} = \eta_2^2(2p_{12} + s_{22}), \quad p_{11} = -\eta_1 x_2 p_{22} + \frac{p_{12}}{\eta_2^2}, \quad p_{21} = p_{12},$$

where $s_{11} = \sigma_{11} + 1$ and $s_{22} = \sigma_{22} + 1$. We select the solution with the negative square root in $p_{12}$ because it is locally (around the origin) positive definite (the other solution is locally negative definite). We can now easily determine the dynamic control law (15). With some abuse of notation, we indicate the cost (17) computed using the dynamic control law as $J(u_d)$. We also determine the optimal control of the linearized system, namely

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; C = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \; D = \begin{bmatrix} 0 \\ \eta_2 \end{bmatrix},$$

which is $u_l = -\left(\bar{R} + D^\top \bar{P} D\right)^{-1}\left(B^\top \bar{P}\right)\bar{x}$, where $\bar{P}$ solves equation (9). Equivalently, $\bar{P}$ can be computed substituting $x_2 = 0$, $\sigma_{11} = 0$ and $\sigma_{22} = 0$ in $p_{11}$, $p_{12}$ and $p_{22}$. With some abuse of notation, we indicate the cost (17) computed using the control law $u_l$ as $J(u_l)$ and the state generated by this policy as $\bar{x}$.

All numerical simulations are performed in MATLAB using Euler integration with step size 0.005. To render the simulations reproducible, we have used the command *rng('default')* which sets the random generator of MATLAB to the Mersenne Twister with seed zero. We select the initial conditions of system (17) as $x_1(0) = 10$ and $x_2(0) = -6$, and the matrix $W$ as $W = 0.5I$. The initial conditions of the dynamic extension have been selected by minimizing $\mathcal{V}(x, \xi)$ with respect to $(\xi_1, \xi_2)$ (note that $P(\xi)$ is independent of $\xi_1(0)$, so $\xi_1(0) = x_1(0)$ always). Finally, we have selected $k = 100$.

In the first simulation we select $\eta_1 = 5$ and $\eta_2 = 0.1$ (similar results have been obtained with other choices of the parameters, as shown below.). Fig. 1 shows, for one specific path (the first generated by MATLAB), the time histories of $x$ (top graph: $x_1$, bottom graph: $x_2$) depicted with a solid/blue line, the time histories of $\xi$ (top graph: $\xi_1$, bottom graph: $\xi_2$) depicted with a dotted/black line and

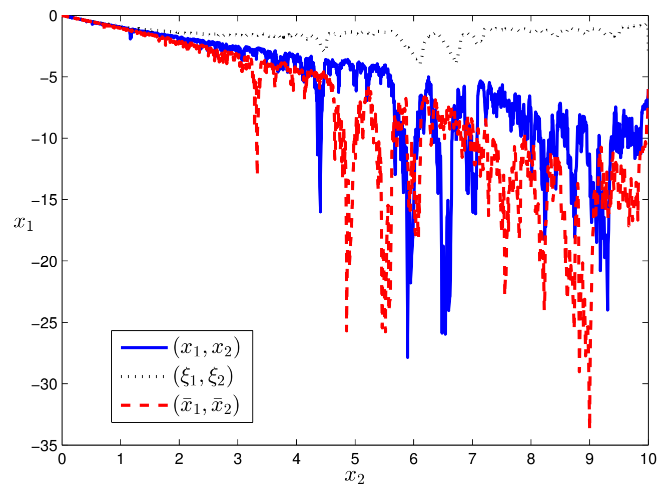


Fig. 2. Phase portrait, for the same path in Fig. 1, of $x$ (solid/blue line), of $\xi$ (dotted/black line) and of $\bar{x}$ (dashed/red line).

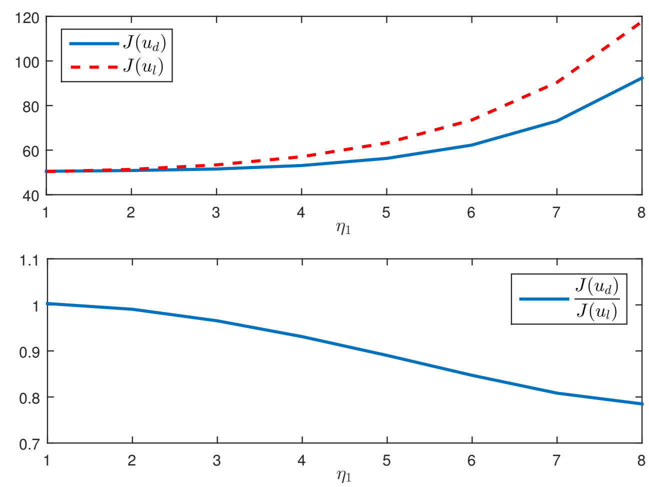


Fig. 3. Top graph: cost $J(u_d)$ (solid/blue line) and cost $J(u_l)$ (dashed/red line) for $\eta_1 \in [1, 8]$. Bottom graph: ratio $\dfrac{J(u_d)}{J(u_l)}$ for $\eta_1 \in [1, 8]$.

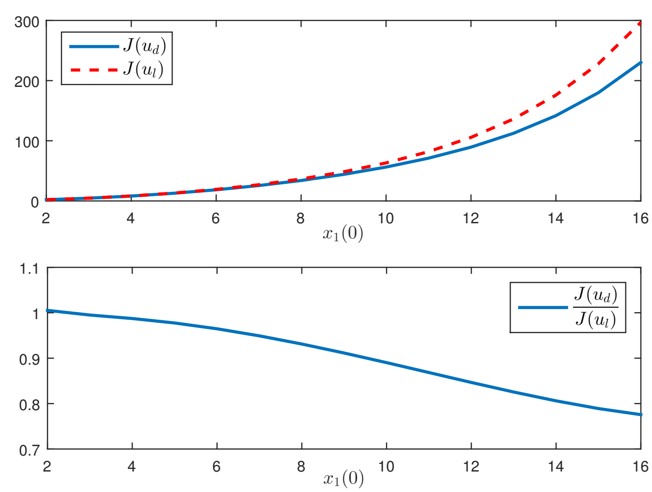


Fig. 4. Top graph: cost $J(u_d)$ (solid/blue line) and cost $J(u_l)$ (dashed/red line) for $x_1(0) \in [2, 16]$. Bottom graph: ratio $\dfrac{J(u_d)}{J(u_l)}$ for $x_1(0) \in [2, 16]$.

the time histories of $\bar{x}$ (top graph: $\bar{x}_1$, bottom graph: $\bar{x}_2$) depicted with a dashed/red line. Fig. 2 shows, for the same path, the phase portrait of $x$ (solid/blue line), phase portrait of $\xi$ (dotted/black line) and the phase portrait of $\bar{x}$ (dashed/red line). We note that both control laws are able to stabilize the system and that the generated trajectories are different, but similar. The simulation is repeated 20 times and the average costs are computed for the two control laws, yielding $J(u_d) = 56.297$ and $J(u_l) = 63.236$ resulting in the ratio $\dfrac{J(u_d)}{J(u_l)} = 0.890$.

Note that a larger value of $\eta_1$ increases the effect of the nonlinearity. Thus, we expect that increasing the value of $\eta_1$ the relative performance of the dynamic control law with respect to the performance of the linearized control law will improve. To this end, we simulate system (17) with $\eta_1 \in [1, 8]$. The average cost is computed over 20 simulations for each scenario. Fig. 3 (top graph) shows the cost $J(u_d)$ in solid/blue line and the cost $J(u_l)$ in dashed/red line for $\eta_1 \in [1, 8]$. The bottom graph shows the ratio $\dfrac{J(u_d)}{J(u_l)}$ for the same values of $\eta_1$. We see that, as expected, for larger values of $\eta_1$ the performance of the dynamic control law is increasingly better than the performance of the linearized control law. Note that a similar effect is caused, for a fixed value of $\eta_1$, also by taking initial conditions further away from the origin. This is shown in Fig. 4 for $x_1(0) \in [2, 16]$.

## VI. Conclusion

In this paper we have addressed the problem of optimal control for stochastic nonlinear systems. We have proposed a method to determine approximate solutions for the infinite-horizon optimal control problem. In particular, we have formulated and solved a relaxed problem which is a local version of the original problem and in which the cost functional has an additional running cost. The proposed technique has two main advantages. First, the difference between the two problems is precisely quantified and it can, in principle, be minimized. Second, the method does not require the solution of any partial differential equation, which is usually the main drawback of the family of methods based on dynamic programming. The proposed technique is illustrated with a numerical example in which the control policy is not weighted in the cost, showing that the method can deal with this class of problems which are peculiar of the stochastic framework.

## References


[1] R. E. Bellman, *Dynamic programming*. Princeton, NJ, USA: Princeton University Press, 1957.

[2] M. G. Crandall and P. Lions, “Viscosity solutions of Hamilton-Jacobi equations,” *Transactions of the American Mathematical Society*, vol. 277, no. 1, pp. 1–42, 1983.

[3] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, 3rd ed., vol. First.

[4] A. E. Bryson and Y. Ho, *Applied Optimal Control: Optimization, Estimation and Control*. Taylor & Francis, 1975.

[5] L. S. Pontryagin, V. G. Boltyansky, R. V. Gamkrelidze, and E. Mishchenko, *The Mathematical Theory of Optimal Processes*. New York: Wiley, 1962.

[6] R. E. Kalman, “Contributions to the theory of optimal control,” *Boletin de la Sociedad Matematica Mexicana*, vol. 5, pp. 102–119, 1960.

[7] R. Vinter, *Optimal Control*, ser. Modern Birkhäuser Classics. Birkhäuser Boston, 2010.

[8] F. Modigliani and F. E. Hohn, “Production planning over time and the nature of the expectation and planning horizon,” *Econometrica*, vol. 23, no. 1, pp. 46–66, 1955.

[9] R. C. Merton, “Lifetime portfolio selection under uncertainty: The continuous-time case,” *The review of Economics and Statistics*, vol. 51, no. 3, pp. 247–257, 1969.

[10] ——, “Optimum consumption and portfolio rules in a continuous-time model,” *Journal of economic theory*, vol. 3, no. 4, pp. 373–413, 1971.

[11] J. F. Reinganum, “On the diffusion of new technology: A game theoretic approach,” *The Review of Economic Studies*, vol. 48, no. 3, pp. 395–405, 1981.

[12] K. Roberts and M. L. Weitzman, “Funding criteria for research, development, and exploration projects,” *Econometrica: Journal of the Econometric Society*, pp. 1261–1288, 1981.

[13] H. Chernoff, “Optimal stochastic control,” *Sankhyā: The Indian Journal of Statistics, Series A*, vol. 30, no. 3, pp. 221–252, 1968.

[14] J. Yong and X. Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, ser. Stochastic Modelling and Applied Probability. Springer New York, 1999.

[15] M. Sassano and A. Astolfi, “Dynamic approximate solutions of the HJ inequality and of the HJB equation for input-affine nonlinear systems,” *IEEE Transactions on Automatic Control*, vol. 57, no. 10, pp. 2490–2503, 2012.

[16] ——, “Approximate finite-horizon optimal control without PDE’s,” *Systems & Control Letters*, vol. 62, no. 2, pp. 97–103, 2013.

[17] T. Mylvaganam, M. Sassano, and A. Astolfi, “Approximate solutions to a class of nonlinear differential games,” in *Proceedings of the 51th IEEE Conference on Decision and Control*, 2012, pp. 6489 – 6494.

[18] ——, “Approximate solutions to a class of nonlinear differential games using a shared dynamic extension,” in *Proceedings of the 12th European Control Conference*, 2013, pp. 710–715.

[19] ——, “Constructive $\epsilon$-Nash equilibria for nonzero-sum differential games,” *IEEE Transactions on Automatic Control*, vol. 60, no. 4, pp. 950–965, 2015.

[20] ——, “A constructive differential game approach to collision avoidance in multi-agent systems,” in *Proceedings of the 2014 American Control Conference*, 2014.

[21] ——, “A differential game approach to multi-agent collision avoidance,” *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 4229–4235, Aug 2017.

[22] T. Mylvaganam and M. Sassano, “Autonomous collision avoidance for wheeled mobile robots using a differential game approach,” *European Journal of Control*, vol. 40, pp. 53–61, 2018.

[23] G. Scarciotti and A. Astolfi, “Approximate finite-horizon optimal control for input-affine nonlinear systems with input constraints,” *Journal of Control and Decision*, vol. 1, no. 2, pp. 149–165, 2014.

[24] S. Chen, X. Li, and X. Y. Zhou, “Stochastic linear quadratic regulators with indefinite control weight costs,” *SIAM Journal on Control and Optimization*, vol. 36, no. 5, pp. 1685–1702, 1998.

[25] T. C. Gard, *Introduction to Stochastic Differential Equations*, ser. Monographs and textbooks in pure and applied mathematics. M. Dekker, 1988.

[26] M. A. Rami and X. Y. Zhou, “Linear matrix inequalities, Riccati equations, and indefinite stochastic linear quadratic controls,” *IEEE Transactions on Automatic Control*, vol. 45, no. 6, pp. 1131–1143, 2000.