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

Power series solutions to the time-varying dynamic programming equations

Cesar O. Aguilar and Arthur J. Krener

Abstract— **In this paper we construct high-order approximate solutions to the value function and optimal control for a finite-horizon optimal control problem for time-varying discrete-time nonlinear systems. The method consists in expanding the dynamic programming equations (DPE) in a power series, collecting homogeneous polynomial terms and solving for the unknown coefficients from the known and previously computed data. The resulting high-order equations are linear difference equations for the unknown homogeneous terms and are solved backwards in time. The method is applied to construct high-order perturbation controllers around a nominal optimal trajectory.**

I. INTRODUCTION

Consider the time-varying discrete-time control system

$$x_{t+1} = f_t(x_t, u_t) \tag{1}$$

where $f_t : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is smooth, i.e., infinitely differentiable, and $f_t(0,0) = 0$, for $t \in \{0,1,\ldots,\}$. Let $t_0 \in \{0,1,\ldots,\}$, let N be a fixed positive integer, and let there be given smooth functions $\ell_t : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and $\phi : \mathbb{R}^n \to \mathbb{R}$. Define for controlled trajectories satisfying (1)

$$\begin{aligned} \mathbf{x} &= (x_{t_0}, x_{t_0+1}, \ldots, x_{t_0+N}) \\ \mathbf{u} &= (u_{t_0}, u_{t_0+1}, \ldots, u_{t_0+N-1}), \end{aligned}$$

with initial condition $x_{t_0} = x^0$, the cost

$$J_{t_0}(x^0, \mathbf{u}) = \phi(x_{t_0+N}) + \sum_{t=t_0}^{t_0+N-1} \ell_t(x_t, u_t). \tag{2}$$

We will say that the control sequence

$$\mathbf{u}^* = (u^*_{t_0}, u^*_{t_0+1}, \ldots, u^*_{t_0+N-1})$$

solves the *optimal control problem* (1)-(2) if

$$J_{t_0}(x^0, \mathbf{u}^*) \leq J_{t_0}(x^0, \mathbf{u})$$

for all control sequences $\mathbf{u} = (u_{t_0}, u_{t_0+1}, \ldots, u_{t_0+N-1})$. We let $\mathbf{x}^* = (x^*_{t_0}, x^*_{t_0+1}, \ldots, x^*_{t_0+N})$ denote the state trajectory corresponding to $\mathbf{u}^*$. Denote the value function to the optimal control problem as

$$\pi_t(x) = \min_{\mathbf{u}} J_t(x, \mathbf{u})$$

Research performed while the first author held a National Research Council Postdoctoral Award at the Department of Applied Mathematics, Naval Postgraduate School, 833 Dyer Rd., Bldg. 232, Monterey, CA 93943. Research supported in part by AFOSR and NSF. coaguila@nps.edu, ajkrener@nps.edu

where now x is the initial condition at time t. Applying Bellman's dynamic principle [2], the value functions π_t satisfy the recurrence relation

$$\pi_t(x) = \min_u \left[\ell_t(x,u) + \pi_{t+1}(f_t(x,u))\right],$$

with final condition $\pi_{t_0+N}(x) = \phi(x)$. If $u = \alpha_t(x)$ is a minimizing controller, then clearly

$$\pi_t(x) = \ell_t(x, \alpha_t(x)) + \pi_{t+1}(f_t(x, \alpha_t(x))). \tag{DPE1}$$

Assuming that π_t is differentiable for each t (Theorem 2.1), the following necessary condition for a minimum holds:

$$0 = \frac{\partial \ell_t}{\partial u}(x, \alpha_t(x)) + \frac{\partial \pi_{t+1}}{\partial x}(f_t(x, \alpha_t(x)))\frac{\partial f_t}{\partial u}(x, \alpha_t(x)). \tag{DPE2}$$

Equations (DPE1)-(DPE2) are the dynamic programming equations for the optimization problem.

Following the method of Al'brekht [1] (see also [8], [10]), we construct polynomial approximations to π_t and α_t as follows. Let f_t, ℓ_t and ϕ have the following Taylor series expansions:

$$f_t(x,u) = A_t x + B_t u + f_t^{(2)}(x,u) + f_t^{(3)}(x,u) + \cdots \tag{3a}$$

$$\ell_t(x,u) = \tfrac{1}{2}x' Q_t x + x' S_t u + \tfrac{1}{2}u' R_t u + \ell_t^{(3)}(x,u) + \cdots \tag{3b}$$

$$\phi(x) = \tfrac{1}{2}x' P x + \phi^{(3)}(x) + \phi^{(4)}(x) + \cdots \tag{3c}$$

where $Q_t = Q_t' \succeq 0$, $R_t = R_t' \succ 0$, and $P = P' \succ 0$ (prime denotes transposition). The term $f_t^{(d)}(x,u)$ denotes a homogeneous polynomial of order d in the components of (x,u) with coefficients depending on t, and similarly for $\ell_t^{(d)}(x,u)$, $\phi^{(d)}(x)$, etc. We assume that π_t and α_t have Taylor series expansions of the form:

$$\pi_t(x) = \tfrac{1}{2}x' P_t x + \pi_t^{(3)}(x) + \pi_t^{(4)}(x) + \cdots \tag{4a}$$

$$\alpha_t(x) = K_t x + \alpha_t^{(2)}(x) + \alpha_t^{(3)}(x) + \cdots \tag{4b}$$

To compute the homogeneous components of $\pi_t(x)$ and $\alpha_t(x)$, we substitute the expansions (3)-(4) into the DPE, collect terms of the same order and solve for the unknown homogeneous terms of $\pi_t(x)$ and $\alpha_t(x)$. For each $d \geq 1$, (DPE1) is used to solve for the $(d+1)$ order homogeneous term of $\pi_t(x)$ and (DPE2) is used to solve for the d order homogeneous term of $\alpha_t(x)$. As will be seen, for $d \geq 2$, the d order term of $\alpha_t(x)$ vanishes in the $(d+1)$ order equations of (DPE1), resulting in a triangular set of equations for $\pi_t^{(d+1)}(x)$ and $\alpha_t^{(d)}$, thereby simplifying the method substantially. The resulting equations are difference equations involving the previously computed lower order terms of $\pi_t(x)$ and $\alpha_t(x)$ and the known data $f_t(x), \ell_t(x)$ and $\phi(x)$. For $d = 1$, the equations that arise are the familiar linear quadratic regulator equations for the linearized dynamics of (1), i.e., the time-varying discrete Riccati equation [5].

Our high-order approximation method is an extension of the method of Al'brekht [1] for continuous time-invariant nonlinear systems. In [1], a method is used to compute high-order polynomial approximations to the value function and optimal control for the Hamilton–Jacobi–Bellman (HJB) equation and a first order necessary condition for optimality similar to (DPE2). The resulting equations for the coefficients of the homogeneous polynomial terms of the value function and optimal controller are algebraic linear equations. Later, an approach similar to [1] was employed in [10] for continuous time-varying nonlinear systems and a finite horizon optimal control problem. In [10], as the HJB equation is time-varying, the coefficients of the homogeneous terms of the value function and optimal controller are time-varying, resulting in ordinary differential equations for the unknown coefficients. Later in [9], the method of Al'brekht was applied to discrete time-invariant nonlinear systems and the resulting equations are algebraic. Hence, our work can be considered as a natural extension to discrete-time systems of the method in [10] on continuous-time systems.

A natural application of our method is the construction of high-order perturbation controllers around a nominal optimal trajectory, the so-called *neighboring extremal* method [3, Ch. 6] or *perturbation control* [5, Section 2.8]. For the case $d = 1$, our method coincides with the unconstrained neighboring extremal method found in [3]. The neighboring extremal method with state and input constraints has been considered in [6], [7] in the development of fast model predictive control (MPC) laws. In this paper we do not treat state and input constraints. In any case, perturbation controllers can be used to approximate optimal trajectories that are nearby a known pre-computed optimal trajectory. Consequently, perturbation controllers can be used to increase the speed of MPC algorithms by providing a more accurate initial guess to nearby optimal trajectories.

II. EXISTENCE OF SMOOTH SOLUTIONS TO THE DPE

Before describing our algorithm for computing polynomial approximate solutions to the DPE, in this section we show for completeness that, under the standard assumptions in the linear quadratic regulator problem [5], there exist sequences of smooth functions $\pi_{t_0}, \pi_{t_0+1}, \ldots, \pi_{t_0+N-1}$ and $\alpha_{t_0}, \alpha_{t_0+1}, \ldots, \alpha_{t_0+N-1}$ solving (DPE1)-(DPE2).

Theorem 2.1: Consider the nonlinear system (1) and cost function (2). Suppose that ℓ_t, f_t, and ϕ are smooth. Assume that ℓ_t and ϕ vanish along with their first derivatives at $(x, u) = (0, 0)$, and that also $f_t(0, 0) = 0$. Assume further that $R_t = \frac{\partial^2 \ell_t}{\partial u^2}(0, 0)$ are positive definite, $Q_t = \frac{\partial^2 \ell_t}{\partial x^2}(0, 0)$ are positive semi-definite, and $P = \frac{\partial^2 \phi}{\partial x^2}(0)$ is positive semi-definite. Then there exist sequences of smooth functions $\pi_{t_0}, \pi_{t_0+1}, \ldots, \pi_{t_0+N-1}$ and $\alpha_{t_0}, \alpha_{t_0+1}, \ldots, \alpha_{t_0+N-1}$, defined locally about $x = 0$, solving (DPE1)-(DPE2).

Proof. We begin with the case $s = t_0 + N - 1$. Define the function $\Psi_s : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ by

$$\Psi_s(x, u) = \ell_s(x, u) + \pi_{s+1}(f_s(x, u))$$

and recall that $\pi_{s+1} = \pi_{t_0+N} = \phi$ is known. Let $B_s = \frac{\partial f_s}{\partial u}(0, 0)$. From the assumptions that

$$\frac{\partial \ell_s}{\partial u}(0, 0) = 0, \quad \frac{\partial \pi_{s+1}}{\partial x}(0) = 0,$$

it follows that the mapping $\frac{\partial \Psi_s}{\partial u} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ vanishes at $(x, u) = (0, 0)$. Furthermore, the $m \times m$ symmetric matrix $\frac{\partial^2 \Psi_s}{\partial u^2}(0, 0)$ is invertible. Indeed, a direct calculations gives that

$$\frac{\partial^2 \Psi_s}{\partial u^2}(0, 0) = R_s + B_s^T P_{s+1} B_s,$$

which is the sum of the positive definite matrix R_s and the positive semi-definite matrix $B_s^T P_{s+1} B_s$, and therefore is also positive definite. By the Implicit Function Theorem applied to $\frac{\partial \Psi_s}{\partial u}$, there exists an open set $\mathcal{V} \subset \mathbb{R}^n$ containing $x = 0$ and an open set $\mathcal{U} \subset \mathbb{R}^m$ containing $u = 0$, and a unique smooth mapping $\alpha_s : \mathcal{V} \to \mathcal{U}$ such that $\alpha_s(0) = 0$ and $\frac{\partial \Psi_s}{\partial u}(x, \alpha_s(x)) = 0$. In other words,

$$0 = \frac{\partial \ell_s}{\partial u}(x, \alpha_s(x)) + \frac{\partial \pi_{s+1}}{\partial x}(f_s(x, \alpha_s(x)) \frac{\partial f_s}{\partial u}(x, \alpha_s(x)).$$

By continuity of the mapping $(x, u) \mapsto \frac{\partial^2 \Psi_s}{\partial u^2}(x, u)$ and the fact that the set of positive definite matrices is open in the set of symmetric matrices, we have that for $x \in \mathcal{V}$ the matrix $\frac{\partial^2 \Psi_s}{\partial u^2}(x, \alpha_s(x))$ is positive definite (here it may be necessary to shrink $\mathcal{V}$). Hence, it follows that for each fixed $x \in \mathcal{V}$, the mapping $u \mapsto \Psi_s(x, u)$ has a minimum at $u = \alpha_s(x)$. Therefore,

$$\pi_s(x) \triangleq \min_u \Psi_s(x, u) = \ell_s(x, \alpha_s(x)) + \pi_{s+1}(f_s(x, \alpha_s(x)))$$

and it is clear that $\pi_s : \mathcal{V} \to \mathbb{R}$ is smooth. Thus, we have proved that α_s and π_s solve (DPE1)-(DPE2) for $s = t_0 + N - 1$ on $\mathcal{V}$. Now, by classical results regarding the discrete-time linear quadratic regulator problem [5, pg. 63], the assumption that Q_s is positive semi-definite implies that the matrix $P_s = \frac{\partial^2 \pi_s}{\partial x^2}(0)$ is positive semi-definite. We can therefore repeat our arguments above for the mapping $\Psi_{s-1} : \mathcal{V}' \times \mathcal{U}' \to \mathbb{R}$ defined as

$$\Psi_{s-1}(x, u) = \ell_{s-1}(x, u) + \pi_s(f_{s-1}(x, u)),$$

where $\mathcal{V}' \subset \mathcal{V}$ and $\mathcal{U}' \subset \mathcal{U}$ are sufficiently small open sets such that $f_{s-1}(\mathcal{V}', \mathcal{U}') \subset \mathcal{V}$. In this way, we obtain the desired sequences $\pi_{t_0}, \pi_{t_0+1}, \ldots, \pi_{t_0+N-1}$ and $\alpha_{t_0}, \alpha_{t_0+1}, \ldots, \alpha_{t_0+N-1}$, and this completes the proof. ■

III. POWER SERIES SOLUTION TO THE DPE

In this section we describe our algorithm for computing the homogeneous polynomial terms of $\pi_t(x)$ and $\alpha_t(x)$ order-by-order from the DPE.

A. Order $d = 1$: Computing P_t and K_t

Substituting the power series expansions (3)-(4) into the DPE and collecting the quadratic terms from (DPE1) and the linear terms from (DPE2) yield the familiar equations from the discrete-time linear quadratic regulator problem:

$$\begin{aligned}\tfrac{1}{2}x'P_t x = \tfrac{1}{2}x'[&Q_t + 2S_tK_t + K_t'R_tK_t \\ &+ (A_t + B_tK_t)'P_{t+1}(A_t + B_tK_t)]x \end{aligned} \tag{5a}$$

$$0 = x'[S_t + K_t'R_t + (A_t + B_tK_t)'P_{t+1}B_t]. \tag{5b}$$

As (5b) holds for all x, it follows that

$$K_t = -(R_t + B_t'P_{t+1}B_t)^{-1}(S_t + A_t'P_{t+1}B_t)'. \tag{6}$$

Substituting (6) into (5a) and simplifying yields the time-varying discrete Riccati equation (DRE)

$$P_t = Q_t + A_t'P_{t+1}A_t - \Gamma_t(R_t + B_t'P_{t+1}B_t)^{-1}\Gamma_t' \tag{7}$$

where $\Gamma_t = (S_t + A_t'P_{t+1}B_t)$. The DRE is solved backwards from $t = t_0 + N$ to $t = t_0$ with known final condition $P_{t_0+N} = P$.

B. Order $d = 2$: Computing $\pi_t^{(3)}$ and $\alpha_t^{(2)}$

Assume we have computed P_t and K_t, and let $F_t = A_t + B_tK_t$ denote the closed-loop matrices. Collecting cubic terms in (DPE1) yields

$$\begin{aligned}\pi_t^{(3)}(x) &= \ell_t^{(3)}(x, K_tx) + \pi_{t+1}^{(3)}(F_tx) \\ &\quad + x'F_t'P_{t+1}f_t^{(2)}(x, K_tx) \\ &\quad + x'\underbrace{[S_t + K_t'R_t + (A_t + B_tK_t)'P_{t+1}B_t]}_{=0 \text{ from (5b)}}\alpha_t^{(2)}(x) \\ &= \pi_{t+1}^{(3)}(F_tx) + \ell_t^{(3)}(x, K_tx) \\ &\quad + x'F_t'P_{t+1}f_t^{(2)}(x, K_tx).\end{aligned}$$

Therefore, we obtain the following recurrence relation for $\pi_t^{(3)}(x)$:

$$\pi_t^{(3)}(x) = \pi_{t+1}^{(3)}(F_tx) + W_t^{(3)}(x) \tag{8}$$

where

$$W_t^{(3)}(x) = \ell_t^{(3)}(x, K_tx) + x'F_t'P_{t+1}f_t^{(2)}(x, K_tx).$$

Notice that $W_t^{(3)}(x)$ depends on the linear part of $\alpha_t(x)$ and on the quadratic part of $\pi_{t+1}(x)$, which have already been computed by assumption. The recurrence relation (8) is solved backwards from $t = t_0 + N$ to $t = t_0$ with known final condition $\pi_{t_0+N}^{(3)}(x) = \phi^{(3)}(x)$.

Collecting quadratic terms in (DPE2) yields

$$\begin{aligned}0 = \alpha_t^{(2)}(x)'R_t &+ \frac{\partial \ell_t^{(3)}}{\partial u}(x, K_tx) + \frac{\partial \pi_{t+1}^{(3)}}{\partial x}(F_tx)B_t \\ &+ [B_t\alpha_t^{(2)}(x) + f_t^{(2)}(x, Kx)]'P_{t+1}B_t \\ &+ x'(A_t + B_tK_t)'P_{t+1}\frac{\partial f_t^{(2)}}{\partial u}(x, K_tx).\end{aligned}$$

Therefore, we can solve for $\alpha_t^{(2)}$ once $\pi_{t+1}^{(3)}$ is known:

$$\alpha_t^{(2)}(x) = -(R_t + B_t'P_{t+1}B_t)^{-1}V_t^{(2)}(x)' \tag{9}$$

where

$$\begin{aligned}V_t^{(2)}(x) = \frac{\partial \ell_t^{(3)}}{\partial u}(x, K_tx) &+ \frac{\partial \pi_{t+1}^{(3)}}{\partial x}(F_tx)B_t \\ &+ x'(F_t)'P_{t+1}\frac{\partial f_t^{(2)}}{\partial u}(x, K_tx) \\ &+ f_t^{(2)}(x, Kx)'P_{t+1}B_t.\end{aligned}$$

Notice that $V_t^{(2)}(x)$ depends on the linear part of $\alpha_t(x)$ and on up to the cubic part of $\pi_{t+1}(x)$, which have already been computed.

C. Order $d \geq 2$: Computing $\pi_t^{(d+1)}$ and $\alpha_t^{(d)}$

Consider now the general case $d \geq 2$. Hence, assume that we have computed $\pi_t(x)$ up to degree d and $\alpha_t(x)$ up to degree $d - 1$. Collecting $d + 1$ order terms from (DPE1) yields the following expression for $\pi_t^{(d+1)}(x)$:

$$\begin{aligned}\pi_t^{(d+1)}(x) &= \pi_{t+1}^{(d+1)}(F_tx) + W_t^{(d+1)}(x) \\ &\quad + x'\underbrace{[S_t + K_t'R_t + (A_t + B_tK_t)'P_{t+1}B_t]}_{=0 \text{ from (5b)}}\alpha_t^{(d)}(x)\end{aligned} \tag{10}$$

where $W_t^{(d+1)}(x)$ is a homogeneous polynomial in x of degree $d + 1$ depending on $\pi_{t+1}(x)$ up to degree d and on $\alpha_t(x)$ up to degree $d-1$, which have already been computed by assumption. We therefore obtain the following recurrence relation for $\pi_t^{(d+1)}(x)$:

$$\pi_t^{(d+1)}(x) = \pi_{t+1}^{(d+1)}(F_tx) + W_t^{(d+1)}(x). \tag{11}$$

The recurrence relation (11) is solved backwards from $t = t_0 + N$ to $t = t_0$ with known final condition $\pi_{t_0+N}^{(d+1)}(x) = \phi^{(d+1)}(x)$.

Next, collecting d order terms from (DPE2) we obtain an expression of the form

$$0 = \alpha_t^{(d)}(x)'(R_t + B_t'P_{t+1}B_t) + V_t^{(d)}(x)$$

where $V_t^{(d)}(x)$ is a homogeneous polynomial in x of degree d depending on π_{t+1} up to degree $d + 1$ and on α_t up to degree $d - 1$. Therefore, we can solve for $\alpha_t^{(d)}(x)$ because $\pi_{t+1}^{(d+1)}$ has already been computed from (11):

$$\alpha_t^{(d)}(x) = -(R_t + B_t'P_{t+1}B_t)^{-1}V_t^{(d)}(x)'. \tag{12}$$

In this way, for a desired order M, the above procedure produces a polynomial approximation to $\pi_t(x)$ of order $M + 1$ and a polynomial approximation to $\alpha_t(x)$ of order M, for $t = t_0, \ldots, t_0 + N - 1$.

Remark 3.1: It is worth emphasizing the importance of (5b) in the computation of $\pi_t^{(d+1)}(x)$ and $\alpha_t^{(d)}(x)$ for $d \geq 2$. As one can observe from (10), the relation (5b) eliminates $\alpha_t^{(d)}(x)$ from the equation for $\pi_t^{(d+1)}(x)$, thereby resulting in a triangular set of equations for $\pi_t^{(d+1)}(x)$ and $\alpha_t^{(d)}(x)$.

Remark 3.2: As can be seen from (11), the computation of $\pi_t^{(d+1)}(x)$ involves only the *evaluation* of the known and previously computed data, i.e, $W_t^{(d+1)}(x)$, and $\pi_{t+1}^{(d+1)}(x)$. The computational work for performing these calculations can be carried out efficiently by using matrix representations of homogeneous polynomials, as opposed to performing symbolic computations.

IV. PERTURBATION CONTROLLERS AROUND A NOMINAL OPTIMAL TRAJECTORY

The method of the previous section can be used to construct perturbation controllers around a nominal optimal trajectory for discrete nonlinear systems of the form (1) and cost function (2). Such perturbation controllers can be used to increase the speed of model predictive controllers (MPC) [6], [7] by providing more accurate initial guesses to nonlinear programming solvers. In the MPC formulation, the terminal cost function ϕ can be chosen to ensure closed-loop stability of the resulting MPC feedback [4].

In this section we construct time-varying systems, and the associated cost function, describing the perturbed dynamics from a pre-computed optimal trajectory. Our high-order method can then be used on the perturbed dynamics to compute approximations to optimal trajectories nearby the pre-computed optimal trajectory.

Let $(\mathbf{x}^*, \mathbf{u}^*)$ be an optimal trajectory starting at time $t = t_0$ with initial condition x^0. Using the method of Lagrange multipliers [5], we augment the constraints (1) to the cost (2), yielding the Hamiltonian function

$$\begin{aligned} H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) &= \phi(x_{t_0+N}) + \lambda_{t_0}^T (x^0 - x_{t_0}) \\ &+ \sum_{t=t_0}^{t_0+N-1} \ell_t(x_t, u_t) + \lambda_{t+1}^T (f_t(x_t, u_t) - x_{t+1}) \end{aligned} \tag{13}$$

where $\boldsymbol{\lambda} = (\lambda_{t_0}, \lambda_{t_0+1}, \dots, \lambda_{t_0+N})$ are the undetermined *Lagrange multipliers*. For convenience, define

$$h_t(x, u, \lambda) = \ell_t(x, u) + \lambda^T f_t(x, u).$$

Re-arranging the expression for H so that the x_t's are lumped together, H can be written as

$$\begin{aligned} H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}) &= \phi(x_{t_0+N}) - \lambda_{t_0+N}^T x_{t_0+N} + \lambda_{t_0}^T x^0 \\ &+ \sum_{t=t_0}^{t_0+N-1} h_t(x_t, u_t, \lambda_{t+1}) - \lambda_t^T x_t. \end{aligned}$$

The necessary first order condition for $(\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*)$ to be a minimizing triple for the Hamiltonian H can be decomposed into the equations

$$0 = \frac{\partial \ell_t}{\partial x}(x_t^*, u_t^*) + \lambda_{t+1}^{*T} \frac{\partial f_t}{\partial x}(x_t^*, u_t^*) - \lambda_t^{*T} \tag{14a}$$

$$0 = \frac{\partial \phi}{\partial x}(x_{t_0+N}^*) - \lambda_{t_0+N}^{*T} \tag{14b}$$

$$0 = \frac{\partial \ell_t}{\partial u}(x_t^*, u_t^*) + \lambda_{t+1}^{*T} \frac{\partial f_t}{\partial u}(x_t^*, u_t^*) \tag{14c}$$

$$0 = x^0 - x_{t_0}^* \tag{14d}$$

$$0 = f_t(x_t^*, u_t^*) - x_{t+1}^* \tag{14e}$$

for $t = t_0, \dots, t_0 + N - 1$. Now define the mappings $\tilde{f}_t : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ by

$$\tilde{f}_t(\tilde{x}, \tilde{u}) = f_t(x_t^* + \tilde{x}, u_t^* + \tilde{u}) - f_t(x_t^*, u_t^*)$$

For a controlled trajectory $(\bar{\mathbf{x}}, \bar{\mathbf{u}})$ of (1), define the perturbed state $\tilde{\mathbf{x}} = \bar{\mathbf{x}} - \mathbf{x}^*$ and perturbed control $\tilde{\mathbf{u}} = \bar{\mathbf{u}} - \mathbf{u}^*$. Then it is easy to see that the pair $(\tilde{\mathbf{x}}, \tilde{\mathbf{u}})$ satisfies

$$\tilde{x}_{t+1} = \tilde{f}_t(\tilde{x}_t, \tilde{u}_t).$$

Define the function $\widetilde{H} : (\mathbb{R}^n)^{N+1} \times (\mathbb{R}^m)^N \times (\mathbb{R}^n)^{N+1} \to \mathbb{R}$ by

$$\widetilde{H}(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}, \boldsymbol{\eta}) = H(\mathbf{x}^* + \tilde{\mathbf{x}}, \mathbf{u}^* + \tilde{\mathbf{u}}, \boldsymbol{\lambda}^* + \boldsymbol{\eta}) - H(\mathbf{x}^*, \mathbf{u}^*, \boldsymbol{\lambda}^*)$$

and functions $\tilde{\phi} : \mathbb{R}^n \to \mathbb{R}$ and $\tilde{\ell}_t : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ by

$$\tilde{\phi}(\tilde{x}) = \phi(x_{t_0+N}^* + \tilde{x}) - \lambda_{t_0+N}^{*T} \tilde{x} - \phi(x_{t_0+N}^*)$$

$$\begin{aligned} \tilde{\ell}_t(\tilde{x}, \tilde{u}) &= \ell_t(x_t^* + \tilde{x}, u_t^* + \tilde{u}) - \ell_t(x_t^*, u_t^*) \\ &+ \lambda_{t+1}^{*T} \tilde{f}_t(\tilde{x}, \tilde{u}) - \lambda_t^{*T} \tilde{x}. \end{aligned}$$

It it straightforward to verify that

$$\begin{aligned} \widetilde{H}(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}, \boldsymbol{\eta}) &= \tilde{\phi}(\tilde{x}_{t_0+N}) + \sum_{t=t_0}^{t_0+N-1} \tilde{\ell}_t(\tilde{x}_t, \tilde{u}_t) \\ &- \eta_{t_0}^T \tilde{x}_{t_0} + \sum_{t=t_0}^{t_0+N-1} \eta_{t+1}^T (\tilde{f}_t(\tilde{x}_t, \tilde{u}_t) - \tilde{x}_{t+1}). \end{aligned}$$

Hence, $\widetilde{H}$ is the Hamiltonian obtained by adjoining the dynamical constraints

$$\tilde{x}_{t+1} = \tilde{f}_t(\tilde{x}_t, \tilde{u}_t) \tag{15}$$

to the cost function

$$\tilde{J}(\tilde{x}^0, \tilde{\mathbf{u}}) = \tilde{\phi}(\tilde{x}_{t_0+N}) + \sum_{t=t_0}^{t_0+N-1} \tilde{\ell}_t(\tilde{x}_t, \tilde{u}_t) \tag{16}$$

where $\tilde{x}_{t_0} = \tilde{x}^0$. From (14b), the function $\tilde{\phi}$ has a Taylor expansion about $\tilde{x} = 0$ beginning with quadratic terms. Similarly, by (14a) and (14c), $\tilde{\ell}_t$ has a Taylor expansion about $(\tilde{x}, \tilde{u}) = (0, 0)$ beginning with quadratic terms, for $t = t_0, \dots, t_0 + N - 1$.

By construction, if $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*)$ is an optimal controlled trajectory for the time-varying system (15) and cost (16) with initial condition $\tilde{x}_{t_0}^* = \tilde{x}^0$, then $\bar{\mathbf{x}}^* = \mathbf{x}^* + \tilde{\mathbf{x}}^*$ and $\bar{\mathbf{u}}^* = \mathbf{u}^* + \tilde{\mathbf{u}}^*$ is an optimal controlled trajectory for the original system (1) and cost (2) with initial condition $\bar{x}^0 = x^0 + \tilde{x}^0$. The method of the previous section can be employed on the time-varying system (15) with the cost (16) to obtain high-order polynomial approximations to $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*)$, and consequently approximations to $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*)$. In the next section we illustrate the results of this approach with two examples.

V. EXAMPLES

We consider two examples illustrating our method.

Example 5.1: The system evolves in $\mathbb{R}$ and given as

$$x_{t+1} = x_t + \Delta t(\sin(x_t) + u_t), \tag{17}$$

and the cost function is

$$J(x^0, \mathbf{u}) = \tfrac{1}{2} p x_N^2 + \Delta t \sum_{t=0}^{N-1} (\tfrac{1}{2} q x_t^2 + \tfrac{1}{2} r u_t^2) \tag{18}$$

where we set $q = 2$, $r = 1$, and $\Delta t = 0.05$. The scalar $p > 0$ is chosen as the solution to the discrete algebraic Riccati equation arising by considering the linearized dynamics of (17) and the infinite-horizon cost

$$J_\infty(x^0, \mathbf{u}) = \Delta t \sum_{t=0}^{\infty} \left(\tfrac{1}{2} q x_t^2 + \tfrac{1}{2} r u_t^2\right).$$

The value of p is approximately $p \approx 2.85$. The optimal controlled trajectory $(\mathbf{x}^*, \mathbf{u}^*)$ for (17) and cost (18) is pre-computed for initial condition $x^0 = 0.5$ and $N = 75$. We now wish to compute the optimal trajectory $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*)$ for (17)-(18) with initial condition $\bar{x}^0 = 1.5$. As described in §IV, we form the dynamics for the perturbed state $\tilde{\mathbf{x}} = \bar{\mathbf{x}} - \mathbf{x}^*$ and perturbed control $\tilde{\mathbf{u}} = \bar{\mathbf{u}} - \mathbf{u}^*$, resulting in a time-varying nonlinear system of the form (15). Using the power series method described in §III, we computed approximations of orders 1-5 for the optimal controlled trajectory $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*)$ with initial condition $\tilde{x}^0 = \bar{x}^0 - x^0 = 1$. In Fig. 1, we plot the state error $\bar{\mathbf{x}}^* - (\mathbf{x}^* + \tilde{\mathbf{x}}^*)$ and the control error $\bar{\mathbf{u}}^* - (\mathbf{u}^* + \tilde{\mathbf{u}}^*)$ using the approximations to $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*)$ of orders 1-5. The optimal trajectory $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*)$ was computed using Matlab's nonlinear solver `fminsearch` using as an initial guess $\bar{\mathbf{u}}^* \approx \mathbf{u}^* + \tilde{\mathbf{u}}^*$ with $\tilde{\mathbf{u}}^*$ approximated with the 5th order approximation. We remark that the Matlab function `fminsearch` failed to converge in computing $\bar{\mathbf{u}}^*$ using the approximations of $\tilde{\mathbf{u}}^*$ of orders $1-4$.

Example 5.2: In this example we illustrate the method on the pendulum-cart system. The system consists of a cart of mass m_c that is free to move horizontally and acted upon a horizontal force u. The pendulum rod is pivoted at the center of mass of the cart and free to swing in a vertical plane about its frictionless pivot point. The center of mass of the pendulum is a distance l from its pivot point and has mass m_p. For simplicity, we only consider the dynamics of the pendulum and ignore the cart. Applying Newton's laws, the dynamical equation for the pendulum rod is

$$\ddot{\theta} = \frac{\frac{g}{l}\sin(\theta) - \frac{1}{2} m_r \dot{\theta}^2 \sin(2\theta) - \frac{m_r}{ml}\cos(\theta) u}{\frac{4}{3} - m_r \cos^2(\theta)} \tag{19}$$

where θ is the angle the pendulum makes with the vertical, $m_r = \frac{m_p}{m_c + m_p}$, and $g = 9.8$ m/s^2 is the acceleration due to gravity. We take the values $m_p = 2$ kg, $m_c = 8$ kg, and $l = 0.5$ m. Let $x = (\theta, \dot{\theta})$ and let $F(x, u) \in \mathbb{R}^2$ denote the controlled vector field resulting by writing (19) as a first order system. The Eulerian discretization of (19) yields

$$x_{t+1} = f(x_t, u_t) = x_t + \Delta t F(x_t, u_t) \tag{20}$$

where Δt is the sampling interval, $x_t = (\theta(t\Delta t), \dot{\theta}(t\Delta t))$ is the state vector, and $u_t = u(t\Delta t)$ is the control force, for $t = 0, 1, \ldots,$. We take the value $\Delta t = 0.05$. As cost function we take

$$J(x^0, \mathbf{u}) = \tfrac{1}{2} x_N' P x_N + \Delta t \sum_{t=0}^{N-1} \left(\tfrac{1}{2} x_t' Q x_t + \tfrac{1}{2} u_t' R u_t\right) \tag{21}$$

where $Q = \mathrm{diag}(q_{11}, q_{22})$ is positive definite and R is a positive scalar. As in the previous example, the matrix P is chosen as the solution to the discrete algebraic Riccati equation arising by considering the linearized dynamics of (20) and the cost

$$J_\infty(x^0, \mathbf{u}) = \Delta t \sum_{t=0}^{\infty} \left(\tfrac{1}{2} x_t' Q x_t + \tfrac{1}{2} u_t' R u_t\right). \tag{22}$$

The optimal controlled trajectory $(\mathbf{x}^*, \mathbf{u}^*)$ for (20) and cost (21) is pre-computed for initial condition $x^0 = (-0.7, -0.5)$ and $N = 25$. We now wish to compute the optimal trajectory $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*)$ for (20)-(21) with initial condition $\bar{x}^0 = (-0.9, -0.6)$. We form the dynamics for the perturbed state $\tilde{\mathbf{x}} = \bar{\mathbf{x}} - \mathbf{x}^*$ and perturbed control $\tilde{\mathbf{u}} = \bar{\mathbf{u}} - \mathbf{u}^*$, and computed approximations of orders 1-4 for the optimal controlled trajectory $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*)$ with initial condition $\tilde{x}^0 = \bar{x}^0 - x^0 = (-0.2, -0.1)$. In Fig. 2, we plot the Euclidean norm of the state error $\bar{\mathbf{x}}^* - (\mathbf{x}^* + \tilde{\mathbf{x}}^*)$ and the control error $\bar{\mathbf{u}}^* - (\mathbf{u}^* + \tilde{\mathbf{u}}^*)$ using the approximations to $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*)$ of orders 1-4. The optimal trajectory $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*)$ was computed using Matlab's nonlinear solver `fminsearch` using as an initial guess $\bar{\mathbf{u}}^* \approx \mathbf{u}^* + \tilde{\mathbf{u}}^*$ with $\tilde{\mathbf{u}}^*$ approximated with the 4th order approximation. In Table I, we show the computational time and the number of Newton iterations required to compute the optimal trajectory $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*)$ using the approximations of orders 1-4 as initial guesses to the nonlinear solver. The first row in Table I corresponds to using the control sequence $u_t = K x_t$ as an initial guess, where K is the optimal gain for the linearized dynamics of (20) and cost (22), and the second row corresponds to using the previously computed control $\mathbf{u}^*$ as the initial guess. All computations were done on a computer with a 2 GHz processor and 2 GB of RAM.

TABLE I

COMPUTATIONAL TIME AND NUMBER OF ITERATIONS REQUIRED TO CONVERGENCE TO OPTIMAL SOLUTION $(\bar{\mathbf{x}}^*, \bar{\mathbf{u}}^*)$ FOR EXP. 2

Order	Time [sec]	Newton Iterations	Improvement
Linear	0.0495	5	N/A
Previous	0.0399	4	19.4%
1	0.0393	4	20.6%
2	0.0314	3	36.6%
3	0.0303	3	38.8%
4	0.0217	2	56.2%

VI. CONCLUSIONS AND FUTURE WORKS

In this paper we presented a method for computing high-order approximate solutions to the value function and optimal control for a finite-horizon optimal control problem for time-varying discrete-time nonlinear systems. The method was applied to construct perturbation controllers around a nominal optimal trajectory. Examples were given illustrating the method. A natural direction of future work would consider state and input constraints.

REFERENCES

[1] E.G. Al'brekht *On the optimal stabilization of nonlinear systems*, J. Appl. Math. Mech., 25 (1961), pp. 1254-1266.

[2] Richard Bellman, *Introduction to the Mathematical Theory of Control Processes*, Vol. II, Academic Press, New York, NY, 1971.

[3] Arthur E. Bryson and Yu-Chi Ho, *Applied Optimal Control*, New York, Hemisphere Publishing, 1975.

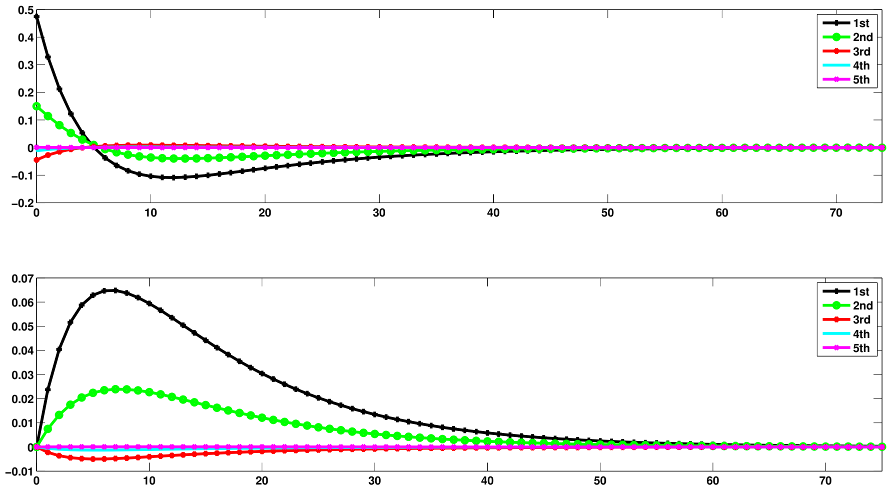

Fig. 1. Error $\bar{\mathbf{u}}^* - (\mathbf{u}^* + \tilde{\mathbf{u}}^*)$ (top) and error $\bar{\mathbf{x}}^* - (\mathbf{x}^* + \tilde{\mathbf{x}}^*)$ (bottom) using approximations to $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*)$ of orders 1-5 for Example 1.

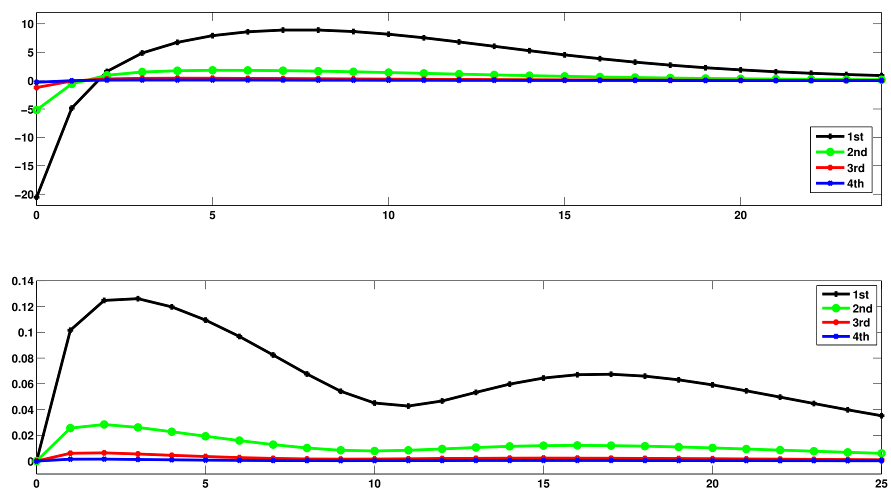

Fig. 2. Error $\bar{\mathbf{u}}^* - (\mathbf{u}^* + \tilde{\mathbf{u}}^*)$ (top) and error $\|\bar{\mathbf{x}}^* - (\mathbf{x}^* + \tilde{\mathbf{x}}^*)\|$ (bottom) using approximations to $(\tilde{\mathbf{x}}^*, \tilde{\mathbf{u}}^*)$ of orders 1-4 for Example 2.

[4] G. De Nicolao, L. Magni, and R. Scattolini, *Stabilizing receding-horizon control of nonlinear time-varying systems*, IEEE Trans. Automat. Control, Vol. 43, No. 7, pp. 1030-1036, 1998.

[5] Frank L. Lewis, *Optimal Control*, John Wiley & Sons, 1986.

[6] R. Ghaemi, J. Sun and I. Kolmanovsky, *An integrated perturbation analysis and Sequential Quadratic programming approach for Model Predictive Control*, Automatica, 45 (2009), pp. 2412-2418.

[7] R. Ghaemi, J. Sun and I. Kolmanovsky, *Neighboring extremal solution for nonlinear discrete-time optimal control problems with state inequality constraints*, IEEE Trans. Automat. Control, Vol. 54, No. 11, pp. 2674-2679, 2009.

[8] D.L. Lukes, *Optimal regulation of nonlinear dynamical systems*, SIAM J. Control and Optimization, 7 (1969), pp. 75-100.

[9] C. Navasca, *Local solutions of the dynamic programming equations and the Hamilton–Jacobi–Bellman PDEs*, PhD Thesis, University of California, Davis, 2002.

[10] A.P. Willemstein *Optimal regulation of nonlinear dynamical systems on a finite interval*, SIAM J. Control Optim., 15 (1977), pp. 1050-1069.