2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Approximate Finite-Horizon Optimal Control without PDE's

M. Sassano and A. Astolfi

***Abstract*— The problem of controlling the state of a system, from a given initial condition, during a fixed time interval minimizing at the same time a criterion of optimality is commonly referred to as *finite-horizon optimal control* problem. It is well-known that one of the standard solutions to the finite-horizon optimal control problem relies upon the solution of the Hamilton-Jacobi-Bellman (HJB) partial differential equation, which may be difficult or impossible to obtain in closed-form. Herein we propose a methodology to avoid the explicit solution of such HJB pde for input-affine nonlinear systems by means of a *dynamic extension*. This results in a dynamic time-varying state feedback yielding an approximate solution to the *finite-horizon optimal control* problem.**

I. INTRODUCTION

The problem of controlling the state of a system during a desired time interval, which is generally decided *a priori*, and to minimize, along the resulting trajectory of the system, a criterion of optimality is crucial in control system applications [4]. The problem informally defined above is commonly referred to as *finite-horizon optimal control* problem.

Two different approaches are available in the literature to solve the problem, namely the Minimum Principle and the Dynamic Programming approach [1], [3], [5], [6], [10], [11]. The former hinges upon the definition of the *Hamiltonian* associated to the optimal control problem that must be minimized by the optimal control law. The latter is based on the *principle of optimality* [2], which formalizes the intuitive requirement that a truncation of the optimal control law must be optimal with respect to the resulting truncated problem. Obviously, each approach has its own advantages and drawbacks.

The solution relying on the Minimum Principle – which provides only necessary conditions for optimality – yields in general an open-loop control law defined in terms of an adjoint state, or *costate*, satisfying an ordinary differential equation with a boundary condition imposed on the value of the costate at the terminal time. Therefore, determining the trajectories of the state and the costate consists in finding the solution of a *two-point boundary value* problem.

The Dynamic Programming approach, on the other hand, provides necessary and sufficient conditions for optimality and the resulting control policy is a memoryless time-varying feedback. This methodology hinges upon the solution of the Hamilton-Jacobi-Bellman partial differential equation, which may be in general difficult or impossible to compute in closed-form.

The main contribution of this paper is a method to construct *dynamically*, $i.e.$ by means of a dynamic extension, an exact solution of a (modified) HJB pde for input-affine nonlinear systems without actually solving any partial differential equation. The *extended* closed-loop system is a system of ordinary differential equations with two-point boundary conditions. However, differently from the Minimum Principle, if an approximate solution is sought then the solution of the *two-point boundary value* problem can be avoided and the cost of this approximation can be explicitly quantified.

The rest of the paper is organized as follows. In Section II the definition of the problem is given and the basic notation is introduced. A *notion* of solution of the Hamilton-Jacobi-Bellman partial differential equation is provided in Section III. The design of a dynamic time-varying state feedback that approximates the solution of the finite-horizon optimal control problem is presented in Section IV. The paper is completed by a numerical example and by some conclusions in Sections V and VI, respectively.

II. DEFINITION OF THE PROBLEM

Consider a nonlinear system described by equations of the form

$$\dot{x} = f(x) + g(x)u\,, \tag{1}$$

where $x(t) \in \mathbb{R}^n$ is the state of the system while $u(t) \in \mathbb{R}^m$ denotes the control input. The mappings $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are assumed to be sufficiently smooth. The finite-horizon optimal control problem consists in determining a control input u that minimizes the cost functional[1]

$$J(x(0), u) \triangleq \frac{1}{2}\int_0^T (q(x)+u^\top u)dt + \frac{1}{2}x(T)^\top Kx(T)\,, \tag{2}$$

[1]For simplicity we consider only a quadratic terminal cost.

This work is partially supported by the Austrian Center of Competence in Mechatronics (ACCM), by the EPSRC Programme Grant Control for Energy and Sustainability EP/G066477 and by the MIUR under PRIN Project Advanced Methods for Feedback Control of Uncertain Nonlinear Systems.

M. Sassano is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK (Email: mario.sassano08@imperial.ac.uk).

A. Astolfi is with the Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, UK and with the Dipartimento di Informatica, Sistemi e Produzione, Università di Roma "Tor Vergata", Via del Politecnico 1, 00133 Roma, Italy (Email: a.astolfi@ic.ac.uk).

where $T > 0$ is the fixed terminal time, $q : \mathbb{R}^n \to \mathbb{R}_+$ is the running cost imposed on the state of the system, $K = K^\top \geq 0$ penalizes the value of the state of the system at the terminal time, and subject to the dynamic constraint (1) and the initial condition $x(0) = x_0$. The Dynamic Programming solution of the fixed terminal time, free-endpoint optimal control problem hinges upon the solution, $V(x,t) : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$, of the Hamilton-Jacobi-Bellman partial differential equation[2]

$$V_t + V_x f(x) + \frac{1}{2}q(x) - \frac{1}{2}V_x g(x)g(x)^\top V_x^\top = 0\,, \quad (3)$$

together with the boundary condition

$$V(x,T) = \frac{1}{2}x^\top K x\,, \quad (4)$$

for all $x \in \mathbb{R}^n$. The solution of the HJB equation (3), if it exists, is the *value function* of the optimal control problem, *i.e.* it is a function that associates to every initial state x_0 the optimal cost, namely

$$V(x_0,0) = \min_u \frac{1}{2}\left\{\int_0^T (q(x) + u^\top u)dt + x(T)^\top K x(T)\right\}. \quad (5)$$

Problem 1: Consider system (1) and the cost (2). The *regional dynamic finite-horizon optimal control* problem consists in determining an integer $\tilde{n} \geq 0$, a dynamic control law described by the equations

$$\begin{aligned} \dot{\zeta} &= \alpha(x,\zeta,t)\,, \\ u &= \beta(x,\zeta,t)\,, \end{aligned} \quad (6)$$

with $\alpha : \mathbb{R}^n \times \mathbb{R}^{\tilde{n}} \times \mathbb{R} \to \mathbb{R}^{\tilde{n}}$, $\beta : \mathbb{R}^n \times \mathbb{R}^{\tilde{n}} \times \mathbb{R} \to \mathbb{R}^m$, and a set $\bar{\Omega} \subset \mathbb{R}^n \times \mathbb{R}^{\tilde{n}}$ containing the origin of $\mathbb{R}^n \times \mathbb{R}^{\tilde{n}}$ such that the closed-loop system

$$\begin{aligned} \dot{x} &= f(x) + g(x)\beta(x,\zeta,t)\,, \\ \dot{\zeta} &= \alpha(x,\zeta,t)\,, \end{aligned} \quad (7)$$

is such that

$$J((x_0,\zeta_0),\beta) \leq J((x_0,\zeta_0),\bar{u})\,, \quad (8)$$

for any $\bar{u}$ and any (x_0,ζ_0) such that the trajectory of the system (7) remain in $\bar{\Omega}$ for all $t \in [0,T]$.

Herein an approximate solution of the *regional dynamic finite-horizon optimal control* problem is determined, as described in the following statement.

Problem 2: Consider system (1) and the cost (2). The *approximate regional dynamic finite-horizon optimal control* problem consists in determining an integer $\tilde{n} \geq 0$, a dynamic control law described by (6), a set $\bar{\Omega} \subset \mathbb{R}^n \times \mathbb{R}^{\tilde{n}}$ containing the origin of $\mathbb{R}^n \times \mathbb{R}^{\tilde{n}}$ and functions $\rho_1 : \mathbb{R}^n \times \mathbb{R}^{\tilde{n}} \times \mathbb{R} \to \mathbb{R}_+$ and $\rho_2 : \mathbb{R}^n \times \mathbb{R}^{\tilde{n}} \to \mathbb{R}_+$ such that the regional dynamic finite-horizon optimal control problem is solved with respect to the running cost

$$\mathcal{L}(x,\zeta,u) \triangleq q(x) + \rho_1(x,\zeta,t) + u^\top u\,, \quad (9)$$

and the terminal cost

$$\mathcal{T}(x(T),\zeta(T)) \triangleq \frac{1}{2}x(T)^\top K x(T) + \rho_2(x(T),\zeta(T))\,. \quad (10)$$

Finally recall that in the linear-quadratic case, *i.e.* the underlying system is

$$\dot{x} = Ax + Bu\,, \quad (11)$$

and the running cost on the state is $q(x) = x^T Q x$, the solution of the finite-horizon optimal control problem is a linear time-varying state feedback of the form $u = -B^\top \bar{P}(t)x$, where $\bar{P}(t)$ is the (symmetric positive semidefinite) solution of the differential Riccati equation

$$\dot{\bar{P}}(t) + \bar{P}(t)A + A^\top \bar{P}(t) - \bar{P}(t)BB^\top \bar{P}(t) + Q = 0\,, \quad (12)$$

with the boundary condition $\bar{P}(T) = K$.

III. Algebraic Solution and Value Function

In this section a notion of solution of the Hamilton-Jacobi-Bellman equation (3) is presented, see [7], [8], [9] for a similar approach in the case $T = +\infty$. To this end, consider the augmented system

$$\dot{z} = F(z) + G(z)u\,, \quad (13)$$

with $z(t) = (x(t)^\top, \tau(t))^\top \in \mathbb{R}^{n+1}$, $F(z) = (f(x)^\top, 1)^\top$ and $G(z) = (g(x)^\top, 0)^\top$. Note that the partial differential equation (3) can be rewritten as

$$V_z F(z) + \frac{1}{2}q(x) - \frac{1}{2}V_z G(z)G(z)^\top V_z^\top = 0\,. \quad (14)$$

Following [8] and [9], consider the HJB equation (14) and suppose that it can be solved *algebraically*, as detailed in the following definition.

Definition 1: Let $\Sigma : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^{n\times n}$, with $x^\top\, \Sigma(x,\tau)\; x > 0$, for all $(x,\tau) \in \mathbb{R}^n \setminus \{0\} \times \mathbb{R}$, and $\sigma : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}_+$. A continuously differentiable mapping $P(x,\tau) = [p(x,\tau)^\top, r(x,\tau)]^\top$, $p : \mathbb{R}^{n+1} \to \mathbb{R}^{1\times n}$, $r : \mathbb{R}^{n+1} \to \mathbb{R}$, is an *algebraic* $\bar{P}$ *solution* of (3) if

$$\begin{aligned} & p(x,\tau)f(x) + r(x,\tau) - \frac{1}{2}p(x,\tau)g(x)g(x)^\top p(x,\tau)^\top \\ & + \frac{1}{2}q(x) + x^\top \Sigma(x,\tau)x + \tau^2\sigma(x,\tau) = 0\,, \end{aligned} \quad (15)$$

and $p(0,\tau) = 0$, $r(0,\tau) = 0$,

$$\left.\frac{\partial p^\top}{\partial \tau}\right|_{(0,\tau)} = 0\,, \quad \left.\frac{\partial r}{\partial \tau}\right|_{(0,\tau)} = 0\,,$$

$$\left.\frac{\partial p^\top}{\partial x}\right|_{(0,\tau)} = \bar{P}(\tau)\,, \quad \left.\frac{\partial^2 r}{\partial x^2}\right|_{(0,\tau)} = \dot{\bar{P}}(\tau)\,,$$

[2]The notation V_x denotes the partial derivative of the scalar function V with respect to the variable x.

where $\bar{P}(\tau)$ is the solution of the differential Riccati equation (12) together with the boundary condition $\bar{P}(T) = K$. ◇

Obviously, since an arbitrary mapping that solves the equation (15) is selected, the mapping $P(x,\tau)$ may not be a *gradient vector*.

Using an *algebraic $\bar{P}$ solution* of the equation (15), define the function

$$V(x,\tau,\xi,s) = p(\xi,s)x + r(\xi,s)\tau + \frac{1}{2}\|x-\xi\|^2_{R(s)} + \frac{1}{2}b\|\tau - s\|^2 , \tag{16}$$

where $\xi \in \mathbb{R}^n$, $s \in \mathbb{R}$, $b > 0$ and $\|v\|^2_{R(s)}$ denotes the Euclidean norm of the vector v weighed by the positive definite matrix $R(s)$, *i.e.* $\|v\|^2_{R(s)} = v^\top R(s) v$.

IV. Main Results

We first present the main ideas of the proposed approach in the case of linear systems. To this end consider the system (11) and the cost (2) with $q(x) = x^\top Q x$. From the definition of *algebraic $\bar{P}$ solution* of the equation (14), we expect p to *approximate* (in the sense defined in (15)) the partial derivative of the value function with respect to the state x whereas r *represents* the partial derivative of V with respect to time. Therefore, in the linear case an *algebraic $\bar{P}$ solution* is given by

$$P(x,\tau) = [x^\top \bar{P}(\tau),\ \frac{1}{2}x^\top \dot{\bar{P}}(\tau)x] , \tag{17}$$

which satisfies the condition (15) if and only if the differential Riccati equation (12) is satisfied by $\bar{P}(\tau)$. Then, define the function

$$V(x,\tau,\xi,s) = \xi^\top \bar{P}(s)x + \frac{1}{2}\xi^\top \dot{\bar{P}}(s)\xi\tau \\ \frac{1}{2}\|x-\xi\|^2_{R(s)} + \frac{1}{2}b\|\tau - s\|^2 . \tag{18}$$

In the following result we show that the function (18) is indeed a *value function* for the system (11) and the cost (2), *i.e.* solves the Hamilton-Jacobi-Bellman partial differential inequality associated to the *extended* system (7), namely

$$\mathcal{HJB} \triangleq V_z F + V_\xi \dot{\xi} + V_s \dot{s} + \frac{1}{2}q - \frac{1}{2}V_z G G^\top V_z^\top \le 0 , \tag{19}$$

and satisfies the boundary condition

$$V(x,T,\xi(T),s(T)) = \frac{1}{2}x^\top K x . \tag{20}$$

Proposition 1: Consider the linear system (11) and the cost (2) with $q(x) = x^\top Q x$. Let $\bar{P}(\tau) = \bar{P}(\tau)^\top > 0$ be the solution of the differential Riccati equation (12) with the boundary condition $\bar{P}(T) = K$. Let $R(\tau) = \bar{P}(\tau)$ for all $\tau \in [0,T]$. Then there exists $\bar{k}$ such that for all $k \ge \bar{k}$, V as in (18) satisfies the partial differential equation (19) with $\dot{s} = 1$ and $\dot{\xi} = -kV_\xi^\top$. Furthermore, selecting $s(0) = 0$ and $\xi(0) = 0$, yields $V(x,T,0,T) = \frac{1}{2}x^\top K x$, *i.e.* the boundary condition (20) is satisfied. Hence

$$\begin{aligned} \dot{s} &= 1 , \\ \dot{\xi} &= -kV_\xi^\top , \\ u &= -B^\top V_x^\top = -B^\top \bar{P}(\tau)x + B^\top(\bar{P}(\tau) - \bar{P}(s))x , \end{aligned} \tag{21}$$

with $s(0) = 0$ and $\xi(0) = 0$, solves the regional dynamic finite-horizon optimal control problem.

Consider now the nonlinear system (1) and the cost (2). To begin with suppose that the matrix R in the function V, defined as in (16), is a constant symmetric positive definite matrix.

To streamline the presentation and provide a concise statement of the main result let

$$\Delta(x,\xi,s) = (R - \Phi(x,\xi,s))\Lambda(\xi,s)^\top , \tag{22}$$

$$\delta(x,\xi,s) = (R - \Phi(x,\xi,s))\lambda(\xi,s)^\top , \tag{23}$$

with $\Lambda(\xi,s) = \Psi(\xi,s)R^{-1}$, $\lambda(\xi,s) = \psi(\xi,s)R^{-1}$, where $\Phi(x,\xi,s) \in \mathbb{R}^{n\times n}$ is a continuous mapping such that

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

and $\Psi(\xi,s) \in \mathbb{R}^{n\times n}$ and $\psi(\xi,s) \in \mathbb{R}^{1\times n}$ are the Jacobian matrices of the mappings $p(\xi,s)$ and $r(\xi,s)$, respectively, with respect to ξ.

Moreover, $A_{cl}(x,\tau) = F(x) - g(x)g(x)N(x,\tau)$, with $N(x,\tau) \in \mathbb{R}^{n\times n}$ such that $p(x,\tau) = x^\top N(x,\tau)^\top$. Finally let $\ell(x,\tau,s)$, $H(x,\xi,s)$, $\Pi(x,\tau,s)$, $W_1(x,\xi,s)$, $W_2(x,s)$, $D_1(x,\xi,s)$ and $D_2(x,s)$ be such that

$$\begin{aligned} r(x,s) - r(x,\tau) &= \ell(x,\tau,s)(s-\tau) , \\ r(\xi,s) - r(x,s) &= (x-\xi)^\top H(x,\xi,s)(x-\xi) , \\ p(x,s) - p(x,\tau) &= x^\top \Pi(x,\tau,s) , \\ \frac{\partial p(\xi,s)}{\partial s} - \frac{\partial p(x,s)}{\partial s} &= W_1(x,\xi,s)(x-\xi) , \\ \frac{\partial p(x,s)}{\partial s} &= W_2(x,s)x , \\ \frac{\partial r(\xi,s)}{\partial s} - \frac{\partial r(x,s)}{\partial s} &= D_1(x,\xi,s)(x-\xi) , \\ \frac{\partial r(x,s)}{\partial s} &= D_2(x,s)x . \end{aligned}$$

Remark 1: The vector field $A_{cl}(x,\tau)x$ describes the closed-loop nonlinear system when only the *algebraic input*, namely $u = -g(x)^\top p(x,\tau)$, is applied. ▲

The following statement provides a solution to the *approximate regional dynamic finite-horizon optimal control problem* considering the cost (2) subject to the dynamical constraint (1).

Proposition 2: Consider system (1) and the cost (2). Let $P(x,\tau)$ be an *algebraic* $\bar{P}$ *solution* of (14).

Let $R = R^\top \geq \sup_{\tau\in[0,T]} \bar{P}(\tau)$ and $b > 0$ be such that

$$\left[\begin{array}{cc} L_1 & L_2 \\ L_2^T & L_3 \end{array}\right] < \left[\begin{array}{cc} \Sigma & 0 \\ 0 & \sigma \end{array}\right] + \frac{1}{2}\left[\begin{array}{c} \Delta^\top \\ \delta^\top \end{array}\right] gg^\top \left[\begin{array}{cc} \Delta & \delta \end{array}\right], \tag{24}$$

for all $(x,\tau,\xi,s) \in \Omega \subset \mathbb{R}^{2(n+1)}$, with

$$L_1 = A_{cl}^\top(x,\tau)\Pi^\top + \gamma W_2 + \Lambda Y + Y^\top\Lambda^\top + \Lambda H \Lambda^\top,$$

$$L_2 = \frac{\gamma}{2}D_2^\top + Y^\top\lambda^\top + \Lambda H\lambda^\top + \frac{\gamma}{2}\Lambda D_1^\top,$$

$$L_3 = \lambda H \lambda^\top + \frac{\gamma}{2}D_1\lambda^\top + \frac{\gamma}{2}\lambda D_1,$$

$$Y = \frac{1}{2}(R-\Phi)^\top A_{cl}(x,s) + \frac{\gamma}{2}W_1,$$

and $\gamma(x,\tau,s) = 1 - \frac{\ell(x,\tau,s)}{b}$.

Then there exists $\bar{k}$ such that for all $k \geq \bar{k}$ the function $V(x,\tau,\xi,s)$ defined in (16) satisfies the Hamilton-Jacobi-Bellman inequality (19) for all $(x,\tau,\xi,s) \in \Omega$, with $\dot{\xi} = -kV_\xi^\top$ and $\dot{s} = \gamma(x,\tau,s)$. Furthermore, selecting $s(T) = T$ and $\xi(T) = 0$, yields $V(x,T,0,T) = \frac{1}{2}x^T R x$, hence

$$\begin{aligned} \dot{s} &= \gamma(x,\tau,s), \\ \dot{\xi} &= -k(\Psi(\xi,s)^\top x - R(x-\xi) + \psi(\xi,s)^\top\tau), \\ u &= -g(x)^\top\left[p(x,\tau)^\top + (R-\Phi)(x-\xi) + \Pi^\top x\right] \end{aligned} \tag{25}$$

with $s(T) = T$, $\xi(T) = 0$, solves the approximate regional dynamic finite-horizon optimal control problem with $\rho_1 = -\mathcal{HJB}$ and $\rho_2 = \frac{1}{2}x(T)^\top(R-K)x(T)$. Finally, suppose that condition (24) is satisfied by $R = K$ then the additional cost on the final state is zero, *i.e.* $\rho_2 = 0$.

Remark 2: Suppose that $P(z)z > 0$ for all $z \neq 0$ and consider the limit of (2) as T tends to infinity. Let $\mathcal{A} = \{(x,\tau,\xi,s) \in \Omega : x = \xi = 0, s = \tau\}$. Then V as in (16) is zero on $\mathcal{A}$, *i.e.* $V|_\mathcal{A} = 0$, and there exist $R = R^\top > 0$, $b > 0$ and a set $\tilde{\Omega} \supset \Omega$ such that $V|_{\tilde{\Omega}\setminus\mathcal{A}} > 0$. If R is such that (19) is satisfied then $\dot{V} \leq 0$, along the trajectories of the closed-loop system (13)-(25), and the set $\mathcal{A}$ is stable for the system (13)-(25). ▲

Remark 3: The closed-loop system (13)-(25) is defined by a system of $2(n+1)$ ordinary differential equations with *two-point boundary conditions*, similar to the problem obtained exploiting the arguments of the Minimum Principle. However with the approach proposed herein, differently from the Minimum Principle, if the solution of the *two-point boundary value* problem is not precisely computed, *i.e.* $\xi(T) = \varepsilon_1 \neq 0$ and $s(T) = \varepsilon_2 \neq T$, then it can be guaranteed, by Remark 2, that the resulting trajectories are not *far away* from the optimal evolutions and moreover the additional cost is given by

$$\rho_2 = p(\varepsilon_1,\varepsilon_2)x(T) + r(\varepsilon_1,\varepsilon_2)T + \frac{1}{2}b(T-\varepsilon_2)^2$$
$$+ \frac{1}{2}x(T)^\top(R-K)x(T) + \frac{1}{2}\varepsilon_1^\top R\varepsilon_1 - x(T)^\top R\varepsilon_1.$$

As a matter of fact, if an approximate solution of the problem is sought, then the solution of the *two-point boundary value* problem can be avoided selecting the initial condition of ξ and s such that the quantity ρ_2 is minimized. ▲

Remark 4: Since $z = 0$ is not an equilibrium point for $F(z)$ defined in (13), there are terms in the partial differential equation (19), namely $\mu(x,\tau,s) \triangleq r(x,s) - r(x,\tau) + b(\tau - s) - b(\tau-s)\dot{s}$, that can not be written as a *quadratic form* in x, $(x-\xi)$ and τ. The choice $\dot{s} = \bar{\gamma}$, with $\bar{\gamma}$ constant, is sufficient to guarantee that the term μ is smaller than zero. In fact, let $\dot{s} = \bar{\gamma}$, $s(0) = (1-\bar{\gamma})T$ and note that $s(t) \geq \tau(t)$ and $s(T) = T$. Let $\bar{\ell} > 0$ be such that $|r(x,s) - r(x,\tau)| < \bar{\ell}(s-\tau)$ for all $(x,\tau,s) \in \Omega$, then $\bar{\gamma} \leq 1 - \frac{\bar{\ell}}{b}$ guarantees that $\mu \leq 0$ for all $(x,\tau,s) \in \Omega$. In this case, the state variable s can be substituted in the dynamic control law (25) by the function of time $s(t) = T + \bar{\gamma}(t-T)$. ▲

Remark 5: The control law (21), in the linear case, is consistent with the conditions given in Proposition 2. In fact, (21) can be obtained from (25) letting $\gamma = \bar{\gamma} = 1$, hence $s(0) = 0$, and noting that, by the assumptions in Proposition 1,

$$\dot{\xi} = -k\left[\dot{\bar{P}}(s)\tau + \bar{P}(\tau)\right]\xi,$$

with an equilibrium point at $\xi = 0$, hence $\xi(0) = 0$ provides a solution to the *two-point boundary value* problem defined in Proposition 2. ▲

Remark 6: The left-hand side of the condition (24) is zero at the origin. Assume additionally that $\Sigma(0,0) = \bar{\Sigma} > 0$ and $\sigma(0,0) = \bar{\sigma} > 0$ in the definition of *algebraic solution*. Then, by continuity of the left-hand side of inequality (24), there exists a non-empty subset $\hat{\Omega} \subset \mathbb{R}^{2(n+1)}$, containing the origin, such that the condition (24) is satisfied for all $(x,\tau,\xi,s) \in \hat{\Omega}$. Therefore, the *algebraic* $\bar{P}$ *solution* of (15), with $\Sigma(0,0) = \bar{\Sigma}$ and $\sigma(0,0) = \bar{\sigma}$, solves the approximate regional dynamic finite-horizon optimal control problem in $\hat{\Omega}$. ▲

Remark 7: The gain k in the dynamics of ξ, namely the second equation of (25), may be defined as a function of (x,τ,ξ,s), *i.e.* $k(x,\tau,\xi,s)$, in order to reduce the additional running cost ρ_1, see [9] and the numerical example for more details. ▲

V. Numerical Example

To illustrate the results of the paper consider the system

$$\begin{aligned} \dot{x}_1 &= x_2\,, \\ \dot{x}_2 &= x_1 x_2 + u\,, \end{aligned} \tag{26}$$

with $x(t) = (x_1(t), x_2(t))^\top \in \mathbb{R}^2$, $u(t) \in \mathbb{R}$ and the cost

$$J(x(0), u) = \frac{1}{2}\int_0^T u(t)^2 dt + \frac{1}{2}\left[x_1(T)^2 + x_2(T)^2\right]. \tag{27}$$

Note that no running cost is imposed on the state of the system, hence only the position of the state at the terminal time is penalized, together with the control effort. Let

$$\bar{P}(\tau) = \begin{bmatrix} \bar{p}_{11}(\tau) & \bar{p}_{12}(\tau) \\ \bar{p}_{12}(\tau) & \bar{p}_{22}(\tau) \end{bmatrix}$$

be the solution of the differential Riccati equation (12) with the boundary condition $\bar{P}(T) = I_2$. Let

$$\begin{aligned} p(x, \tau) &= x^\top \bar{P}(\tau) + \mathcal{Q}(x, \tau)\,, \\ r(x, \tau) &= \tfrac{1}{2} x^\top \dot{\bar{P}}(\tau) x\,, \end{aligned} \tag{28}$$

with $\mathcal{Q} = [\mathcal{Q}_1,\, \mathcal{Q}_2] \in \mathbb{R}^{1\times 2}$ and

$$\mathcal{Q}_1(x, \tau) = -\bar{p}_{12}(\tau)x_1^2 - \bar{p}_{22}(\tau)x_1 x_2\,,$$

$$\mathcal{Q}_2(x, \tau) = -\bar{p}_{12}(\tau)x_1 - \bar{p}_{22}(\tau)x_2 + x_1 x_2 + \sqrt{\chi(x_1, x_2, \tau)}\,, \tag{29}$$

with

$$\chi(x_1, x_2, \tau) \triangleq \bar{p}_{12}(\tau)^2 x_1^2 + 2\bar{p}_{12}(\tau)\bar{p}_{22}(\tau)x_1 x_2 + \bar{p}_{22}(\tau)^2 x_2^2$$
$$- 2\bar{p}_{12}(\tau)x_1^2 x_2 - 2\bar{p}_{22}(\tau)x_1 x_2^2 + x_1^2 x_2^2 + 2\vartheta x_1^2 + 2\vartheta x_2^2\,.$$

Letting $\Sigma = \vartheta I_2$, $\vartheta > 0$, it can be shown that $P(x, \tau) = [p(x, \tau),\, r(x, \tau)]$, with p and r defined in (28), is an *algebraic $\bar{P}$ solution* for the system (26) as defined in (15).

The dynamic solution (25) proposed herein is compared with the optimal solution of the linearized problem, namely $u_l = -B^\top \bar{P}(\tau)x$. Note that the control law u_l is designed for the linear part of the system (26), which is described by a double integrator, and performs poorly on the nonlinear system (26). The following simulations show that the performances of the control law u_l may be improved determining a solution of the algebraic equation (15), constructing the *augmented value function* V as in (16) and obtaining the dynamic control (25). To compare the performances of the two control laws we introduce the ratio

$$\eta = \frac{J(x(0), u_d)}{J(x(0), u_l)}\,,$$

where J is defined in (27) and u_d denotes the dynamic control law (25). A ratio smaller than one implies that the cost *yielded* by the dynamic control law is smaller than the cost yielded by the linear control law u_l.

In the first simulation we selcect $(x_1(0),\, x_2(0)) = (\frac{3}{2},\, \frac{3}{2})$, $R = \alpha I_2$, with $\alpha > 1$, $\bar{\gamma} = 1$ and the terminal time $T = 1$.

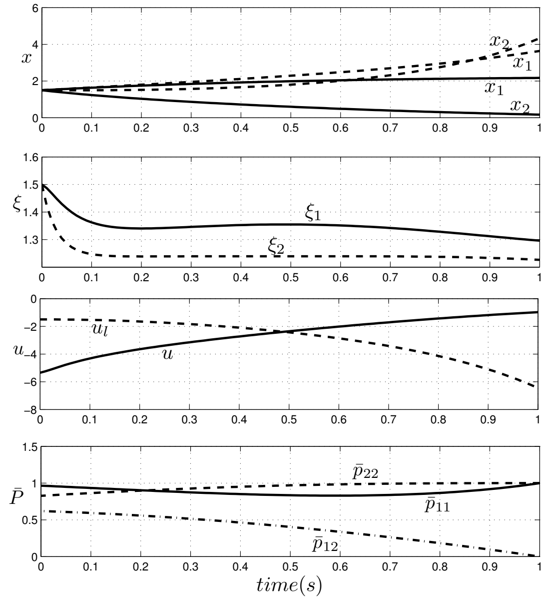

Fig. 1. Top Graph: Time histories of the state of the system (26) in closed-loop with the linear control law u_l (dashed line) and with the dynamic control law (25) (solid line). Upper Middle Graph: Time histories of the state of the dynamic extension ξ, with $\xi(0)$ such that ρ_2 is minimized. Lower Middle Graph: Time histories of the control action u_l (dashed line) and of the dynamic control law (25) (solid line). Bottom Graph: Time histories of the elements of the matrix $\bar{P}(t)$, solution of the differential Riccati equation (12).

The gain $k(x, \tau, \xi, s)$ is selected as in Remark 7 and it is such that the partial differential equation (19) and consequently the additional cost ρ_1 are identically zero. The top graph of Figure 1 displays the time histories of the state of the system (26) in closed-loop with the linear control law u_l (solution of the linearized problem) and with the dynamic control law (25). The time histories of the control signals u_l and (25) are depicted in the lower middle graph of Figure 1. Note that, considering ρ_1 as defined in Proposition 2 and ρ_2 obtained as in Remark 3, the ratio η is equal to 0.4234. The upper middle graph of Figure 1 shows the time histories of the dynamic extension ξ with $\xi(0)$ selected such that ρ_2, namely the additional cost on the terminal *augmented* state, is minimized. Finally, the time histories of the elements of the matrix $\bar{P}(\tau)$, solution of the differential Riccati equation (12) are displayed in the bottom graph of Figure 1.

In the second simulation we consider several initial conditions. Figure 2 displays the phase portraits of the system (26) in closed-loop with the linear control law u_l and with the dynamic control law (25).

Figure 3 shows the time histories of the system (26) in closed-loop with the the linear control law u_l and with the dynamic control law (25) for different terminal times, namely $T = 1, 1.5, 2, 2.5$, and the initial condition $(x_1(0),\, x_2(0)) = (2,\, -2)$.

In a *cheap control* scenario we let $K = cI$, $c > 1$. Figure 4

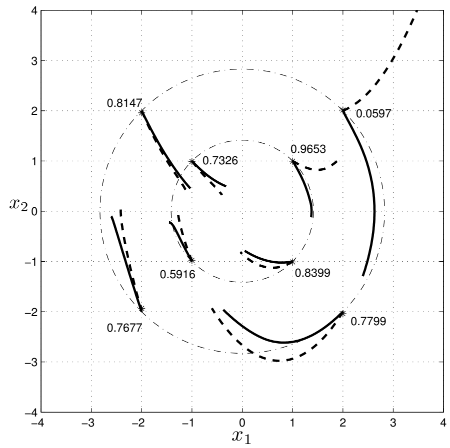

Fig. 2. Phase portraits of the system (26) in closed-loop with the linear control law u_l (dashed line) and with the dynamic control law (25) (solid line). The values describe the ratio η of the cost of the dynamic control law over the cost of the linear control law u_l for the corresponding initial condition.

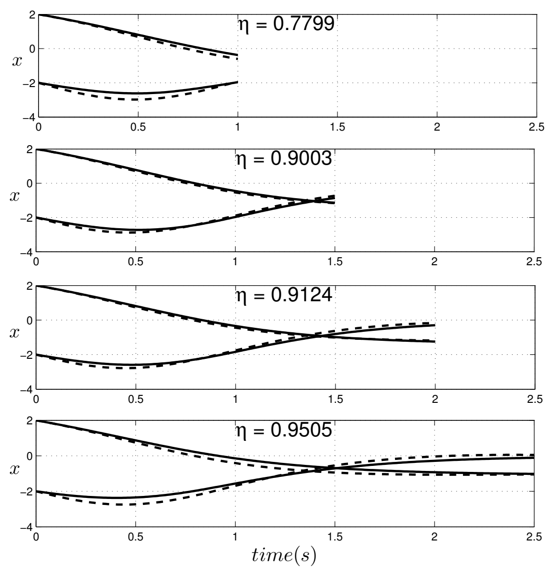

Fig. 3. Time histories of the system (26) in closed-loop with the linear control law u_l (dashed line) and with the dynamic control law (25) (solid line) for different terminal times.

shows the trajectories of the system (26) from the initial condition $(x_1(0), x_2(0)) = (2, 2)$, with $K = cI$ for different values of c, in closed-loop with u_l and with the dynamic control law (25).

VI. Conclusions

The finite-horizon optimal control problem for input-affine nonlinear systems is solved within the framework of dynamic programming. It is shown that the computation of the solution of the Hamilton-Jacobi-Bellman partial differential equation can be avoided provided an additional

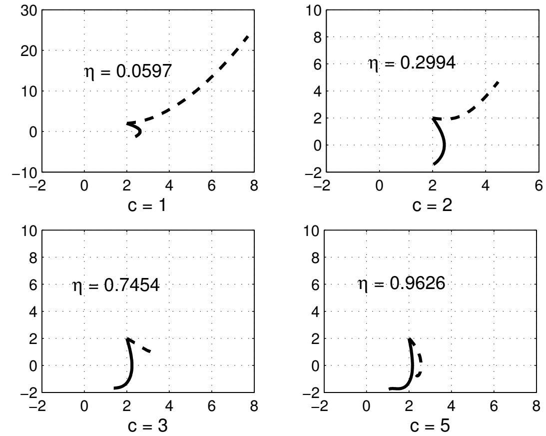

Fig. 4. Trajectories of the system (26), with $K = cI$ for different values of c, in closed-loop with u_l (dashed line) and with the dynamic control law (25) (solid line), respectively.

cost is paid. The methodology makes use of a *dynamic extension* yielding a dynamic control law that solves the approximate regional dynamic finite-horizon optimal control problem. The closed-loop system is defined in terms of a *two-point boundary value* problem. Finally, differently from the Minimum Principle approach, if the solution of the *two-point boundary value* problem is not exactly determined then the error generated by the approximation can be explicitly quantified and minimized.

References

[1] B. D. O. Anderson and J. B. Moore. *Optimal Control: Linear Quadratic Methods*. Prentice-Hall, 1989.

[2] R. E. Bellman. *Dynamic Programming*. Princeton University Press, New Jersey, 1957.

[3] D. P. Bertsekas. *Dynamic Programming and Optimal Control - Volume I, third edition*. Athena Scientific, 2005.

[4] A. E. Bryson and Y. C. Ho. *Applied Optimal Control: optimization, estimation, and control*. Taylor and Francis, New York, 1975.

[5] G. Knowles. *An Introduction to Applied Optimal Control*. Academic Press, New York, 1981.

[6] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko. *The Mathematical Theory of Optimal Processes*. Interscience, New Jersey, 1962.

[7] M. Sassano and A. Astolfi. Dynamic approximate solutions of the HJB equation and of the HJ inequality. *Submitted to IEEE Transactions on Automatic Control*.

[8] M. Sassano and A. Astolfi. Dynamic solution of the Hamilton-Jacobi inequality in the L_2-disturbance attenuation problem. In *Proc. of the 8th IFAC Symposium on Nonlinear Control Systems, Bologna, Italy*, pages 3081–3086, 2010.

[9] M. Sassano and A. Astolfi. Dynamic solution of the HJB equation and the optimal control of nonlinear systems. In *Proc. of the 49th Conference on Decision and Control, Atlanta, GA*, 2010.

[10] H. J. Sussmann and J. C. Willems. 300 years of optimal control: from the Brachistochrone to the Maximum Principle. *IEEE Control Systems Magazine*, 17:32–44, 1997.

[11] R. B. Vinter. *Optimal Control*. Birkhauser Boston, 2000.