2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Closed-loop Neighboring Extremal Optimal Control Using HJ Equation

Ayush Rai, Shaoshuai Mou, and Brian D. O. Anderson

Abstract— **This study introduces a method to obtain a neighboring extremal optimal control (NEOC) solution for a broad class of nonlinear systems with nonquadratic performance indices by investigating the variation to a known closed-loop optimal control law caused by small, known variations in the system parameters or in the performance index. The NEOC solution can formally be obtained by solving a linear partial differential equation similar to those arising in an iterative solution procedure for a nonlinear Hamilton-Jacobi equation. Motivated by numerical procedures for solving such an equation, we also propose a numerical algorithm based on the Galerkin algorithm that uses basis functions to solve the underlying Hamilton-Jacobi equation. This approach allows the determination of the minimum performance index as a function of both the system state and parameters and extends to allow the determination of the adjustment to an optimal control law given a small adjustment of parameters in the system or the performance index, effectively by computing the derivative of the law with respect to those parameters. The validity of the claims and theory is supported by numerical simulations.**

I. INTRODUCTION

Neighboring extremal optimal control (NEOC) is a term referring to a systematic process for modifying an optimal control strategy to accommodate small perturbations in parameters. Such parameters may be in the system itself or in the performance index. Initially, they were typically associated with the initial or terminal conditions in the original optimal control problem. See in particular Breakwell et al.'s work in 1963, marking the initial contribution to the field of neighboring-extremal optimization techniques [1]. The general problem is to develop a neighboring-optimal feedback control scheme, meaning an adjustment to an already known optimal control, for an *open-loop* optimal control problem in which the state variables are subject to initial and possibly terminal constraints. The prevailing method has been to use the second variation (linear-quadratic) theory to minimize the second variation of the performance index; it requires linearization of the system about a nominal trajectory and a 'quadraticization' of the performance index. To this end, a Riccati transformation is often employed, along with a backward sweep method, to calculate linear feedback gains that can handle state variations along the original optimal control path [1]–[3]. In later work, [4], an extension of the above ideas was used to address dynamic optimization problems with variation in the system equation's parameters or the performance index, in contrast to variable initial or terminal states. This paper used a modified backward sweep method to obtain linear feedback laws defining the control variation as a function of the parameter variation, which had to be small to secure the linearity of the laws. The problem context gave rise to the NEOC terminology.

Previous research in this field has largely focused on scenarios where the base optimal control problem generates an open-loop optimal control (even when the adjustment due to variation might be in closed-loop form). Closed-loop feedback however offers a significant advantage over open-loop control in dealing with the inevitable disturbances that can compromise the optimal performance of the system. In this work, we examine what might happen given a small parameter variation when the original optimal control problem has a solution characterized using the Hamilton-Jacobi equation for the optimal performance, and an associated closed-loop feedback law for the optimal control, rather than an open-loop time function dependent on the initial state. We aim to provide a theoretical tool for determining a change in the closed-loop control law resulting from a small perturbation in the parameters. To the best of the authors' knowledge, this is the first investigation in this area and is applicable to a wide variety of problems.

The article is divided into 7 sections. Section II aims to provide an overview of the problem formulation, by explaining the problem and high-level aspects of the solution using a Hamilton-Jacobi equation and optimal control law formula associated with the original unperturbed problem. Because the solution to the problem of interest can be interpreted as a type of variation on a single iteration in one particular approach to solving a Hamilton-Jacobi equation, the solution of the unperturbed problem and its solution via this approach is presented in Section III, while Section IV covers the perturbed problem, with the previous section providing some kind of a template for the solution to the problem of interest. To this point, the solutions are all analytic or contained in formulas such as integrals over a semi-infinite time interval of trajectories of a known system. An actual numerical approach to solving the problem of interest using the Galerkin algorithm (which has earlier been widely used for Hamilton-Jacobi solution, see e.g. [5]) is detailed in Section V, with illustrative examples provided in Section VI. Finally, the conclusions are provided in Section VII.

Notations: We use the notation $\mathbb{R}^n$ and $\mathbb{R}^{n\times p}$ to denote the set of all $n \times 1$ real vectors and $n \times p$ real matrices, respectively. The transpose of a matrix or vector is denoted by $(\cdot)^\top$. The squared norm of a vector v with respect to

The research is supported in part by a grant from the NASA University Leadership Initiative (80NSSC20M0161) and a gift funding from Northrop Grumman Corporation.

Ayush Rai and Shaoshuai Mou are with School of Aeronautics and Astronautics, Purdue University; {rai29, mous}@purdue.edu

Brian D. O. Anderson is with School of Engineering, The Australian National University, Canberra; brian.anderson@anu.edu.au

979-8-3503-0123-6/23/$31.00 ©2023 IEEE

metric R is denoted by $\|v\|_R^2$, which is equal to $v^\top Rv$. The gradient of a continuous differentiable scalar function $\bar{h}(\cdot)$ with respect to vector $x \in \mathbb{R}^n$ is defined as a column vector denoted by $\nabla_x \bar{h}(\cdot) \in \mathbb{R}^n$, while the derivative of a continuous differentiable vector function $h(\cdot) \in \mathbb{R}^n$, with respect to vector $\alpha \in \mathbb{R}^q$ is defined using a Jacobian matrix $J_{h,\alpha}(\cdot) = [\partial h(\cdot)/\partial \alpha] \in \mathbb{R}^{n\times q}$ whose element in the ith row and jth column is $\partial h(\cdot)_i/\partial \alpha_j$. We use the notation $\{(\cdot)_l\}_{vec}$ to denote a column vector whose lth entry is given by $(\cdot)_l$, whereas we use $\{(\cdot)_l\}_{mat}$ to denote a matrix whose lth column is given by vector $(\cdot)_l$. Lastly, for two any functions ω_1, ω_2, assumed square-integrable on given set Ω, we define the inner product as

$$\langle \omega_1, \omega_2 \rangle_\Omega = \int_\Omega \omega_1(x)\omega_2(x)dx. \tag{1}$$

II. Problem formulation

We set up our problem similar to that of [6], but with a dependence on a set of parameters, call it $\alpha \in \mathbb{R}^q$. We assume that perturbations in α are bounded. With this parameter dependence, the underlying control affine system dynamics is given as

$$\dot{x}(t,\alpha) = f(x,\alpha) + g(x,\alpha)u \quad x(0,\alpha) = x_0(\alpha). \tag{2}$$

where $x \in \Omega \subset \mathbb{R}^n$, $f(x,\alpha) : \Omega \times \mathbb{R}^q \to \mathbb{R}^n$, $g(x,\alpha) : \Omega \times \mathbb{R}^q \to \mathbb{R}^{n\times p}$, and $u(x,\alpha) : \Omega \times \mathbb{R}^q \to \mathbb{R}^p$. We assume $f(\cdot)$ and $g(\cdot)$ are smooth, Lipschitz continuous on Ω that contains origin as an interior point, and the equation has well-defined solutions in Ω for any closed-loop feedback control law $u(x,\alpha)$ of interest. We also assume $f(0,\alpha) = 0$ for any $\alpha \in \mathbb{R}^q$ and that the system is completely controllable, in the sense that given any $(x(t_0,\alpha), t_0)$ and $(x(t_1,\alpha), t_1)$ there exists a smooth control u defined on $[t_0, t_1]$ which will move the first state to the second. We assume that the set Ω is compact, and we restrict attention to trajectories (and associated control laws) which ensure it is invariant, i.e. $x(0,\alpha) \in \Omega$ means $x(t,\alpha) \in \Omega$ for all t.

The performance index we consider, and we are interested in optimization for all initial conditions $x(0,\alpha) \in \Omega$ through a feedback law, is

$$V(x_0, u(\cdot), \alpha) = \lim_{T\to\infty} \int_0^T [\|u(x(t,\alpha),\alpha)\|_R^2 + m(x(t,\alpha),\alpha)]dt \quad \text{s.t.} \quad x(T,\alpha) = 0, \tag{3}$$

where $m(x,\alpha)$ is a smooth, positive definite[1], radially increasing function and R is a positive definite matrix. As shown in [5], it is not enough for a system with a control law u to be stabilizing for the integral (3) to be finite. Therefore, it is necessary to introduce the concept of an *admissible* control law.

Definition 2.1: A control law $u : \Omega \times \mathbb{R}^q \to \mathbb{R}^m$ is considered admissible with respect to the performance index (3) for a given system dynamics (2) if it satisfies the following conditions: it is continuous on Ω, $u(0,\alpha) = 0$, it stabilizes the system (2), ensuring that $x(t,\alpha) \in \Omega\ \forall t$, and it results in a finite integral in (3) for all $x(0,\alpha)$ in Ω.

We define the minimum performance index ϕ as the minimum of the cost function at optimal u as[2]

$$\phi(x,\alpha) = \min_u V(x, u(\cdot), \alpha). \tag{4}$$

From the (steady-state) Hamilton-Jacobi equation [7], for any given α, we know that $\phi(x,\alpha)$ satisfies

$$\begin{aligned}&[\nabla_x \phi(x,\alpha)]^\top f(x,\alpha) + m(x,\alpha) \\ &-\frac{1}{4}[\nabla_x \phi(x,\alpha)]^\top g(x,\alpha) R^{-1} g(x,\alpha)^\top \nabla_x \phi(x,\alpha) = 0,\end{aligned} \tag{5}$$

and the optimal control law, which is provably stabilizing[3] the system (2), i.e. ensures $x(t,\alpha) \to 0$ when $t \to \infty$, is given by

$$u^* = -\frac{1}{2} R^{-1} g(x,\alpha)^\top \nabla_x \phi(x,\alpha). \tag{6}$$

Suppose that the above calculations are done for a specific value, $\alpha = \bar{\alpha}$ say, of α, giving us unperturbed closed-loop feedback. Our task is to say what happens to the optimal performance index ϕ and the control law (6) when α is changed to $\bar{\alpha} + \delta\alpha$ for some small known $\delta\alpha$. The problem can be regarded as one of establishing what the derivatives of ϕ and the optimal control law function are with respect to α.

III. Unperturbed problem

In this section, we begin by recalling one approach to solving the unperturbed problem where the parameter α is set to a fixed constant $\bar{\alpha}$. Since the parameter is fixed, we drop the α dependence for this section.

Approximate solutions of the Hamilton-Jacobi equation can also be obtained, as suggested in [8], by solving an iterative sequence of simpler equations, which in the limit yield the actual solution. (The calculation at a single iteration with the simpler equation serves to inspire our approach to solving the closed-loop NEOC problem.) In a broader sense, such approaches are known as policy iterations [9], [10]. The word ‘policy’ refers to the control law in this case. The architecture of such algorithms involves policy evaluation and policy improvement. Starting with admissible control input, in each iteration, one computes the associated performance index corresponding to the current control law (policy evaluation) and subsequently updates the control law based on the performance index (policy improvement). In greater detail and for the particular form of system and performance index that we are working with, to start the iteration, one supposes the existence of an admissible (but not necessarily optimal) control $u_0 = -\frac{1}{2}R^{-1}g^\top \nabla\phi_0(x)$ for some $\phi_0(x)$. Subsequently, the iterative process ensures

[1] Some relaxation of this assumption to permit some non-negative definite functions is possible, but for convenience in the subsequent analysis, we stay with the positive definiteness assumption.

[2] We assume the existence condition to hold true, meaning that there exists a control law u^* that achieves the minimum of the performance index.

[3] By considering $\phi(x)$ as a Lyapunov function, one can demonstrate that the optimal control obtained from the Hamilton-Jacobi equation is stabilizing.

that at step i, an admissible control law, in the form of $u_i = -\frac{1}{2} g^\top \nabla\phi_i(x)$, is known and obtains the next iterate in a predefined manner. (The linear-quadratic version of this approach is actually known as the Kleinman algorithm, which replaces the solution of a time-invariant Riccati equation by the solution of a sequence of linear matrix equations, [11]). More specifically, the algorithm defines $\phi_{i+1}(x)$ by the linear partial differential equation

$$\begin{aligned}[\nabla_x\phi_{i+1}(x)]^\top [f(x) - \frac{1}{2} g(x) R^{-1} g(x)^\top \nabla_x\phi_i(x)] = \\ -\frac{1}{4}[\nabla_x\phi_i(x)]^\top g(x) R^{-1} g(x)^\top \nabla_x\phi_i(x) - m(x)\end{aligned} \tag{7}$$

Because of the admissibility assumption, this is equivalent to setting

$$\phi_{i+1}(x) = \int_0^\infty \left[\frac{1}{4}[\nabla_y\phi_i(y)]^\top g(y) R^{-1} g(y)^\top \nabla_y\phi_i(y) + m(y)\right] dt \tag{8}$$

where the integration is performed along the trajectory with $y(\cdot)$ defined by

$$\dot{y} = f(y) - \frac{1}{2} g(y) R^{-1} g(y)^\top \nabla_y\phi_i(y) \quad y_i(0) = x. \tag{9}$$

Evidently, $\phi_{i+1}(x)$ represents the value of the performance index when the closed-loop control $u_i(y) = -\frac{1}{2} R^{-1} g^\top(y) \nabla_y\phi_i(y)$ is used and the initial condition is $y(0) = x$.

Additional insight is obtained by considering the change from $\nabla_x\phi_i(x)$ to $\nabla_x\phi_{i+1}(x)$. For this purpose, define

$$\eta_i(x) = [\nabla_x\phi_i(x)]^\top f(x) + m(x) - \frac{1}{4}\left\|R^{-1} g(x)^\top \nabla_x\phi_i(x)\right\|_R^2 \tag{10}$$

The function η_i can be interpreted as an error associated with $\nabla_x\phi_i$ being an approximate rather than exact solution of the steady state Hamilton-Jacobi equation. One straightforwardly obtains

$$\begin{aligned}[\nabla_x\phi_{i+1}(x) - \nabla_x\phi_i(x)]^\top [f(x) - \frac{1}{2} g(x) R^{-1} g(x)^\top \nabla_x\phi_i(x)] \\ = -\eta_i(x)\end{aligned} \tag{11}$$

which is equivalent to

$$\phi_{i+1}(x) - \phi_i(x) = \int_0^\infty \eta_i(y(s)) ds \tag{12}$$

with $y(\cdot)$ as in (9) above. It is possible to show that given an admissible control input as an initial guess, the recursive algorithm outlined in (7) maintains the admissible behavior while also ensuring a monotonically decreasing minimum performance index.

IV. Perturbation Problem

The results obtained in section III give us a nominal closed-loop optimal feedback law, and they rely on repeatedly solving equations like (7), which has a solution obtainable using (8) and (9). For the perturbation problem, a *single* calculation parallel to (7), (8), and (9) can be used. To study the consequence of small perturbation in the parameter, we define a vector function $\xi(x,\alpha) \in \mathbb{R}^q$ by

$$\xi(x,\alpha) = \nabla_\alpha\phi(x,\alpha), \tag{13}$$

which means also that

$$J_{\xi,x}(x,\alpha) = J_{\nabla_\alpha\phi,x}(x,\alpha). \tag{14}$$

with the Jacobian matrix $J_{\xi,x}(x,\alpha) = [\partial\xi(x,\alpha)/\partial x] \in \mathbb{R}^{q\times n}$.

Now simple differentiation of the parameterized steady-state Hamilton-Jacobi equation (5) yields

$$\begin{aligned} & J_{\xi,x}(x,\alpha)\left[f(x,\alpha) - \frac{1}{2} g(x,\alpha) R^{-1} g(x,\alpha)^\top \nabla_x\phi(x,\alpha)\right] \\ & = -J_{f,\alpha}(x,\alpha)^\top \nabla_x\phi(x,\alpha) - \nabla_\alpha m(x,\alpha) \\ & + \frac{1}{2}\left\{\nabla_x\phi(x,\alpha)^\top \frac{\partial g(x,\alpha)}{\partial\alpha_l} R^{-1} g(x,\alpha)^\top \nabla_x\phi(x,\alpha)\right\}_{vec}, \end{aligned} \tag{15}$$

where $\{(\cdot)\}_{vec}$ denotes a column vector whose lth entry is given by $(\cdot)\ \forall\, l = 1,\ldots,q$. This means that formally there holds

$$\begin{aligned}\xi(x,\alpha) = \int_0^\infty \Bigg(& J_{f,\alpha}(y,\alpha)^\top \nabla_y\phi(y,\alpha) + \nabla_\alpha m(y,\alpha) \\ & -\frac{1}{2}\left\{\nabla_y\phi(y,\alpha)^\top \frac{\partial g(y,\alpha)}{\partial\alpha_l} R^{-1} g(y,\alpha)^\top \nabla_y\phi(y,\alpha)\right\}_{vec}\Bigg) dt,\end{aligned} \tag{16}$$

with $y(\cdot)$ defined by

$$\dot{y} = f(y,\alpha) - \frac{1}{2} g(y,\alpha) R^{-1} g(y,\alpha)^\top \nabla_y\phi(y,\alpha);\ y(0,\alpha) = x.$$

The change in optimum performance due to a small change $\delta\alpha$ away from an initial value $\bar{\alpha}$ is evidently given by $\xi(x,\bar{\alpha})^\top \delta\alpha$ and the change in optimal control law is given by adding to the original feedback law the adjustment term

$$\begin{aligned}\delta u = -\frac{1}{2} R^{-1} \Bigg(& g(x,\bar{\alpha})^\top J_{\xi,x}(x,\bar{\alpha})^\top + \\ & \left\{\left[\frac{\partial g(x,\bar{\alpha})}{\partial\alpha_l}\right]^\top \nabla_x\phi(x,\bar{\alpha})\right\}_{mat}\Bigg)\delta\alpha.\end{aligned} \tag{17}$$

where $\{(\cdot)\}_{mat}$ denote a matrix whose lth column is given by $(\cdot)\ \forall\, l = 1,\ldots,q$.

V. Numerical Algorithm

Although we have a solution in the form of equation (16) and (17) that allows us to determine the sensitivity of the optimal performance index and optimal control law to a parameter change, obtaining an analytical or numerical solution can be challenging because it involves solving a linear partial differential equation. Due to the resemblance between the equations encountered in our problem and those encountered during the iterative solution of the HJ equation, we have employed a numerical approach commonly utilized for solving HJ equations in our problem. Successful policy iteration methods for optimal control problems have been implemented to solve HJ equations, including offline neural networks [12], sequential actor-critic neural networks [13], and online actor-critic algorithms [14]. In [5], researchers introduced an iterative algorithm based on the Galerkin spectral approximation method to solve the generalized Hamilton-Jacobi-Bellman equation. This method offers a simple yet powerful approach to obtain the performance index in a functional form using basis functions. In this work, we extend and implement the Galerkin spectral approximation method to obtain the solution to the NEOC problem.

We explain first the original use of the method for an iterative Hamilton-Jacobi equation solution. It is assumed that in the iterative algorithm presented, the ith approximant of the minimum performance index, denoted as $\phi_i(x,\alpha)$, can be represented as a sum of an infinite series of smoothly differentiable linearly independent basis functions $\{\psi_j(x)\}_{j=1}^{\infty}$ along with their respective coefficients $\{w_{ij}(\alpha)\}_{j=1}^{\infty}$ that depend on the parameter α. We select the basis function such that $\phi_i(x,\alpha)$ is in the Hilbert space $L^2(\Omega)$, ensuring square integrable properties.[4] To make the computation feasible, we limit the summation of the infinite series to a finite number of terms which allows us to approximate $\phi_i(x,\alpha)$ to any desired degree of precision, giving us

$$\begin{aligned}\phi_i(x,\alpha) &\approx \sum_{j=1}^{N} w_{ij}(\alpha)\psi_j(x). \\ &= w_i(\alpha)^\top \Psi(x),\end{aligned} \tag{18}$$

where $w_i(\alpha) = [w_{i1}(\alpha),...,w_{iN}(\alpha)]^\top \in \mathbb{R}^N$ and $\Psi(x) = [\psi_1(x),...,\psi_N(x)]^\top \in \mathbb{R}^N$.

The associated optimal control law approximation is

$$u_i(x,\alpha) = -\frac{1}{2}R^{-1}g(x,\alpha)^\top J_{\Psi,x}(x)^\top w_i(\alpha) \tag{19}$$

where $J_{\Psi,x}(x) = [\partial\Psi(x)/\partial x] \in \mathbb{R}^{N\times n}$.

The Galerkin projection method which we now describe constitutes a variant of the earlier iterative procedure; more precisely, assume that one has an expansion $w_i(\alpha)^\top\Psi(x)$ after i steps as an approximation to the optimal performance index; one uses it to define the associated optimal control law approximation (19). Instead of then pursuing directly the calculation of $\phi_{i+1}(x,\alpha)$ as the associated performance index, one determines (with different calculations) an approximation $w_{i+1}(\alpha)^\top\Psi(x)$ to the index, which through appropriate choice of the coefficient vector $w_{i+1}(\alpha)$ is a least squares approximation over the whole set Ω to that index. Since $\phi_{i+1}(x,\alpha)$ in the normal iterative procedure is defined by

$$\nabla_x\phi_{i+1}(x)^\top[f(x,\alpha)+g(x,\alpha)u_i(x,\alpha)] = -\|u_i(x,\alpha)\|_R^2 - m(x,\alpha)$$

this means that $J_{\Psi,x}(x)^\top w_{i+1}(\alpha)$ must be an approximate solution of

$$\begin{aligned}w_{i+1}(\alpha)^\top J_{\Psi,x}(x)[f(x,\alpha)+g(x,\alpha)u_i(x,\alpha)]& \\ \approx -\|u_i(x,\alpha\|_R^2 - m(x,\alpha)&\end{aligned} \tag{20}$$

[4]For more details refer [5].

The best least square approximate solution is obtained by choosing the coefficient vector $w_{i+1}(\alpha)$ to ensure that the error between the left side and the right side is orthogonal to the basis functions, i.e. for each $j = 1,2,\ldots,N$, there holds

$$\begin{aligned}\Big\langle w_{i+1}(\alpha)^\top J_{\Psi,x}(x)[f(x,\alpha)+g(x,\alpha)u_i(x,\alpha)],\psi_j(x)\Big\rangle_\Omega& \\ = -\langle\|u_i(x,\alpha\|_R^2 + m(x,\alpha),\psi_j(x)\rangle_\Omega,&\end{aligned}$$

where the inner product is defined as per (1). There are N linear equations in N unknowns, viz. the entries of $w_{i+1}(\alpha)$. It is established in [5] that the equation set is nonsingular, so that $w_{i+1}(\alpha)$ is well defined.

The algorithm's convergence results as $N\to\infty$ and $i\to\infty$ are presented in [5]. In particular, one can assume that the sequence of iterations for some sufficiently large but fixed N and some $\alpha=\bar{\alpha}$ will converge in practical terms once the iteration number i reaches a value k. Then we can re-write the last equation as

$$\begin{aligned}&\big\langle w_k(\alpha)^\top J_{\Psi,x}(x)f(x,\alpha),\psi_j(x)\big\rangle_\Omega + \langle m(x,\alpha),\psi_j(x)\rangle_\Omega \\ &+\frac{1}{4}\Big\langle\big\|g(x,\alpha)^\top J_{\Psi,x}(x)^\top w_k(\alpha)\big\|_R^2,\psi_j(x)\Big\rangle_\Omega = 0,\end{aligned} \tag{21}$$

for each $j = 1,2,\ldots,N$. These N scalar equations will be the basis for determining the sensitivity of the optimal performance index and associated control law (or more precisely, their approximations $w_k(\alpha)^\top\Psi(x)$ and $-\frac{1}{2}R^{-1}g(x,\alpha)^\top J_{\Psi,x}(x)^\top w_k(\alpha)$) to variation in the parameter α. In particular, by utilizing a calculation akin to that of (15), we differentiate equation (21) with respect to $\alpha\in\mathbb{R}^q$ to involve derivatives of the weighting coefficients with respect to the parameters. The following equation for each $j = 1,2,\ldots,N$, is to be understood as shorthand for q scalar equations obtained by setting $J_{(\cdot),\alpha_{\bar{i}}}$ in place of $J_{(\cdot),\alpha}$ for each $\bar{i} = 1,2,\ldots,q$. The first term of (21) can be differentiated as follows:

$$\begin{aligned}&\big\langle J_{w_k,\alpha}(\alpha)^\top J_{\Psi,x}(x)f(x,\alpha),\psi_j(x)\big\rangle_\Omega \\ &+\big\langle J_{f,\alpha}(x,\alpha)^\top J_{\Psi,x}(x)^\top w_k(\alpha),\psi_j(x)\big\rangle_\Omega.\end{aligned}$$

For the second term of (21), we have

$$\langle\nabla_\alpha m(x,\alpha),\psi_j(x)\rangle_\Omega.$$

Lastly, the derivative of the third term of (21) requires the formation of a tensor and can be simplified to

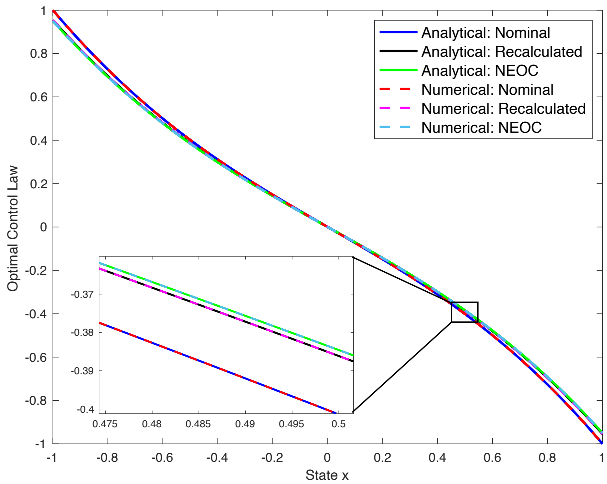
$$\begin{aligned}&-\frac{1}{2}\Big\langle J_{w_k,\alpha}(\alpha)^\top J_{\Psi,x}(x)gR^{-1}g^\top J_{\Psi,x}(x)^\top w_k(\alpha),\psi_j\Big\rangle_\Omega \\ &-\frac{1}{2}\left\{\left\langle w_k(\alpha)^\top J_{\Psi,x}(x)\frac{\partial g}{\partial\alpha_l}R^{-1}g^\top J_{\Psi,x}(x)^\top w_k(\alpha),\psi_j\right\rangle_\Omega\right\}_{vec}.\end{aligned}$$

We observe that this differentiation of (21) yields a linear equation set, and it can be solved for the entries of $J_{w_k,\alpha} = [\partial w_k(\alpha)/\partial\alpha]\in\mathbb{R}^{N\times q}$. These linear equations are provably solvable, as shown in Lemma 14 of [5].

The variation in optimal control law can be obtained using (17) under small perturbation $\delta\alpha$ from a constant value $\bar{\alpha}$, which yields

$$\delta u = -\frac{1}{2}R^{-1}\Bigg(g(x,\bar{\alpha})^\top J_{\Psi,x}(x)^\top J_{w_k,\alpha}(\bar{\alpha}) + \left\{\left[\frac{\partial g(x,\bar{\alpha})}{\partial \alpha_l}\right]^\top J_{\Psi,x}(x)^\top w_k(\alpha)\right\}_{mat}\Bigg)\delta\alpha. \tag{22}$$

VI. Illustrative Examples



(a) NEOC performance

Fig. 1: Example 1- Comparison of the analytical and numerical results for nominal, recalculated, and NEOC optimal control law for $\delta\alpha = 0.2$.

Example 6.1: First, we consider a general class of single-input single-output (i.e. scalar state) systems used in [8] with dynamics

$$\dot{x}(t,\alpha) = f(x,\alpha) + u$$

and a non-quadratic cost function-

$$V = \lim_{T\to\infty}\int_0^T [u^2(x(t),\alpha) + m(x(t),\alpha)]dt \quad \text{s.t.} \quad x(T,\alpha) = 0,$$

where m is positive definite and radially increasing.[5]

In this case, the derivative $\nabla_x\phi(x)$ is a scalar, and the Hamilton Jacobi equation (5) is simply a quadratic equation for $\nabla_x\phi(x)$. One of the two solutions is easily seen to define a stabilizing control law. In particular, we find that

$$\nabla_x\phi_\infty(x,\alpha) = 2f(x,\alpha) + 2\sqrt{f(x,\alpha)^2 + m(x,\alpha)},$$

and the optimal control law is given using $u_\infty(x,\alpha) = -\frac{1}{2}\nabla_x\phi_\infty(x,\alpha)$ as

$$u_\infty(x,\alpha) = -f(x,\alpha) - \sqrt{f(x,\alpha)^2 + m(x,\alpha)}.$$

This provides us with the analytical closed-loop optimal control law, which is the function of the parameters. In case of the presence of perturbation $\delta\alpha$ from the nominal value $\bar{\alpha}$, the optimal solution can be obtained by evaluating $u_\infty(x,\bar{\alpha}+\delta\alpha)$.

[5]Our results are concerned with closed-loop laws, so we can simply regard the initial state as fixed but arbitrary.

Using the NEOC approach, given that $u_\infty(x,\bar{\alpha})$ has already been calculated, in the presence of known perturbation $\delta\alpha$ we add the adjustment term given by (17), which gives us

$$\delta u_\infty(x,\bar{\alpha},\delta\alpha) = -\left(\frac{\partial f}{\partial \alpha} + \frac{2\frac{\partial f}{\partial \alpha} + \frac{\partial m}{\partial \alpha}}{2\sqrt{f(x,\alpha)^2 + m(x,\alpha)}}\right)\Bigg|_{\alpha=\bar{\alpha}}\delta\alpha.$$

This provides the analytical NEOC solution for the perturbation case as

$$u_{NEOC}(x,\bar{\alpha}+\delta\alpha) = u_\infty(x,\bar{\alpha}) + \delta u_\infty(x,\bar{\alpha},\delta\alpha).$$

To assess the effectiveness of the numerical algorithm outlined in Section V, we examine the performance of the algorithm for specific functions f and m, where f is defined as $-\alpha x$ and m is defined as $(1+\alpha)x^2 + x^4$ and the value of α is centered at 1. The corresponding analytical solution is given by

$$u_\infty(x,\alpha) = \alpha x - x\sqrt{1+\alpha+\alpha^2+x^2},$$

and the variation in the optimal control law around the nominal value of $\alpha = \bar{\alpha}$ can be expressed as

$$\delta u_\infty(x,\bar{\alpha},\delta\alpha) = \left(x - \frac{(1+2\bar{\alpha})x}{2\sqrt{1+\bar{\alpha}+\bar{\alpha}^2+x^2}}\right)\delta\alpha.$$

For the numerical algorithm, to approximate the minimum performance index, we utilize the basis function given by

$$\{\psi_j\} = \{x^2, x^4, x^6, x^8, x^{10}\}.$$

For both, the analytical and numerical approaches we compare three optimal control laws: the *nominal* control law which is obtained for $\alpha = \bar{\alpha}$; the *recalculated* optimal control law for the perturbed scenario where we re-run the complete algorithm using $\alpha = \bar{\alpha} + \delta\alpha$; and the *NEOC* law where the nominal control law calculated for $\alpha = \bar{\alpha}$ is updated using a single calculation. To numerically obtain the nominal and recalculated control laws, we execute the Galerkin algorithm for 100 iterations with an initial admissible control input $u_0 = -5x$. The results obtained for the current examples are shown in Fig. 1 for $\delta\alpha = 0.2$. The approximate optimal control laws obtained using NEOC are found to closely match the recalculated optimal control laws for both analytical and numerical cases. In this case, the analytical and numerical results cannot be visually distinguished because we are dealing with a simple example. However, in general, the accuracy of the results depends on the number of basis functions and iterations used to approximate ϕ.

Example 6.2: We next consider a simplified class of bilinear systems used in [15], given by

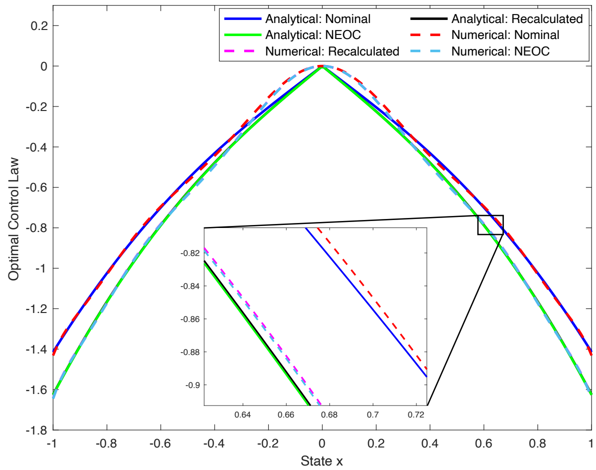
$$\dot{x}(t,\alpha) = g(x,\alpha)u$$

and a non-quadratic cost function-

$$V = \lim_{T\to\infty}\int_0^T [u^2(x(t),\alpha) + m(x(t),\alpha)]dt \quad \text{s.t.} \quad x(T,\alpha) = 0,$$

where m is positive definite and radially increasing.

Similarly to the previous case, the Hamilton-Jacobi equation can be analytically solved, leading to the optimal solution given by



(a) NEOC performace

Fig. 2: Example 2- Comparison of the analytical and numerical results for nominal, recalculated, and NEOC optimal control law for $\delta\alpha = 0.2$.

$$\nabla_x \phi_\infty(x,\alpha) = \frac{2\sqrt{m(x,\alpha)}}{|g(x,\alpha)|}.$$

can be observed using equation (7). The optimal control law is given using $u_\infty(x,\alpha) = -\frac{1}{2}g(x,\alpha)\nabla_x\phi_\infty(x,\alpha)$ as-

$$u_\infty(x,\alpha) = -\operatorname{sgn}(g(x,\alpha))\sqrt{m(x,\alpha)},$$

providing us with the analytical closed-loop optimal law.

Using the NEOC approach, in the presence of a known perturbation $\delta\alpha$, we add the adjustment term given by (17) giving us

$$\delta u_\infty(x,\bar{\alpha},\delta\alpha) = -\operatorname{sgn}(g(x,\bar{\alpha}))\left(\frac{\frac{\partial m(x,\alpha)}{\partial \alpha}}{2\sqrt{m(x,\alpha)}}\right)\Bigg|_{\alpha=\bar{\alpha}}\delta\alpha.$$

To examine the performance numerically, we consider $g(x,\alpha) = \alpha x^2$ and $m(x,\alpha) = \alpha x^2 + \alpha^2 x^4$, where the nominal value of α is set to 1. The analytical solution is given by

$$u_\infty(x,\alpha) = -\sqrt{\alpha x^2 + \alpha^2 x^4},$$

and the variation in the optimal control around the nominal value of α is given by

$$\delta u_\infty(x,\bar{\alpha},\delta\alpha) = -\frac{|x|(1+2\bar{\alpha}x^2)}{2\sqrt{\bar{\alpha}+\bar{\alpha}^2x^2}}\delta\alpha.$$

For the numerical algorithm using the Galerkin method, the identical set of basis functions used in Example 1 is employed. An initial admissible control input of $u_0 = -x^2$ is chosen. The numerical algorithm is run for 100 iterations, and the optimal controls achieved using analytical and numerical methods are compared shown in Fig. 2 for $\delta\alpha = 0.2$. Similar to the previous example, the NEOC solution matches well with the recalculated solution for both analytical and numerical cases, with the latter case being visually indistinguishable. The analytical solution reveals that the optimal control law has a non-smooth point at $x = 0$. Due to this non-smoothness, when the numerical solution is approximated using polynomial basis functions, significant errors are observed around $x = 0$.

VII. Conclusions

The problem of neighboring extremal optimal control (NEOC) is significant in various applications. This study aims to solve the NEOC problem in situations where the nominal solution is a closed-loop feedback law, rather than an open-loop control associated with a particular initial state. Our approach entails analyzing how the perturbation of a parameter affects the optimal control law by studying the variations in the optimal performance index. We obtain a closed-loop adjustment term added to the nominal closed-loop feedback law to achieve neighboring extremal solutions and enhance system performance in the presence of small known parameter variations or perturbations. For numerical implementation on general dynamics, we propose an algorithm that utilizes the Galerkin iterative approach to solve the HJ equations. Various examples are presented to illustrate the algorithm's effectiveness and its ability to handle various scenarios. Our future efforts will focus on: 1) designing efficient online algorithms to solve the NEOC for closed-loop control, and 2) developing a NEOC solution based on a closed-loop feedback law for differential games with performance indices that are not necessarily quadratic.

References

[1] J. V. Breakwell, J. L. Speyer, and A. E. Bryson, "Optimization and control of nonlinear systems using the second variation," *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, vol. 1, no. 2, pp. 193–223, 1963.

[2] C. Hymas, R. CAVIN III, and D. Colunga, "Neighboring extremals for optimal control problems." *AIAA Journal*, vol. 11, no. 8, pp. 1101–1109, 1973.

[3] A. E. Bryson, *Applied optimal control: optimization, estimation, and control*. Routledge, 2018, ch. 4.

[4] A. Lee and A. Bryson, "Neighbouring extremals of dynamic optimization problems with parameter variations," *Optimal Control Applications and Methods*, vol. 10, no. 1, pp. 39–52, 1989.

[5] R. W. Beard, G. N. Saridis, and J. T. Wen, "Galerkin approximations of the generalized Hamilton-Jacobi-Bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159–2177, 1997.

[6] P. Moylan and B. Anderson, "Nonlinear regulator theory and an inverse optimal control problem," *IEEE Transactions on Automatic Control*, vol. 18, no. 5, pp. 460–465, 1973.

[7] B. D. Anderson and J. B. Moore, *Optimal control: linear quadratic methods*. Courier Corporation, 2007.

[8] R. Leake and R.-W. Liu, "Construction of suboptimal control sequences," *SIAM Journal on Control*, vol. 5, no. 1, pp. 54–63, 1967.

[9] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[10] R. A. Howard, "Dynamic programming and Markov processes." 1960.

[11] D. Kleinman, "On an iterative technique for Riccati equation computations," *IEEE Transactions on Automatic Control*, vol. 13, no. 1, pp. 114–115, 1968.

[12] M. Abu-Khalaf and F. L. Lewis, "Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach," *Automatica*, vol. 41, no. 5, pp. 779–791, 2005.

[13] D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. L. Lewis, "Adaptive optimal control for continuous-time linear systems based on policy iteration," *Automatica*, vol. 45, no. 2, pp. 477–484, 2009.

[14] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.

[15] R. W. Beard, *Improving the closed-loop performance of nonlinear systems*. Rensselaer Polytechnic Institute, 1995.