An Iterative Method for Optimal Control of Nonlinear Quadratic Tracking Problems

Xin Ning, Walter Bomela, and Jr-Shin Li

Abstract— **In this paper, we investigate an iterative method for computing optimal controls for general affine nonlinear quadratic tracking problems. The control law is computed iteratively by solving a sequence of linear quadratic tracking problems and, in particular, it consists of solving a set of coupled differential equations derived from the Hamilton-Jacobi-Bellman equation. The convergence of the iterative scheme is shown by constructing a contraction mapping and using the fixed-point theorem. The versatility and effectiveness of the proposed method is demonstrated in numerical simulations of three structurally different nonlinear systems.**

I. Introduction

Optimal control of nonlinear systems is a challenging problem that has been widely investigated. While some optimal control problems, in particular for linear systems, can be solved analytically [1], this is often not feasible for many nonlinear systems. As a result, many researchers have dedicated a considerable amount of effort in investigating effective numerical control algorithms [2]–[5].

The finite and infinite horizon optimal control problems for bilinear systems were examined in [2], where the feedback control law was obtained by solving the Hamilton-Jacobi equation and, more importantly, the bilinear control problem was solved by constructing a sequence of linear problems. On the other hand, [3] proposed an iterative procedure for solving an unconstrained finite-horizon bilinear quadratic control problem, and the convergence of the procedure was shown by using the fixed-point theorem. Later, [4] formulated an iterative method for the fixed-endpoint optimal control problem for bilinear systems. A transformation of coordinates was exploited to transform the bilinear system problem into an equivalent time-varying linear system that could be solved iteratively using the sweep method.

More recently, an interesting computational method for synthesizing optimal control for nonholonomic dynamical systems was proposed in [5]. This method solves the problem in two main steps, first a feasible control is found by solving an unconstrained quadratic programming problem, while the second step consists of iteratively computing the optimal control input by solving a linear constrained quadratic programme. Many other computational methods have been proposed for nonlinear dynamical systems, which can be categorized as direct or indirect methods [6]. For direct methods, for instance, dynamic programming [7], [8] and pseudo-spectral methods have been widely applied to optimal control problems with final constraints, including optimal control of quantum ensembles [9], [10], and neuron ensembles [11]. Unfortunately, these methods can hardly overcome the curse of dimensionality because a suitable discretization grid is required, which increases the size of the optimization problem to solve, hence making these approaches less practical for large control problems. In addition, the convergence speed of the solver can greatly degrade with the size of the control problem.

In this paper, we propose an iterative scheme for solving optimal tracking problem for affine nonlinear system, that is an extension of the method presented in [12] which only considered the case of bilinear systems. One of the main advantage of our scheme is the ability to tune the control matrices to not only achieve the desired performance, but also to improve the convergence rate. The remainder of the paper is organized as follows. In Section II, we formulate the optimal tracking control problem for control-affine nonlinear systems, and then derive the control law by solving the Hamilton-Jacobi-Bellman (HJB) equation. The iterative algorithm together with the proof of convergence are presented in Section III, with more detailed derivations given in the Appendix. Various numerical examples of nonlinear systems are presented in Section IV that demonstrate the effectiveness of the proposed method. Finally, the conclusion is given in Section V.

II. Optimal Tracking Control

A. Problem Formulation

Consider the optimal tracking control problem of the control-affine nonlinear system of the form

$$\dot{x}(t) = f(x) + \sum_{i=1}^{m} b_i(x)u_i(t), \quad y(t) = l(x), \qquad (1)$$

where $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ and $y(t) \in \mathbb{R}^r$. We start our analysis by considering a class of nonlinear systems that have a drift term of the form, $f(x) = A(x)x + \bar{f}(x)$, and a linear output function, $y = Cx$. Furthermore, the matrices $A(x)$, $B(x)$, C and the vector $\bar{f}(x)$ are of appropriate dimensions, and for convenience we denote $\sum_{i=1}^{m} b_i(x)u_i(t) = B(x)u(t)$. In Section IV, we provide three interesting examples where each illustrates a different structure of the drift

This work was supported in part by the NSF awards CMMI-1763070, CMMI-1933976, and ECCS-1810202, the AFOSR award FA9550-17-1-0166, and the NIH grant R01GM131403.

X. Ning is with the Department of Electrical and Systems Engineering, Washington University, St. Louis, MO 63130, USA xin.ning@wustl.edu

W. Bomela is with the Department of Electrical and Systems Engineering, Washington University, St. Louis, MO 63130, USA wbomela@wustl.edu

J.-S. Li is with the Department of Electrical and Systems Engineering, Washington University, St. Louis, MO 63130, USA jsli@wustl.edu

term, $f(x)$. For simplicity, in the following we will denote $A(x)$ and $\bar{f}(x)$ as A and $\bar{f}$, respectively.

Let $z(t) \in \mathbb{R}^r$ denote the reference trajectory. We wish to design a control input $u(t)$ such that the output $y(t)$ of the nonlinear system (1) tracks the reference trajectory as close as possible, while minimizing the performance index

$$\begin{aligned} J(u) =& \tfrac{1}{2} e'(T) F e(T) \\ &+ \tfrac{1}{2} \int_0^T e'(t) Q(t) e(t) + u'(t) R(t) u(t) dt, \end{aligned} \tag{2}$$

where $e(t) = z(t) - y(t)$ is the error, and $F \succeq 0$, $Q(t) \succeq 0$ are $r \times r$ positive semi-definite matrices and $R(t) \succ 0$ is $m \times m$ positive definite $\forall t \in [0, T]$. Here, we assume that the entries of A, $\bar{f}$, C, Q and R are continuous functions over $[0, T]$, and $B(x(t))$ is Lipschitz continuous in $x(t)$. In addition, the nonlinear system in (1) is assumed controllable and observable.

B. Optimal Solution from the HJB Equation

Assume that $V(t, x)$ is the value function associated with the optimal control problem (2), and denote $\mathcal{U}_{(t,x)} \subseteq \mathbb{R}^m$ as the set of all admissible controls for the problem with the pair (t, x), then we have

$$V(t, x) = \inf_{u \in \mathcal{U}_{(t,x)}} J(u).$$

Furthermore, suppose that V is differentiable with respect to (t, x), then the sufficient condition of optimality is given by the HJB equation [13] as

$$V_t + \min_{u \in \mathcal{U}} \{ V_x (Ax + \bar{f} + B(x) u) + \tfrac{1}{2} e' Q e + \tfrac{1}{2} u' R u \} \equiv 0, \tag{3}$$

where V_t, V_x are the partial derivatives $\frac{\partial V}{\partial t}$ and $\frac{\partial V}{\partial x}$, respectively. The necessary condition of optimality admits the optimal control as

$$u(t) = -R^{-1} B(x)' V_x'. \tag{4}$$

Substituting $e(t)$ and (4) into (3) yields

$$\begin{aligned} V_t + \tfrac{1}{2}(V_x A x + x' A' V_x') + V_x \bar{f} - \tfrac{1}{2} V_x E(x) V_x' \\ + \tfrac{1}{2} z' Q z + \tfrac{1}{2} x' D x - \tfrac{1}{2}(x' W z + z' W' x) \equiv 0, \end{aligned} \tag{5}$$

where, for simplicity, we let $E(x) = B(x) R^{-1}(t) B'(x)$, $D(t) = C' Q(t) C$ and $W(t) = C' Q(t)$. A candidate solution to (5) is of the following form

$$V(t, x^*) = \tfrac{1}{2} x^{*\prime} P(t) x^* - x^{*\prime} g(t) + h(t), \tag{6}$$

where x^* denotes the optimal state trajectory under the optimal control, and $P(t) \in \mathbb{R}^{n \times n}$, $g(t) \in \mathbb{R}^n$ and $h(t) \in \mathbb{R}$. Then, the partial derivatives of (6) are obtained as

$$\begin{aligned} V_x' &= P(t) x^* - g(t), \\ V_t &= \tfrac{1}{2} x^{*\prime} \dot{P}(t) x^* - x^{*\prime} \dot{g}(t) + \dot{h}(t). \end{aligned} \tag{7}$$

Taking (7) into (5) yields

$$\begin{aligned} &\tfrac{1}{2} x^{*\prime} (\dot{P} + PA + A'P - P'EP + D) x^* + \dot{h} + \tfrac{1}{2} z' Q z \\ &x^{*\prime} (-\dot{g} - A' g + P' E g - W z + P \bar{f}) - \tfrac{1}{2} g' E g - g' \bar{f} \equiv 0, \end{aligned}$$

which should be satisfied for all $x^*(t), z(t)$, and $\forall t \in [0, T]$. Then, one can obtain the following state-dependent equations which characterize the optimal solution of the tracking problem (2)

$$\begin{aligned} \dot{P}(t) &= -P(t) A - A' P(t) + P(t) E(x) P(t) - D(t), \\ \dot{g}(t) &= -(A' - P(t) E(x)) g(t) + P(t) \bar{f} - W(t) z, \\ \dot{h}(t) &= g'(t) \bar{f} + \tfrac{1}{2} g'(t) E(x) g(t) - \tfrac{1}{2} z' Q(t) z, \end{aligned} \tag{8}$$

where the boundary conditions are $P(T) = C'(T) F C(T)$, $g(T) = C'(T) F z(T)$ and $h(T) = \frac{1}{2} z'(T) F z(T)$, which are derived by setting $V(T, x(T)) = \frac{1}{2} e'(T) F e(T)$. Consequently, combining (4) and (6), we obtain the optimal feedback control law

$$u^*(t) = -R^{-1} B'(x^*) [P(t) x^*(t) - g(t)], \tag{9}$$

with x^* given by

$$x^*(t) = [A^* - E(x^*) P(t)] x^*(t) + \bar{f}^* + E(x^*) g(t), \tag{10}$$

where $A^*, \bar{f}^*$ denote $A(x^*)$ and $\bar{f}(x^*)$, respectively. Then the optimal cost for the problem is $J^* = J(u^*) = \frac{1}{2} x^{*\prime}(0) P(0) x^*(0) - x^{*\prime}(0) g(0) + h(0)$.

Note that the derived result is consistent with the one presented in previous work [12], when $f(x)$ in (1) is linear, i.e., $f(x) = A(t) x$. In this case, the ODEs in (8) are equivalent to the ones presented in [12].

III. An Iterative Algorithm and its Convergence Properties

In this section, we present an algorithm for solving the optimal tracking control problem of interest and provide the proof of convergence. The main idea is to solve (8) in an iterative manner.

A. The Iterative Algorithm

Starting with an initial trajectory $x_k = x_0(t)$ at the k^{th} iteration with a desired trajectory $z(t)$, the algorithm evolves by solving the following updating iteration equations,

$$\begin{aligned} \dot{P}_{k+1}(t) =& -P_{k+1}' A_k - A_k' P_{k+1} + P_{k+1}' E_k P_{k+1} - D, \\ \dot{g}_{k+1}(t) =& -[A_k' - P_{k+1} E_k] g_{k+1} + P_{k+1} \bar{f}_k - W z(t), \\ \dot{h}_{k+1}(t) =& g_{k+1}' \bar{f}_k + \tfrac{1}{2} g_{k+1}' E_k g_{k+1} - \tfrac{1}{2} z'(t) Q z(t), \\ u_{k+1}(t) =& -R^{-1} B_k' [P_{k+1}(t) x_{k+1}(t) - g_{k+1}(t)], \\ \dot{x}_{k+1}(t) =& A_k x_{k+1}(t) + \bar{f}_k + B_k u_{k+1}(t), \end{aligned} \tag{11}$$

with the boundary conditions $P_{k+1}(T) = P(T)$, $g_{k+1}(T) = g(T)$ and $h_{k+1}(T) = h(T)$, where $A_k = A(x_k)$, $\bar{f}_k = \bar{f}(x_k)$, $B_k = B(x_k)$ and $E_k = E(x_k)$. This procedure is summarized in Algorithm 1.

B. Convergence Analysis

In this section, we use contraction mapping and the fixed-point theorem to show the convergence of the proposed iterative method. To facilitate the proof, the following Banach spaces are introduced,

$$\begin{aligned} &\mathcal{B}_1 := \mathcal{C}([0, t_f]; \mathbb{R}^n), \mathcal{B}_2 := \mathcal{C}([0, t_f]; \mathbb{R}^{n \times n}), \\ &\mathcal{B}_3 := \mathcal{C}([0, t_f]; \mathbb{R}^n), \end{aligned}$$

Algorithm 1 Iterative Method

Input: Reference trajectory $z(t)$, parameters Q, R, F
Output: optimal solution $u^*(t)$, $x^*(t)$
Initialization : initial state trajectory $x_0(t)$, $k=1$
while $\|x_k - x_{k-1}\| > \epsilon$ **do**
compute $P_{k+1}(t), g_{k+1}(t)$ by (11);
calculate $x_{k+1}(t)$, $u_{k+1}(t)$;
$k = k+1$;
end while
return x^*, u^*, P^*, g^*, h^*

with the norms defined as

$$\|f\|_\alpha := \sup_{t\in[0,t_f]} \|f(t)\| \exp(-\alpha t), f \in \mathcal{B}_1,$$
$$\|P\|_\alpha := \sup_{t\in[0,t_f]} \|P(t)\| \exp(-\alpha(t_f - t)), P \in \mathcal{B}_2,$$
$$\|g\|_\alpha := \sup_{t\in[0,t_f]} \|g(t)\| \exp(-\alpha(t_f - t)), g \in \mathcal{B}_3,$$

where $\|\cdot\|$ denoting the Frobenius norm, i.e., $\|f\| = (\sum_i f_i^2)^{\frac{1}{2}}$, $\|P\| = (\sum_{ij} P_{ij}^2)^{\frac{1}{2}}$ and $\|g\| = (\sum_i g_i^2)^{\frac{1}{2}}$, and the parameter α serves as an additional degree of freedom to control the rate of convergence [3], whose utility is explained in [14]. We introduce the operators $T_1 : \mathcal{B}_1\times\mathcal{B}_2\times\mathcal{B}_3 \to \mathcal{B}_1$, $T_2 : \mathcal{B}_1\times\mathcal{B}_2\times\mathcal{B}_3 \to \mathcal{B}_2$, $T_3 : \mathcal{B}_1\times\mathcal{B}_2\times\mathcal{B}_3 \to \mathcal{B}_3$ to characterize the dynamics of x, P and g, as described in (8) and (10),

$$\begin{aligned}
\frac{d}{dt}T_1[x,P,g] &= [A - E(x)T_2(x,P,g)]T_1(x,P,g) \\
&\quad + E(x)T_3(x,P,g) + \bar{f}, \\
\frac{d}{dt}T_2[x,P,g] &= -T_2(x,P,g)A - A'T_2(x,P,g) \\
&\quad + T_2(x,P,g)E(x)T_2(x,P,g) - D, \\
\frac{d}{dt}T_3[x,P,g] &= -[A' - T_2[x,P,g]E(x)]T_3[x,P,g] \\
&\quad + T_2[x,P,g]\bar{f} - Wz,
\end{aligned} \tag{12}$$

with the constraints given as $T_1[x,P,g](0) = x_0$, $T_2[x,P,g](T) = P(T)$ and $T_3[x,P,g](T) = g(T)$.

Theorem 1: Consider the tracking problem defined in (2), and the mappings T_1, T_2, and T_3 defined in (12), then Algorithm 1 evolving according to

$$\begin{aligned}
x_{k+1}(t) &= T_1[x_k, P_k, g_k], \\
P_{k+1}(t) &= T_2[x_k, P_k, g_k], \\
g_{k+1}(t) &= T_3[x_k, P_k, g_k],
\end{aligned} \tag{13}$$

is convergent if R is chosen sufficiently large.

Proof: Here, we only provide an outline of the main steps of the proof. Interested readers can refer to [3] for more details on how to use contraction mapping and the fixed-point theorem to prove convergence. To show the convergence, the norm of the difference between consecutive iterations needs to be calculated and shown to be bounded. This requires that the following terms be bounded (see Appendix for details),

$$\begin{aligned}
&\|B_k - B_{k-1}\| \le L\|\xi_{k-1}\|, \\
&\|E_k - E_{k-1}\| \le \gamma_1\|\xi_{k-1}\|, \\
&\|E_k P_{k+1} - E_{k-1}P_k\| \le \gamma_2\|P_{k+1} - P_k\| + \gamma_3\|\xi_{k-1}\|,
\end{aligned} \tag{14}$$

where $\xi_{k-1} = x_k - x_{k-1}$. The first line is a consequence of the Lipschitz continuity condition imposed on $B(x)$, and the parameters $\gamma_1 = L(\|R^{-1}B_k'\| + \|R^{-1}B_{k-1}'\|)$, $\gamma_2 = \|E_k\|$, and $\gamma_3 = \gamma_1\|P_k\|$ are time-varying. Furthermore, we assume that in a compact interval $[0,T]$, $A(x)$ and $\bar{f}(x)$ are continuous and bounded functions, hence we have

$$\begin{aligned}
&\|A_k - A_{k-1}\| \le L_1\|x_k - x_{k-1}\|, \\
&\|\bar{f}_k - \bar{f}_{k-1}\| \le L_2\|x_k - x_{k-1}\|.
\end{aligned} \tag{15}$$

Finally, after combining (14) and (15) into (13), one obtains the following result

$$\begin{bmatrix} \|T_1[x_k,P_k,g_k] - T_1[x_{k-1},P_{k-1},g_{k-1}]\|_\alpha \\ \|T_2[x_k,P_k,g_k] - T_2[x_{k-1},P_{k-1},g_{k-1}]\|_\alpha \\ \|T_3[x_k,P_k,g_k] - T_3[x_{k-1},P_{k-1},g_{k-1}]\|_\alpha \end{bmatrix} \le M\|\xi_{k-1}\|_\alpha.$$

More details are given in the Appendix. The components of M are proportional to R^{-1}, as a result, R can be chosen large enough such that the map T_1 is contractive. Then, the fixed-point theorem guarantees the iterative trajectories x_k converge to the unique fixed point solution x^*. ∎

It should be noted that Theorem 1 requires the boundedness of terms P_k, g_k, h_k, this will be proved in the following subsection.

C. Existence and Optimality of Convergent Solutions

The iterative algorithm will find a converging solution under the assumption that the solution to the coupled set of ordinary differential equations (8) exists and is bounded. In fact, we will show that once the existence and boundedness of P_k has been established, g_k and h_k will follow as bounded.

Theorem 2: If there exists a solution $V(t,x)$ of class C^2 of the HJB equation, which satisfies the following conditions

$$V(T,x) = \phi(T,x), \quad L_{uu}(t,x,u) = \frac{\partial^2 L}{\partial u^2} \succ 0, \forall t,x,u,$$

where $L = e'(t)Q(t)e(t) + u'(t)R(t)u(t)$ is the Lagrangian of the performance index, then V is the optimal performance index for the optimal tracking problem, and the corresponding optimal law is given by (9) [15].

Theorem 3: Consider the value function defined in (6) for the optimal tracking problem presented in (2), and assume that the problem is well-defined. If

(i) $R(t) \succ 0, \forall t \in [0,T]$,
(ii) $D(t) \succeq 0, \forall t \in [0,T]$,
(iii) $F \succeq 0$,

then $V(t,x)$ is well-defined and satisfies the HJB equation. Furthermore, if $V \in C^2(\mathbb{R}\times\mathbb{R}^n)$, then V is the optimal performance index.

Proof: By the existence theorem in [15], the conditions (i)-(iii) (also called Kalman conditions) ensures that for a nonnegative definite $P_k(T)$ at T, the riccati equation in (11) has a unique solution denoted as $P_k(t) = \Pi(t; P_k(T), T)$, for all $t \in (\tau, T]$. It should be noted that if the phenomenon of finite escape time occurs, T can always be chosen small enough to guarantee the existence of the solution. Moreover, at each iteration $k+1$, we are solving the optimal control problem

$$\min_u \ J_{k+1}(u) = \tfrac{1}{2} e'_{k+1}(T) F e_{k+1}(T) + \tfrac{1}{2} \int_0^T e'_{k+1} Q e_{k+1} + u' R u dt, \tag{16}$$

subject to the following state-dependent linear system

$$\dot{x}_{k+1} = A_k x_{k+1} + \bar{f}_k + B_k u,$$

with the value function of the following form

$$V_{k+1} = \frac{1}{2} x'_{k+1} P_{k+1} x_{k+1} - x'_{k+1} g_{k+1} + h_{k+1},$$

where P_{k+1}, g_{k+1} and h_{k+1} satisfy (11). By applying the same technique as the proof in [15], one can show that the following inequality holds for $t \in (\tau, T]$ and all $k \geq 0$,

$$V_{k+1}(t) \leq \tfrac{1}{2} [e'_{k+1}(T) F e_{k+1}(T) + \int_t^T e'_{k+1}(s) Q e_{k+1}(s) ds],$$

where we assume $u \equiv 0$ in the cost function, J_{k+1}, on the right hand side (RHS) of (16). Therefore, one can conclude that P_{k+1} is bounded in the interval $(\tau, T]$, thus exists for all $t \leq T$. Therefore, P_{k+1} is bounded on the compact time interval $[0, T]$ for all the iterations. Hence, $g_{k+1}(t)$ and $h_{k+1}(t)$ also exist, and they are unique and bounded. Consequently, V_{k+1} exists, which solves the HJB equation in each iteration.

By the proof of convergence from Section III-B, one can conclude that $V(t,x) = \lim_{k \to \infty} V_k(t,x)$ exists and satisfies the HJB equation. Furthermore, if V is twice continuously differentiable, we have $L_{uu} = R(t) \succ 0$, then according to Theorem 2, $V(t,x)$ is the optimal performance index for the tracking problem. ■

IV. Numerical Examples

In this section, we present three examples of control-affine nonlinear systems, with each one of them representing a specific class of systems. In particular, the third one represents systems with nonlinear output functions.

Example 1: Consider the dynamics of a nonholonomic integrator (system without drift term) described by

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -x_2 & x_1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix},$$

where the output is given by $y = Cx$ ($C = I$, an $n \times n$ identity matrix). This system, also referred to as the Brockett integrator [5], is partially controllable and it can be shown that starting from the origin, there exists a control, $u^*(t) = (u_1, u_2)'$, that can drive the system to the endpoint $(0, 0, a)'$ in time T. Furthermore, by Potryangin's maximum principle one can analytically derive the optimal control law of the form $u^*(t) = (u_1(0)\cos(2\pi t/T) - u_2(0)\sin(2\pi t/T), u_1(0)\sin(2\pi t/T) + u_2(0)\cos(2\pi t/T))$, with $u_1^2(0) + u_2^2(0) = 2a\pi/T^2$.

To test the performance of our control algorithm, the reference trajectory was generated by applying $u^*(t)$ to steer the system from the origin to $(0, 0, 2)'$. Then, we used our iterative algorithm to compute a tracking control, $u(t)$, that closely tracks the reference while minimizing the cost functional (2). We achieved excellent tracking as shown in Fig. 1(c), with a control cost $\int_0^T u^2(t) dt = 3.057$ that is lower than the energy expended by u^*, which is π.

Note that the system in Example 1 has no drift, hence $A_k = 0$ and $\bar{f}_k = 0$. We observe that introducing a stable matrix $A_k = -I$ and $\bar{f}_k = x_k$ (that cancel each other), improves the convergence rate as shown in Fig. 1(b), in which we refer to this approach as method 1 while method 2 corresponds to the case $A_k = 0$ and $\bar{f}_k = 0$, respectively.

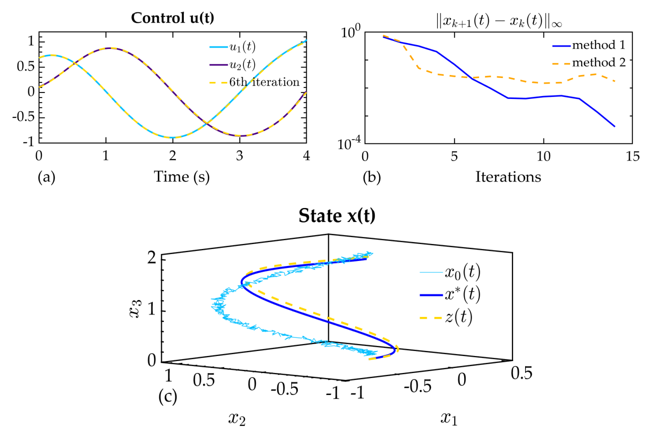

Fig. 1. Brockett integrator system. (a) shows that the convergence of the control in 6 iterations. (b) compares the convergence behavior of method 1 and 2. (c) shows the initial trajectory $x_0(t)$, the reference $z(t)$ and the optimal trajectories $x^*(t)$, respectively. The design parameters are $Q = 100I_3$, $F = 200I_3$ and $R = 9.5I_2$ (with I_n an $n \times n$ identity matrix).

Example 2: The model of a continuous stirred-tank chemical reactor is described by

$$\begin{aligned} \dot{x}_1 &= -(x_1 + 0.25)(2 + u) + (x_2 + 0.5) e^{\frac{25x_1}{x_1+2}}, \\ \dot{x}_2 &= 0.5 - x_2 - (x_2 + 0.5) e^{\frac{25x_1}{x_1+2}}, \end{aligned} \tag{17}$$

where x_1 and x_2 represent the deviation from the steady state temperature and concentration, respectively, and are controlled by the flow of a coolant denoted as u [16]. The goal is to keep the deviation as small as possible without much control effort. For the nonlinear system in (17), in order to apply the algorithm proposed in this paper, the A matrix and $\bar{f}$ vector describing the drift term are written as $A(x) = \begin{pmatrix} -2 & e^{\frac{25x_1}{x_1+2}} \\ 0 & -e^{\frac{25x_1}{x_1+2}} - 1 \end{pmatrix}$, $\bar{f}(x) = \begin{pmatrix} -0.5 + 0.5e^{\frac{25x_1}{x_1+2}} \\ 0.5 - 0.5e^{\frac{25x_1}{x_1+2}} \end{pmatrix}$, and $B(x) = (-x_1 - 0.25, 0)'$.

The simulation of the optimal tracking performance is shown in Fig. 2. The benchmark solution obtained with the algorithm in [16] is presented in Fig. 2(a) and Fig. 2(b) in yellow. From Fig. 2, we observe that one can obtain a satisfying result after 5 iterations with a control that forces the states to be $(-0.0010, -0.0325)'$, which is much closer to the steady value as compared to the benchmark result $(0.0392, -0.0773)'$. The optimal control obtained by our algorithm achieved high accuracy, while expending slightly more energy (0.2535) than the control in [16] (0.2346).

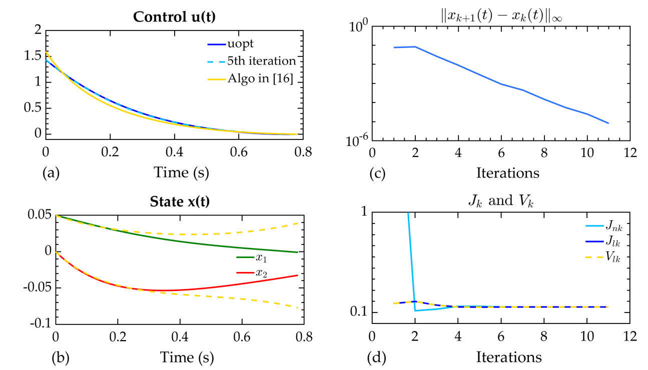

Fig. 2. Chemical reactor control. (a) shows the optimal control u_{opt} and the convergent control obtained by the algorithm in [16]. (b) compares our optimal trajectories (x_1, x_2) to that of [16] (dashed yellow line). (c) shows convergence behavior. (d) cost index J and value function V. J_{lk} and V_{lk} denote the cost index and value function of the linear system, respectively, and J_{nk} the cost index for the original nonlinear system controlled by the computed solution. The design parameters are $Q = 123I_2, R = 0.25, F = 0$.

The preceding examples demonstrate the applicability of the proposed algorithm to nonlinear systems with linear output functions. However, for the general nonlinear system with the output given as $y_i = l_i(x)$, it is sometimes desirable to design output feedback that also considers the nonlinearity, this is, for example, the case of the Pendubot [17], [18]. It appears that many approaches have been proposed for controlling the Pendubot [18], however, no approaches considered the output regulator problem before the work presented in [17].

In the following example, we present how one can apply our proposed algorithm to synthesize the optimal control for a general nonlinear system.

Example 3: We consider a phase model of the form

$$\dot{\theta}(t) = \omega + B(\theta)u(t), \quad y = l(\theta),$$

where, we take $l(\theta) = \sin(\theta)$ and $B(\theta) = 1 - \cos(\theta)$, which is the phase response curve of a SNIPER neuron.

The main steps for solving the optimal tracking control problem associated with this dynamical system remain the same as presented in Section III, however, the value function takes a slightly different form

$$V(t, \theta) = \tfrac{1}{2} l(\theta)P(t)l(\theta) - l(\theta)g(t) + h(t).$$

The necessary equations for computing the optimal control are given below

$$\begin{aligned} \dot{P}(t) &= P^2 \cos^2(\theta)E(\theta) - Q, \\ \dot{g}(t) &= PE(\theta)\cos^2(\theta)g + P\cos(\theta)\omega - Qz, \\ \dot{h}(t) &= \cos(\theta)g\omega - \tfrac{1}{2}zQz + \tfrac{1}{2}\cos^2(\theta)E(\theta)g^2, \end{aligned} \tag{18}$$

where $E(\theta) = B(\theta)R^{-1}B(\theta)$ and the boundary conditions are $P(T) = F$, $g(T) = Fz(T)$ and $h(T) = \frac{1}{2}z(T)Fz(T)$, respectively. Accordingly, the optimal control law and state is $u^*(t) = -R^{-1}B'(\theta^*)(\cos(\theta^*)\sin(\theta^*)P(t) - \cos(\theta^*)g(t))$, $\dot{\theta}^*(t) = \omega - E(\theta^*)(\cos(\theta^*)\sin(\theta^*)P(t) - \cos(\theta^*)g(t))$. Consequently, the algorithm is also changed according to (18). The result is shown in Fig. 3 where the tracking target is set as $z = \sin(2.1t)$, for a phase model with a natural frequency of $\omega = 2$.

This example demonstrates the applicability of the proposed algorithm to a general nonlinear system. However, note that in this case the candidate value function, $V(t, \theta)$, directly involves the output function, $l(\theta)$, instead of the state variable as it was in the case for nonlinear systems with linear output functions. As a result, the equations for computing the optimal control are not exactly the same as before. Nonetheless, one of the main advantages of our method is that, once the iteration equations have been derived, applying our algorithm is straightforward and the convergence is fast.

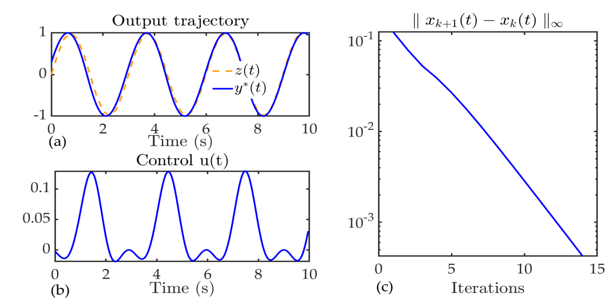

Fig. 3. Control of a dynamical system with a nonlinear output function. (a) shows the reference trajectory (orange) and the controlled optimal trajectory (blue). The optimal control is shown in (b) and convergence behavior in (c). The design parameters are $R = 10$, $Q = 1$ and $F = 10$ with $y^*(0) = 0.3$ and $z(0) = 0$.

V. Conclusion

This paper presents an iterative algorithm for solving optimal tracking control problems of control-affine nonlinear systems, and in particular, the case of a system with a nonlinear output function is explored. It is shown that, regardless of the structure of the system, the iterative procedure remains the same. However, the set of coupled differential equations (derived from the HJB equation) solved at each iteration changes according to the type of system as demonstrated in Section IV. Furthermore, we analyse the convergence of the iterative scheme by using contraction mapping and fixed-point theorem, and provide the conditions under which the convergent solution is the optimal one. The simplicity and easy of implementation of the proposed algorithm allows it to be used in a wide range of applications as illustrated by numerical examples, that also demonstrate its effectiveness.

Appendix

In this section, we show how the components of the matrix M that appears in the proof of convergence in Section III-

B are obtained. By the variation of constants formula, the solution to $x_{k+1}(t)$ in (11) is given as

$$x_{k+1}(t) = \Phi_{k+1}(t,0)x_0 + \int_0^t \Phi_{k+1}(t,\sigma)(E_k(\sigma)g_{k+1}(\sigma) + \bar{f}(\sigma))d\sigma,$$

where $\Phi_{k+1}(t,t_0)$ is the transition matrix of the homogeneous system $\dot{x}_{k+1} = [A_k - E_k P_{k+1}]x_{k+1}$. Now, we proceed to derive the dynamic of the difference $x_{k+1}(t) - x_k(t)$, denoted as $\xi_k(t)$

$$\begin{aligned}\frac{d}{dt}\xi_k &= [A_k - E_k P_{k+1}]\xi_k + [E_{k-1}P_k - E_k P_{k+1}]x_k \\ &\quad + E_{k-1}(g_{k+1} - g_k) + [E_k - E_{k-1}]g_{k+1} \\ &\quad + (A_k - A_{k-1})x_k + \bar{f}_k - \bar{f}_{k-1},\end{aligned}$$

since $\xi_k(0) = 0$, the above immediately admits the integral form

$$\begin{aligned}\xi_k(t) = \int_0^t &\Phi_{k+1}(t,\sigma)\{[E_{k-1}P_k - E_k P_{k+1}]x_k \\ &+ E_{k-1}(g_{k+1} - g_k) + [E_k - E_{k-1}]g_{k+1} \\ &+ (A_k - A_{k-1})x_k + \bar{f}_k - \bar{f}_{k-1}\}(\sigma)d\sigma.\end{aligned}$$

Thus, it follows that

$$\begin{aligned}&\|\xi_k(t)\| \le \int_0^t \|\Phi_{k+1}(t,\sigma)\|\{\|E_{k-1}\|\|g_{k+1} - g_k\| \\ &+ (\gamma_3\|x_k\| + \gamma_1\|g_{k+1}\| + L_1\|x_k\| + L_2)\|x_k - x_{k-1}\| \\ &+ \gamma_2\|x_k\|\|P_{k+1} - P_k\|\}(\sigma)d\sigma,\end{aligned}$$

where the terms $\|E_k P_{k+1} - E_{k-1}P_k\|$, $\|E_k - E_{k-1}\|$, and $\|A_k - A_{k-1}\|$, $\|\bar{f}_k - \bar{f}_{k-1}\|$ are replaced by their respective upper bounds in (14) and (15), respectively. Similarly, with $P_{k+1}(T) = P_k(T)$, and $g_{k+1}(T) = g_k(T)$, we derive the following

$$\begin{aligned}(P_{k+1} - P_k)(t) &= \int_T^t \Phi'_{k+1}(\sigma,t)[P_k(E_k - E_{k-1})P_{k+1} \\ &\quad - P_k(A_k - A_{k-1}) - (A_k - A_{k-1})'P_k]\Phi_{k+1}(\sigma,t)d\sigma \\ (g_{k+1} - g_k)(t) &= \int_T^t \Phi'_{k+1}(t,\sigma)\{[E_k P_{k+1} - E_{k-1}P_k \\ &\quad - (A_k - A_{k-1})]g_k + P_{k+1}\bar{f}_k - P_k\bar{f}_{k-1}\}(\sigma)d\sigma.\end{aligned}$$

After some algebraic manipulation and combining the results into (13), we can derive the inequalities,

$$\begin{aligned}&\|T_2[x_k, P_k, g_k] - T_2[x_{k-1}, P_{k-1}, g_{k-1}]\|_\alpha \\ &\le \delta_1\|x_k - x_{k-1}\|_\alpha, \\ &\|T_3[x_k, P_k, g_k] - T_3[x_{k-1}, P_{k-1}, g_{k-1}]\|_\alpha \\ &\le \delta_2\|x_k - x_{k-1}\|_\alpha + \delta_3\|P_{k+1} - P_k\|_\alpha, \\ &\|T_1[x_k, P_k, g_k] - T_1[x_{k-1}, P_{k-1}, g_{k-1}]\|_\alpha \\ &\le \delta_4\|x_k - x_{k-1}\|_\alpha + \delta_5\|P_{k+1} - P_k\|_\alpha + \delta_6\|g_{k+1} - g_k\|_\alpha,\end{aligned}$$

where δ are defined as

$$\begin{aligned}\delta_1 &\propto \sup\|\Phi_{k+1}(\sigma,T)\|^2(\gamma_1\|P_k\|\|P_{k+1}\| + L_1\|P_k\| \\ &\qquad + L_2\|P_k\|) \\ \delta_2 &\propto \sup\|\Phi_{k+1}(\sigma,T)\|[(\gamma_3 + L_1)\|g_k\| + L_2\|P_k\|] \\ \delta_3 &\propto \sup\|\Phi_{k+1}(\sigma,T)\|(\|\gamma_2\|g_k + \|\bar{f}_k\|) \\ \delta_4 &\propto \sup\|\Phi_{k+1}(t,\sigma)\|(\gamma_3\|x_k\| + \gamma_1\|g_{k+1}\| + L_1\|x_k\| \\ &\qquad + L_2) \\ \delta_5 &\propto \sup\gamma_2\|\Phi_{k+1}(t,\sigma)\|\|x_k\| \\ \delta_6 &\propto \sup\|\Phi_{k+1}(t,\sigma)\|\|E_{k-1}\|\end{aligned}$$

and the components of the matrix M are accordingly defined as $M_1 = \delta_4 + \delta_1\delta_5 + \delta_6(\delta_2 + \delta_1\delta_3)$, $M_2 = \delta_1$, $M_3 = \delta_2 + \delta_1\delta_3$.

References

[1] A. E. Bryson, *Applied optimal control: optimization, estimation and control*. Routledge, 2018.

[2] W. Cebuhar and V. Costanza, “Approximation procedures for the optimal control of bilinear and nonlinear systems,” *Journal of Optimization Theory and Applications*, vol. 43, no. 4, pp. 615–627, 1984.

[3] E. Hofer and B. Tibken, “An iterative method for the finite-time bilinear-quadratic control problem,” *Journal of optimization Theory and applications*, vol. 57, no. 3, pp. 411–427, 1988.

[4] S. Wang and J.-S. Li, “Fixed-endpoint minimum-energy control of bilinear ensemble systems,” in *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 1078–1083.

[5] S. Zeng, “Iterative optimal control syntheses illustrated on the brockett integrator,” in *Proceedings of the 11th IFAC Symposium on Nonlinear Control Systems (NOLCOS)*. IFAC, 2019.

[6] A. V. Rao, “A survey of numerical methods for optimal control,” *Advances in the Astronautical Sciences*, vol. 135, no. 1, pp. 497–528, 2009.

[7] P. Elbert, S. Ebbesen, and L. Guzzella, “Implementation of dynamic programming for n-dimensional optimal control problems with final state constraints,” *IEEE Transactions on Control Systems Technology*, vol. 21, no. 3, pp. 924–931, 2012.

[8] O. Sundström, D. Ambühl, and L. Guzzella, “On implementation of dynamic programming for optimal control problems with final state constraints,” *Oil & Gas Science and Technology–Revue de l'Institut Français du Pétrole*, vol. 65, no. 1, pp. 91–102, 2010.

[9] J. Ruths and J.-S. Li, “A multidimensional pseudospectral method for optimal control of quantum ensembles,” *The Journal of chemical physics*, vol. 134, no. 4, p. 044128, 2011.

[10] J.-S. Li, J. Ruths, T.-Y. Yu, H. Arthanari, and G. Wagner, “Optimal pulse design in quantum control: A unified computational method,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 5, pp. 1879–1884, 2011.

[11] J.-S. Li, I. Dasanayake, and J. Ruths, “Control and synchronization of neuron ensembles,” *IEEE Transactions on automatic control*, vol. 58, no. 8, pp. 1919–1930, 2013.

[12] W. Bomela and J.-S. Li, “An iterative method for computing optimal controls for bilinear quadratic tracking problems,” in *2016 American Control Conference (ACC)*. IEEE, 2016, pp. 2912–2917.

[13] H. Schättler and U. Ledzewicz, *Geometric optimal control: theory, methods and examples*. Springer Science & Business Media, 2012, vol. 38.

[14] L. Collatz, *Differential Equations: An Introduction With Applications*. John Wiley & Sons Inc, 1986.

[15] R. E. Kalman *et al.*, “Contributions to the theory of optimal control,” *Bol. soc. mat. mexicana*, vol. 5, no. 2, pp. 102–119, 1960.

[16] M. Filo and B. Bamieh, “Function space approach for gradient descent in optimal control,” in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 3447–3453.

[17] T.-J. Tarn, P. Sanposh, D. Cheng, and M. Zhang, “Output regulation for nonlinear systems: some recent theoretical and experimental results,” *IEEE Transactions on Control Systems Technology*, vol. 13, no. 4, pp. 605–610, 2005.

[18] M. Zhang and T.-J. Tarn, “Hybrid control of the pendubot,” *IEEE/ASME transactions on mechatronics*, vol. 7, no. 1, pp. 79–86, 2002.