Available online at www.sciencedirect.com
SCIENCE DIRECT®
automatica
Automatica 41 (2005) 779–791
www.elsevier.com/locate/automatica

Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach☆

Murad Abu-Khalaf*, Frank L. Lewis

Automation and Robotics Research Institute, The University of Texas at Arlington, TX 76118, USA

Received 25 November 2002; received in revised form 3 June 2004; accepted 23 November 2004

Abstract

The Hamilton–Jacobi–Bellman (HJB) equation corresponding to constrained control is formulated using a suitable nonquadratic functional. It is shown that the constrained optimal control law has the largest region of asymptotic stability (RAS). The value function of this HJB equation is solved for by solving for a sequence of cost functions satisfying a sequence of Lyapunov equations (LE). A neural network is used to approximate the cost function associated with each LE using the method of least-squares on a well-defined region of attraction of an initial stabilizing controller. As the order of the neural network is increased, the least-squares solution of the HJB equation converges uniformly to the exact solution of the inherently nonlinear HJB equation associated with the saturating control inputs. The result is a nearly optimal constrained state feedback controller that has been tuned a priori off-line.

Keywords: Actuator saturation; Constrained input systems; Infinite horizon control; Hamilton–Jacobi–Bellman; Lyapunov equation; Least squares; Neural network

1. Introduction

The control of systems with saturating actuators has been the focus of many researchers for many years. Several methods for deriving control laws considering the saturation phenomena are found in Saberi, Lin, and Teel (1996), Sussmann, Sontag, and Yang (1994) and Bernstein (1995). However, most of these methods do not consider optimal control laws for general nonlinear systems. In this paper, we study this problem through the framework of the Hamilton–Jacobi–Bellman (HJB) equation appearing in optimal control theory (Lewis & Syrmos, 1995). The solution of the HJB equation is challenging due to its inherently nonlinear nature. For linear systems, the HJB becomes the well-known Riccati equation which results in the linear quadratic regulator (LQR) controller. When the linear system is input constrained, then the closed loop dynamics become nonlinear and the LQR is not optimal anymore.

The HJB equation generally cannot be solved. There has been a great deal of effort to solve this equation. Approximate HJB solutions have been found using many techniques such as those developed by Saridis and Lee (1979), Beard, Saridis, and Wen (1997, 1998), Beard (1995), Murray, Cox, Lendaris, and Saeks (2002), Lee, Teo, Lee, and Wang (2001), Huang, Wang, and Teo (2000), Bertsekas and Tsitsiklis (1996), Munos, Baird, and Moore (1999), Kim, Lewis, and Dawson (2000), Han and Balakrishnan (2000), Liu and Balakrishnan (2000), Lyshevski and Meyer (1995), Lyshevski (1996, 1998, 2001a,b) and Huang and Lin (1995) for HJI equation which is closely related to the HJB equation.

In this paper, we focus on the HJB solution using a sequence of Lyapunov equations developed by Saridis and Lee (1979). Saridis and Lee (1979) successively improve a given initial stabilizing control. This method reduces to the

☆ Supported by NSF ECS-0140490 and ARO DAAD19-02-1-0366. This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor K.L. Teo under the direction of Editor I. Petersen.
* Corresponding author. Tel./fax: +1 817 272 5938.
E-mail addresses: abukhalaf@arri.uta.edu (M. Abu-Khalaf), flewis@arri.uta.edu (F.L. Lewis).

0005-1098/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2004.11.034

well-known Kleinman (1968) iterative method for solving the Algebraic Riccati equation using Lyapunov matrix equations. For nonlinear systems, it is unclear how to solve the LE. Therefore, successful application of the successive approximation method was limited until the novel work of Beard et al. (1997, 1998) where they used a Galerkin spectral approximation method at each iteration to find approximate solutions to the LE and this is called usually GHJB. This requires however the computation of a large number of integrals and it is not obvious how to handle explicit constraints on the controls, which is the interest of this paper.

Lyshevski (1998) proposed the use of nonquadratic functionals to confront constraints on inputs. Using nonquadratic functionals, the HJB equation was formulated and its solution results in a smooth saturated controller. It remains however difficult to actually solve for the value function of this HJB equation.

In this paper, we use neural networks to obtain an approximate solution to the value function of the HJB equation that uses nonquadratic functionals. This in turn results in a nearly optimal constrained input state feedback controller suitable for saturated actuators.

Neural networks applications to optimal control via the HJB equation, Adaptive Critics, were first proposed by Werbos in Miller, Sutton, and Werbos (1990). Parisini and Zoppoli (1998) used neural networks to derive optimal control laws for discrete-time stochastic nonlinear system. Successful neural network controllers have been reported in Chen and Liu (1994), Lewis, Jagannathan, and Yesildirek (1999), Polycarpou (1996), Rovithakis and Christodoulou (1994), Sadegh (1993) and Sanner and Slotine (1991). It has been shown that neural networks can effectively extend adaptive control techniques to nonlinearly parameterized systems. The status of neural network control as of 2001 appears in Narendra and Lewis (2001).

2. Background in optimal control and constrained input systems

Consider an affine in the control nonlinear dynamical system of the form

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

where $x \in \mathbb{R}^n$, $f(x) \in \mathbb{R}^n$, $g(x) \in \mathbb{R}^{n\times m}$. And the input $u \in U$, $U = \{u = (u_1, \ldots, u_m) \in \mathbb{R}^m : \alpha_i \leqslant u_i \leqslant \beta_i,\ i = 1, \ldots, m\}$. where $\alpha_i,\ \beta_i$ are constants. Assume that $f + gu$ is Lipschitz continuous on a set $\Omega \subseteq \mathbb{R}^n$ containing the origin, and that system (1) is stabilizable in the sense that there exists a continuous control on Ω that asymptotically stabilizes the system. It is desired to find u, which minimizes a generalized nonquadratic functional

$$V(x_0) = \int_0^\infty [Q(x) + W(u)]\,\mathrm{d}t. \tag{2}$$

$Q(x)$, $W(u)$ are positive definite, i.e. $\forall x \neq 0\ Q(x) > 0$ and $x = 0 \Rightarrow Q(x) = 0$. For unconstrained control inputs, a common choice for $W(u)$ is $W(u) = u^{\mathrm{T}}Ru$, where $R \in \mathbb{R}^{m\times m}$. Note that the control u must not only stabilize the system on Ω, but guarantee that (2) is finite. Such controls are defined to be *admissible* (Beard et al., 1997).

Definition 1 (*Admissible Controls*). A control u is defined to be admissible with respect to (2) on Ω, denoted by $u \in \Psi(\Omega)$, if u is continuous on Ω, $u(0) = 0$, u stabilizes (1) on Ω, and $\forall x_0 \in \Omega\ V(x_0)$ is finite.

Eq. (2) can be expanded as follows:

$$V(x_0) = \int_0^T [Q(x) + W(u)]\,\mathrm{d}t + \int_T^\infty [Q(x) + W(u)]\,\mathrm{d}t$$
$$= \int_0^T [Q(x) + W(u)]\,\mathrm{d}t + V(x(T)). \tag{3}$$

If $V \in C^1$, then Eq. (3) becomes

$$\lim_{T\to 0}[V(x_0) - V(x(T))]/T = \lim_{T\to 0}\frac{1}{T}\int_0^T [Q(x) + W(u)]\,\mathrm{d}t,$$
$$\dot{V} = V_x^{\mathrm{T}}(f + gu) = -Q(x) - W(u). \tag{4}$$

Eq. (4) is the infinitesimal version of Eq. (2) and is a sort of a Lyapunov equation for nonlinear systems that we refer to as LE in this paper.

$$LE(V,u) \triangleq V_x^{\mathrm{T}}(f + gu) + Q + W(u) = 0,\ \ V(0) = 0. \tag{5}$$

For unconstrained control inputs, $W(u) = u^{\mathrm{T}}Ru$, the LE becomes the well-known HJB equation (Lewis & Syrmos, 1995) on substitution of the optimal control

$$u^*(x) = -\tfrac{1}{2}R^{-1}g^{\mathrm{T}}V_x^* \tag{6}$$

$V^*(x)$ is the value function of the optimal control problem that solves the HJB equation

$$\mathrm{HJB}(V^*) \triangleq V_x^{*\mathrm{T}}f + Q - \tfrac{1}{4}V_x^{*\mathrm{T}}gR^{-1}g^{\mathrm{T}}V_x^* = 0,$$
$$V^*(0) = 0. \tag{7}$$

Lyshevski and Meyer (1995) showed that the value function obtained from (7) serves as a Lyapunov function on Ω.

To confront bounded controls, Lyshevski (1998) introduced a generalized nonquadratic functional

$$W(u) = 2\int_0^u (\boldsymbol{\phi}^{-1}(v))^{\mathrm{T}}R\,\mathrm{d}v,$$

$$\boldsymbol{\phi}(v) = [\phi(v_1)\cdots\phi(v_m)]^{\mathrm{T}},$$
$$\boldsymbol{\phi}^{-1}(u) = [\phi^{-1}(u_1)\cdots\phi^{-1}(u_m)], \tag{8}$$

where $v \in \mathbb{R}^m$, $\boldsymbol{\phi} \in \mathbb{R}^m$, and $\phi(\cdot)$ is a bounded one-to-one function that belongs to C^p $(p \geqslant 1)$ and $L_2(\Omega)$. Moreover, it is a monotonic odd function with its first derivative bounded by a constant M. An example is the hyperbolic tangent $\phi(\cdot) = \tanh(\cdot)$. R is positive definite and assumed to be symmetric for simplicity of analysis. Note that $W(u)$ is positive definite because $\phi^{-1}(u)$ is monotonic odd and R is positive definite.

The LE when (8) is used becomes

$$V_x^{\mathrm{T}}(f+g\cdot u)+Q+2\int_0^u \boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v=0,$$
$$V(0)=0. \tag{9}$$

Note that the LE becomes the HJB equation upon substituting the constrained optimal feedback control

$$u^*(x)=-\boldsymbol{\phi}(\tfrac{1}{2}\cdot R^{-1}g^{\mathrm{T}}V_x^*), \tag{10}$$

where $V^*(x)$ solves the following HJB equation:

$$V_x^{*\mathrm{T}}(f-g\boldsymbol{\phi}(\tfrac{1}{2}\cdot R^{-1}g^{\mathrm{T}}V_x^*))+Q$$
$$+2\int_0^{-\phi(\frac{1}{2}\cdot R^{-1}g^{\mathrm{T}}V_x^*)}\boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v=0,$$
$$V^*(0)=0. \tag{11}$$

Eq. (11) is a nonlinear differential equation for which there may be many solutions. Existence and uniqueness of the value function has been shown in Lyshevski (1996). This HJB equation cannot generally be solved. There is currently no method for rigorously solving for the value function of this constrained optimal control problem. Moreover, current solutions are not well defined over a specific region of the state space.

Remark 1. Optimal control problems do not necessarily have smooth or even continuous value functions, (Huang et al., 2000; Bardi & Capuzzo-Dolcetta, 1997). Lio (2000) used the theory of viscosity solutions to show that for infinite horizon optimal control problems with unbounded cost functionals, under certain continuity assumptions of the dynamics, the value function is continuous on some set Ω, $V^*(x)\in C(\Omega)$. Bardi and Capuzzo-Dolcetta (1997) showed that if the Hamiltonian is strictly convex and if the continuous viscosity solution is semiconcave, then $V^*(x)\in C^1(\Omega)$ satisfying the HJB equation everywhere. For affine input systems (1), the Hamiltonian is strictly convex if the system dynamics gain matrix $g(x)$ is constant, there is no bilinear terms of the state and the control, and if the integrand of (2) does not involve cross terms of the states and the control. In this paper, all derivations are performed under the assumption of smooth solutions to (9) and (11) with all what this requires of necessary conditions. See (Van Der Schaft, 1992; Saridis & Lee, 1979) for similar framework of solutions. If the smoothness assumption is released, then one needs to use the theory of viscosity solutions to show that the continuous cost solutions of (9) converge to the continuous value function of (11).

3. Successive approximation of HJB for saturated controls

The LE is linear in $V(x)$, while the HJB is nonlinear in $V^*(x)$. Solving the LE for $V(x)$ requires solving a linear differential equation, while the HJB solution involves a nonlinear differential equation that may be impossible to solve. This is the reason for introducing the successive approximation theory developed by Saridis and Lee (1979).

Successive approximation using the LE has not yet been rigorously applied to bounded controls. In this section, we show that the successive approximation theory can be extended to constrained input systems when nonquadratic performance functionals are used.

The successive approximation technique is now applied to (9), (10). The following lemma shows how Eq. (10) can be used to improve the control law.

Lemma 1. *If $u^{(i)}\in\Psi(\Omega)$, and $V^{(i)}\in C^1(\Omega)$ satisfies the equation $LE(V^{(i)},u^{(i)})=0$ with the boundary condition $V^{(i)}(0)=0$, then the new control derived as*

$$u^{(i+1)}(x)=-\boldsymbol{\phi}(\tfrac{1}{2}\cdot R^{-1}g^{\mathrm{T}}V_x^{(i)}) \tag{12}$$

is an admissible control for (1) *on Ω. Moreover, if the bounding function $\phi(\cdot)$ is monotone odd, and $V^{(i+1)}$ is the unique positive-definite function satisfying equation* $\mathrm{LE}(V^{(i+1)},u^{(i+1)})=0$, *with the boundary condition $V^{(i+1)}(0)=0$, then $V^*(x)\leqslant V^{(i+1)}(x)\leqslant V^{(i)}(x)$ $\forall x\in\Omega$.*

Proof. To show the admissibility part, since $V^{(i)}\in C^1(\Omega)$, the continuity assumption on g implies that $u^{(i+1)}$ is continuous. Since $V^{(i)}$ is positive definite it attains a minimum at the origin, and thus, $\mathrm{d}V^{(i)}/\mathrm{d}x$ must vanish there. This implies that $u^{(i+1)}(0)=0$. Taking the derivative of $V^{(i)}$ along the system $f+gu^{(i+1)}$ trajectory we have

$$\dot V^{(i)}(x,u^{(i+1)})=V_x^{(i)\mathrm{T}}f+V_x^{(i)\mathrm{T}}gu^{(i+1)}, \tag{13}$$

$$V_x^{(i)\mathrm{T}}f=-V_x^{(i)\mathrm{T}}gu^{(i)}-Q-2\int_0^{u^{(i)}}\boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v. \tag{14}$$

Therefore Eq. (13) becomes

$$\dot V^{(i)}(x,u^{(i+1)})=-V_x^{(i)\mathrm{T}}gu^{(i)}+V_x^{(i)\mathrm{T}}gu^{(i+1)}-Q$$
$$-2\int_0^{u^{(i)}}\boldsymbol{\phi}^{-T}(v)R\,\mathrm{d}v. \tag{15}$$

Since $V_x^{(i)\mathrm{T}}g(x)=-2\boldsymbol{\phi}^{-\mathrm{T}}(u^{(i+1)})R$, we get

$$\dot V^{(i)}(x,u^{(i+1)})=-Q+2\Bigg[\boldsymbol{\phi}^{-\mathrm{T}}(u^{(i+1)})R(u^{(i)}-u^{(i+1)})$$
$$-\int_0^{u^{(i)}}\boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v\Bigg]. \tag{16}$$

The second term in (16) is negative because ϕ and hence ϕ^{-1} is monotone odd. To see this, note that R is symmetric positivedefinite, this means we can rewrite it as $R = \Lambda\Sigma\Lambda$ where Σ is a diagonal matrix with its values being the singular values of R and Λ is an orthogonal symmetric matrix. Substituting for R in (16) we get

$$\dot{V}^{(i)}(x, u^{(i+1)}) = -Q + 2\left[\boldsymbol{\phi}^{-\mathrm{T}}(u^{(i+1)})\Lambda\Sigma\Lambda(u^{(i)} - u^{(i+1)}) - \int_0^{u^{(i)}} \boldsymbol{\phi}^{-\mathrm{T}}(v)\Lambda\Sigma\Lambda\,\mathrm{d}v\right]. \tag{17}$$

Applying the coordinate change $u = \Lambda^{-1}z$ to (17)

$$\begin{aligned}\dot{V}^{(i)}(x, u^{(i+1)}) &= -Q + 2\boldsymbol{\phi}^{-\mathrm{T}}(\Lambda^{-1}z^{(i+1)})\Lambda\Sigma\Lambda(\Lambda^{-1}z^{(i)} - \Lambda^{-1}z^{(i+1)}) - 2\int_0^{z^{(i)}} \boldsymbol{\phi}^{-\mathrm{T}}(\Lambda^{-1}\zeta)\Lambda\Sigma\Lambda\Lambda^{-1}\,\mathrm{d}\zeta \\ &= -Q + 2\boldsymbol{\phi}^{-\mathrm{T}}(\Lambda^{-1}z^{(i+1)})\Lambda\Sigma(z^{(i)} - z^{(i+1)}) - 2\int_0^{z^{(i)}} \boldsymbol{\phi}^{-\mathrm{T}}(\Lambda^{-1}\zeta)\Lambda\Sigma\,\mathrm{d}\zeta \\ &= -Q + 2\boldsymbol{\pi}^{\mathrm{T}}(z^{(i+1)})\Sigma(z^{(i)} - z^{(i+1)}) - 2\int_0^{z^{(i)}} \pi^{\mathrm{T}}(\zeta)\Sigma\,\mathrm{d}\zeta,\end{aligned} \tag{18}$$

where $\boldsymbol{\pi}^{\mathrm{T}}(z^{(i)}) = \boldsymbol{\phi}^{-1}(\Lambda^{-1}z^{(i)})^{\mathrm{T}}\Lambda$.

Since Σ is a diagonal matrix, we can now decouple the transformed input vector such that

$$\begin{aligned}\dot{V}^{(i)}(x, u^{(i+1)}) &= -Q + 2\boldsymbol{\pi}^{\mathrm{T}}(z^{(i+1)})\Sigma(z^{(i)} - z^{(i+1)}) - 2\int_0^{z_k^{(i)}} \boldsymbol{\pi}^{\mathrm{T}}(\zeta)\Sigma\,\mathrm{d}\zeta \\ &= -Q + 2\sum_{k=1}^{m}\Sigma_{kk}\left[\pi(z_k^{(i+1)})(z_k^{(i)} - z_k^{(i+1)}) - \int_0^{z_k^{(i)}} \pi(\zeta_k)\,\mathrm{d}\zeta_k\right].\end{aligned} \tag{19}$$

Since $R > 0$, then we have the singular values Σ_{kk} being all positive. Also, from the geometrical meaning of

$$\pi(z_k^{(i+1)})(z_k^{(i)} - z_k^{(i+1)}) - \int_0^{z_k^{(i)}} \pi(\zeta_k)\,\mathrm{d}\zeta_k,$$

this term is always negative if $\pi(z_k)$ is monotone and odd. Because $\phi^{-1}(\cdot)$ is monotone and odd one-to-one function, and since $\boldsymbol{\pi}^{\mathrm{T}}(z^{(i)}) = \boldsymbol{\phi}^{-1}(\Lambda^{-1}z^{(i)})^{\mathrm{T}}\Lambda$, it follows that $\pi(z_k)$ is monotone and odd. This implies that $\dot{V}^{(i)}(x, u^{(i+1)}) < 0$ and that $V^{(i)}(x)$ is a Lyapunov function for $u^{(i+1)}$ on Ω. From Definition 1, $u^{(i+1)}$ is admissible on Ω.

The second part of the lemma follows by considering that along the trajectories of $f + gu^{(i+1)}$, $\forall x_0$ we have

$$\begin{aligned}V^{(i+1)}(x_0) - V^{(i)}(x_0) &= \int_0^{\infty}\left\{Q(x(\tau, x_0, u^{(i+1)})) + 2\int_0^{u^{(i+1)}\left(x(\tau,x_0,u^{(i+1)})\right)} \boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v\right\}\mathrm{d}\tau \\ &\quad - \int_0^{\infty}\left\{Q(x(\tau, x_0, u^{(i+1)})) + 2\int_0^{u^{(i)}\left(x(\tau,x_0,u^{(i+1)})\right)} \boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v\right\}\mathrm{d}\tau \\ &= -\int_0^{\infty}\frac{\mathrm{d}(V^{(i+1)} - V^{(i)})^{\mathrm{T}}}{\mathrm{d}x}[f + g\,u^{(i+1)}]\,\mathrm{d}\tau.\end{aligned} \tag{20}$$

Because $\mathrm{LE}(V^{(i+1)}, u^{(i+1)}) = 0$, $\mathrm{LE}(V^{(i)}, u^{(i)}) = 0$

$$V_x^{(i)\mathrm{T}} f = -V_x^{(i)^{\mathrm{T}}} g u^{(i)} - Q - 2\int_0^{u^{(i)}} \boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v, \tag{21}$$

$$V_x^{(i+1)\mathrm{T}} f = -V_x^{(i+1)\mathrm{T}} g u^{(i+1)} - Q - 2\int_0^{u^{(i+1)}} \boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v. \tag{22}$$

Substituting (21), (22) in (20) we get

$$V^{(i+1)}(x_0) - V^{(i)}(x_0) = -2\int_0^{\infty}\left\{\boldsymbol{\phi}^{-\mathrm{T}}(u^{(i+1)})R(u^{(i+1)} - u^{(i)}) - \int_{u^{(i)}}^{u^{(i+1)}} \boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v\right\}\mathrm{d}\tau. \tag{23}$$

By decoupling Eq. (23) using $R = \Lambda\Sigma\Lambda$, it can be shown that $V^{(i+1)}(x_0) - V^{(i)}(x_0) \leqslant 0$ when $\phi(\cdot)$ is monotone odd function. Moreover, it can be shown by contradiction that $V^*(x_0) \leqslant V^{(i+1)}(x_0)$. □

The next theorem is a key result which justifies applying the successive approximation theory to constrained input systems.

Definition 2 (*Uniform Convergence*). A sequence of functions $\{f_n\}$ converges uniformly to f on a set Ω if $\forall\varepsilon > 0,\ \exists N(\varepsilon) :\ n > N \Rightarrow |f_n(x) - f(x)| < \varepsilon\ \forall x \in \Omega$, or equivalently $\sup_{x\in\Omega}|f_n(x) - f(x)| < \varepsilon$.

Theorem 1. *If $u^{(0)} \in \Psi(\Omega)$, then $u^{(i)} \in \Psi(\Omega)$, $\forall i \geqslant 0$. Moreover, $V^{(i)} \to V^*$, $u^{(i)} \to u^*$ uniformly on Ω.*

Proof. From Lemma 1, it can be shown by induction that $u^{(i)} \in \Psi(\Omega)$, $\forall i \geqslant 0$. Furthermore, Lemma 1 shows that $V^{(i)}$ is a monotonically decreasing sequence and bounded below by $V^*(x)$. Hence, $V^{(i)}$ converges pointwise to $V^{(\infty)}$.

Because Ω is compact, then uniform convergence follows immediately from Dini's theorem (Apostol, 1974). Due to the uniqueness of the value function (Lewis & Syrmos, 1995; Lyshevski, 1996), it follows that $V^{(\infty)} = V^*$. Controllers $u^{(i)}$ are admissible, therefore they are continuous and have unique trajectories due to the locally Lipschitz continuity assumption on the dynamics. Since (2) converges uniformly to V^*, this implies that the system trajectories converge $\forall x_0 \in \Omega$. Therefore $u^{(i)} \to u^{(\infty)}$ uniformly on Ω. If $\mathrm{d}V^{(i)}/\mathrm{d}x \to \mathrm{d}V^*/\mathrm{d}x$ uniformly, we conclude that $u^{(\infty)}=u^*$. To prove that $\mathrm{d}V^{(i)}/\mathrm{d}x \to \mathrm{d}V^*/\mathrm{d}x$ uniformly on Ω, note that $\mathrm{d}V^{(i)}/\mathrm{d}x$ converges uniformly to some continuous function J. Since $V^{(i)} \to V^*$ uniformly and $\mathrm{d}V^{(i)}/\mathrm{d}x$ exists $\forall i$, it follows that the sequence $\mathrm{d}V^{(i)}/\mathrm{d}x$ is term-by-term differentiable (Apostol, 1974), and $J = \mathrm{d}V^*/\mathrm{d}x$. □

Beard (1995) has shown that improving the control law does not reduce the RAS for the case of unconstrained controls. Similarly, we show that this is the case for systems with constrained controls.

Corollary 1. *If Ω^* denotes the RAS of the constrained optimal control u^*, then Ω^* is the largest RAS of any other admissible control law.*

Proof. The proof is by contradiction. Lemma 1 shows that the saturated control u^* is asymptotically stable on $\Omega^{(0)}$, where $\Omega^{(0)}$ is the stability region of the saturated control $u^{(0)}$. Assume that u_{Largest} is an admissible controller with the largest RAS Ω_{Largest}. Then, there is $x_0 \in \Omega_{\text{Largest}},\ x_0 \notin \Omega^*$. From Theorem 1, $x_0 \in \Omega^*$ which completes the proof. □

Note that there may be stabilizing saturated controls that have larger stability regions than u^*, but are not admissible with respect to $Q(x)$ and the system (f, g).

4. Neural network least-squares approximate HJB solution

Although Eq. (9) is a linear differential equation, when substituting (10) into (9), it is still difficult to solve for the cost function $V^{(i)}(x)$. In this section, neural networks are used along with the theory of successive approximation, to solve for the value function of (11) over Ω, by approximating the solution for the cost function $V^{(i)}(x)$ at each successive iteration i in a least-squares sense. Moreover, to approximate integration, a mesh is introduced on Ω. This yields an efficient, practical, and computationally tractable solution algorithm to find nearly optimal state feedback controllers of constrained input nonlinear systems.

4.1. Neural network approximation of $V(x)$

It is well known that neural networks can be used to approximate smooth functions on prescribed compact sets (Lewis et al., 1999). Since the analysis is restricted to the RAS of some initial stabilizing controller, neural networks are natural for this application. Therefore, to successively solve (9) and (10), $V^{(i)}$ is approximated by

$$V_L^{(i)}(x) = \sum_{j=1}^{L} w_j^{(i)} \sigma_j(x) = \mathbf{w}_L^{T(i)} \boldsymbol{\sigma}_L(x) \tag{24}$$

which is a neural network with the activation functions $\sigma_j(x) \in C^1(\Omega)$, $\sigma_j(0) = 0$. The neural network weights are $w_j^{(i)}$. L is the number of hidden-layer neurons. $\boldsymbol{\sigma}_L(x) \equiv [\sigma_1(x)\ \sigma_2(x)\ \cdots\ \sigma_L(x)]^{\mathrm{T}}$ is the vector activation function, $\mathbf{w}_L^{(i)} \equiv [w_1^{(i)}\ w_2^{(i)}\ \cdots\ w_L^{(i)}]^{\mathrm{T}}$ is the vector weight. The weights are tuned to minimize the residual error in a least-squares sense over a set of points sampled from a compact set Ω inside the RAS of the initial stabilizing control.

For $\mathrm{LE}(V, u) = 0$, the solution V is replaced with V_L having a residual error

$$LE\left(V_L(x) = \sum_{j=1}^{L} w_j \sigma_j(x),\ u\right) = e_L(x). \tag{25}$$

To find the least-squares solution, the method of weighted residuals is used (Finlayson, 1972). The weights, $\mathbf{w}_L$, are determined by projecting the residual error onto $\mathrm{d}e_L(x)/\mathrm{d}\mathbf{w}_L$ and setting the result to zero $\forall x \in \Omega$ using the inner product, i.e.

$$\left\langle \frac{\mathrm{d}e_L(x)}{\mathrm{d}\mathbf{w}_L}, e_L(x) \right\rangle = 0, \tag{26}$$

where $\langle f, g\rangle = \int_\Omega fg\,\mathrm{d}x$ is a Lebesgue integral. One has

$$\langle \nabla\boldsymbol{\sigma}_L(f+gu), \nabla\boldsymbol{\sigma}_L(f+gu)\rangle \mathbf{w}_L + \left\langle Q + 2\int_0^u \left(\boldsymbol{\phi}^{-1}(v)\right)^{\mathrm{T}} R\,\mathrm{d}v, \nabla\boldsymbol{\sigma}_L(f+gu)\right\rangle = 0. \tag{27}$$

The following technical results are needed.

Lemma 2. *If the set $\{\sigma_j\}_1^L$ is linearly independent and $u \in \Psi(\Omega)$, then the set*

$$\{\nabla\sigma_j^{\mathrm{T}}(f+gu)\}_1^L \tag{28}$$

is also linearly independent.

Proof. See (Beard et al., 1997). □

Because of Lemma 2, $\langle \nabla\boldsymbol{\sigma}_L(f+gu), \nabla\boldsymbol{\sigma}_L(f+gu)\rangle$ is of full rank, and thus is invertible. Therefore, a unique solution for $\mathbf{w}_L$ exists and is computed as

$$\mathbf{w}_L = -\langle \nabla\boldsymbol{\sigma}_L(f+gu), \nabla\boldsymbol{\sigma}_L(f+gu)\rangle^{-1} \bullet \left\langle Q + 2\int_0^u (\boldsymbol{\phi}^{-1}(v))^{\mathrm{T}} R\,\mathrm{d}v,\ \nabla\boldsymbol{\sigma}_L(f+gu)\right\rangle. \tag{29}$$

Having solved for $\mathbf{w}_L$, the improved control is given by

$$u = -\boldsymbol{\phi}\left(\tfrac{1}{2} R^{-1} g^{\mathrm{T}}(x) \nabla\boldsymbol{\sigma}_L^{\mathrm{T}} \mathbf{w}_L\right). \tag{30}$$

Eqs. (29) and (30) are successively solved at each iteration (i) until convergence.

4.2. Convergence of the method of least squares

In what follows, we show convergence results of the method of least squares when neural networks are used to solve for the cost function of the LE. Note that (24) is a Fourier series expansion. The following definitions are required.

Definition 3 (*Convergence in the mean*). A sequence of functions $\{f_n\}$ that is Lebesgue integrable on a set Ω, $L_2(\Omega)$, is said to converge in the mean to f on Ω if $\forall \varepsilon > 0,\ \exists N(\varepsilon) : n > N \Rightarrow \|f_n(x) - f(x)\|_{L_2(\Omega)} < \varepsilon$, where $\|f\|^2_{L_2(\Omega)} = \langle f, f\rangle$.

The convergence proofs of the least-squares method are done in the Sobolev function space setting (Adams & Fournier, 2003). This allows defining functions that are $L_2(\Omega)$ with their partial derivatives.

Definition 4. Sobolev space $H^{m,p}(\Omega)$: Let Ω be an open set in $\mathbb{R}^n$ and let $u \in C^m(\Omega)$. Define a norm on u by

$$\|u\|_{m,p} = \sum_{0\leqslant|\alpha|\leqslant m}\left(\int_\Omega |D^\alpha u(x)|^p\,\mathrm{d}x\right)^{1/p}, \quad 1 \leqslant p < \infty.$$

This is the Sobolev norm in which the integration is Lebesgue. The completion of $u \in C^m(\Omega) : \|u\|_{m,p} < \infty$ with respect to $\|\|_{m,p}$ is the Sobolev space $H^{m,p}(\Omega)$. For $p = 2$, the Sobolev space is a Hilbert space.

The LE can be written using the linear operator A defined on the Hilbert space $H^{1,2}(\Omega)$

$$\overbrace{V_x^{\mathrm{T}}(f + gu)}^{AV} = \overbrace{-Q - W(u)}^{P}.$$

In Mikhlin (1964), it is shown that if the set $\{\sigma_j\}_1^L$ is complete, and the operator A and its inverse are bounded, then $\|AV_L - AV\|_{L_2(\Omega)} \to 0$ and $\|V_L - V\|_{L_2(\Omega)} \to 0$. For the LE however, it can be shown that these sufficiency conditions are violated.

Neural networks based on power series have an important property that they are differentiable. This means that they can approximate uniformly a continuous function with all its partial derivatives up to order m using the same polynomial, by differentiating the series termwise. This type of series is m-uniformly dense as shown in Lemma 3. Other m-uniformly dense neural networks, not necessarily based on power series, are studied in Hornik et al. (1990).

Lemma 3. *High-order Weierstrass approximation theorem. Let $f(x) \in C^m(\Omega)$, then there exists a polynomial, $P(x)$, such that it converges uniformly to $f(x) \in C^m(\Omega)$, and such that all its partial derivatives up to order m converges uniformly* (Finlayson, 1972), Hornik et al. (1990).

Lemma 4. *Given N linearly independent set of functions $\{f_n\}$. Then for the Fourier series $\alpha_N^{\mathrm{T}} f_N$, it follows that $\|\alpha_N^{\mathrm{T}} f_N\|^2_{L_2(\Omega)} \to 0 \Leftrightarrow \|\alpha_N\|^2_{l_2} \to 0$.*

Proof. To show the sufficiency part, note that the Gram matrix, $G = \langle f_N, f_N\rangle$, is positive definite. Hence, $\alpha_N^{\mathrm{T}} G_N \alpha_N \geqslant \lambda(G_N)\|\alpha_N\|^2_{l_2}$, $\lambda(G_N) > 0\ \forall N$. If $\alpha_N^{\mathrm{T}} G_N \alpha_N \to 0$,then $\|\alpha_N\|^2_{l_2} = \alpha_N^{\mathrm{T}} G_N \alpha_N / \lambda(G_N) \to 0$ because $\lambda(G_N) > 0\ \forall N$.

To show the necessity part, note that

$$\begin{aligned}&\|\alpha_N\|^2_{L_2(\Omega)} - 2\|\alpha_N^{\mathrm{T}} f_N\|^2_{L_2(\Omega)} + \|f_N\|^2_{L_2(\Omega)}\\ &\quad = \|\alpha_N - f_N\|^2_{L_2(\Omega)},\end{aligned}$$

$$\begin{aligned}2\|\alpha_N^{\mathrm{T}} f_N\|^2_{L_2(\Omega)} &= \|\alpha_N\|^2_{L_2(\Omega)} + \|f_N\|^2_{L_2(\Omega)}\\ &\quad - \|\alpha_N - f_N\|^2_{L_2(\Omega)}.\end{aligned}$$

Using the Parallelogram Law

$$\begin{aligned}&\|\alpha_N - f_N\|^2_{L_2(\Omega)} + \|\alpha_N + f_N\|^2_{L_2(\Omega)}\\ &\quad = 2\|\alpha_N\|^2_{L_2(\Omega)} + 2\|f_N\|^2_{L_2(\Omega)},\end{aligned}$$

as $N \to \infty$

$$\begin{aligned}&\|\alpha_N - f_N\|^2_{L_2(\Omega)} + \|\alpha_N + f_N\|^2_{L_2(\Omega)}\\ &\quad = \overbrace{2\|\alpha_N\|^2_{L_2(\Omega)}}^{\to 0} + 2\|f_N\|^2_{L_2(\Omega)},\\ &\quad \Rightarrow \|\alpha_N - f_N\|^2_{L_2(\Omega)} \to \|f_N\|^2_{L_2(\Omega)},\ \ \|\alpha_N + f_N\|^2_{L_2(\Omega)}\\ &\quad \to \|f_N\|^2_{L_2(\Omega)},\end{aligned}$$

$$\begin{aligned}2\|\alpha_N f_N\|^2_{L_2(\Omega)} &= \overbrace{\|\alpha_N\|^2_{L_2(\Omega)}}^{\to 0} + \|f_N\|^2_{L_2(\Omega)}\\ &\quad - \overbrace{\|\alpha_N - f_N\|^2_{L_2(\Omega)}}^{\to \|f_N\|^2_{L_2(\Omega)}} \to 0.\end{aligned}$$

Therefore, $\|\alpha_N\|^2_{l_2} \to 0 \Rightarrow \|\alpha_N f_N\|^2_{L_2(\Omega)} \to 0$. □

The following assumptions are required.

Assumption 1. The LE solution is positive definite. This is guaranteed for stabilizable dynamics and when the performance functional satisfies zero-state observability defined in (Van Der Schaft, 1992).

Assumption 2. The system dynamics and the performance integrands $Q(x) + W(u(x))$ are such that the solution of the LE is continuous and differentiable. Therefore, belongs to the Sobolev space $V \in H^{1,2}(\Omega)$.

Assumption 3. We can choose a complete coordinate elements $\{\sigma_j\}_1^\infty \in H^{1,2}(\Omega)$ such that the solution $V \in H^{1,2}(\Omega)$ and $\{\partial V/\partial x_1, \ldots, \partial V/\partial x_n\}$ can be uniformly approximated by the infinite series built from $\{\sigma_j\}_1^\infty$.

Assumption 4. The sequence $\{\psi_j = A\sigma_j$ is linearly independent and complete.

The first three assumptions are standard in optimal control and neural network control literature. Lemma 2 assures the linear independence required in the fourth assumption. Completeness follows from Lemma 3 and (Hornik et al., 1990),

$$\forall V, \varepsilon \exists L, \mathbf{w}_L \because |V_L - V| < \varepsilon, \forall i |\mathrm{d}V_L/\mathrm{d}x_i - \mathrm{d}V/\mathrm{d}x_i| < \varepsilon.$$

This implies that as $L \to \infty$

$$\sup_{x\in\Omega} |AV_L - AV| \to 0 \Rightarrow \|AV_L - AV\|_{L_2(\Omega)} \to 0,$$

and therefore completeness of the set $\{\psi_j\}$ is established.

The next theorem uses these assumptions to conclude convergence results of the least-squares method which is placed in the Sobolev space $H^{1,2}(\Omega)$.

Theorem 2. *If Assumptions* 1–4 *hold, then approximate solutions exist for the LE using the method of least squares and are unique for each L. In addition*

(R1) $\|LE(V_L(x)) - LE(V(x))\|_{L_2(\Omega)} \to 0,$
(R2) $\|V_{x_L} - V_x\|_{L_2(\Omega)} \to 0,$
(R3) $\|u_L(x) - u(x)\|_{L_2(\Omega)} \to 0.$

Proof. *Existence* of a least-squares solution for the LE can be easily shown. The least-squares solution V_L is nothing but the solution of the minimization problem

$$\|AV_L - P\|^2 = \min_{\tilde{V}\in S_L} \|A\tilde{V} - P\|^2 = \min_{w} \|\mathbf{w}_L^{\mathrm{T}}\boldsymbol{\psi}_L - P\|^2,$$

where S_L is the span of $\{\sigma_1, \ldots, \sigma_L\}$.

Uniqueness follows from the linear independence of $\{\psi_1, \ldots, \psi_L\}$. (R1) follows from the *completeness* of $\{\psi_j\}$.

To show the second result (R2), write the LE in terms of its series expansion on Ω with coefficients c_j

$$LE\left(V_L = \sum_{i=1}^{N} w_i\sigma_i\right) - \overbrace{LE\left(V = \sum_{i=1}^{\infty} c_i\sigma_i\right)}^{=0} = \varepsilon_L(x),$$

$$(\mathbf{w}_L - \mathbf{c}_L)^{\mathrm{T}}\nabla\boldsymbol{\sigma}_L(f + gu) = \varepsilon_L(x) + \overbrace{\sum_{i=L+1}^{\infty} c_i \frac{\mathrm{d}\sigma_i}{\mathrm{d}x}(f + gu)}^{e_L(x)}.$$

Note that $e_L(x)$ converges uniformly to zero due to Lemma 3, and hence converges in the mean. On the other hand $\varepsilon_L(x)$ converges in the mean due to (R1). Therefore,

$$\|(\mathbf{w}_L - \mathbf{c}_L)^{\mathrm{T}}\nabla\boldsymbol{\sigma}_L(f + gu)\|^2_{L_2(\Omega)} = \|\varepsilon_L(x) + e_L(x)\|^2_{L_2(\Omega)} \leqslant 2\|\varepsilon_L(x)\|^2_{L_2(\Omega)} + 2\|e_L(x)\|^2_{L_2(\Omega)} \to 0.$$

Because $\nabla\boldsymbol{\sigma}_L(f + gu)$ is linearly independent, using Lemma 4, we conclude that $\|\mathbf{w}_L - \mathbf{c}_L\|^2_{l2} \to 0$. Furthermore, the set $\{\mathrm{d}\sigma_i/\mathrm{d}x\}$ is linearly independent, and hence from Lemma 4 it follows that $\|(\mathbf{w}_L - \mathbf{c}_L)^{\mathrm{T}}\nabla\boldsymbol{\sigma}_L\|^2_{L_2(\Omega)} \to 0$. Because the infinite series with c_j converges uniformly. It follows that as $L \to \infty$, $\|\mathrm{d}V_L/\mathrm{d}x - \mathrm{d}V/\mathrm{d}x\|_{L_2(\Omega)} \to 0$.

Finally, (R3) follows from (R2) and from the fact that $g(x)$ is continuous and therefore bounded on Ω. Hence

$$\| -\tfrac{1}{2} \cdot R^{-1}g^{\mathrm{T}}(V_{x_L} - V_x)\|^2_{L_2(\Omega)} \leqslant \| -\tfrac{1}{2} \cdot R^{-1}g^{\mathrm{T}}\|^2_{L_2(\Omega)}\|(V_{x_L} - V_x)\|^2_{L_2(\Omega)} \to 0.$$

Denote $\alpha_{j,L}(x) = -\frac{1}{2} \cdot g_j^{\mathrm{T}}V_{x_L}$, $\alpha_j(x) = -\frac{1}{2} \cdot g_j^{\mathrm{T}}V_x$

$$\begin{aligned} u_L - u &= \boldsymbol{\phi}(-\tfrac{1}{2} \cdot g^{\mathrm{T}}V_{x_L}) - \boldsymbol{\phi}(-\tfrac{1}{2} \cdot g^{\mathrm{T}}V_x), \\ &= [\phi(\alpha_{1,L}(x)) - \phi(\alpha_1(x)) \cdots \phi(\alpha_{m,L}(x)) \\ &\quad - \phi(\alpha_m(x))]. \end{aligned}$$

Because $\phi(\cdot)$ is smooth, and under the assumption that its first derivative is bounded by a constant M, then we have $\phi(\alpha_{j,L}) - \phi(\alpha_j) \leqslant M(\alpha_{j,L}(x) - \alpha_j(x))$. Therefore

$$\|\alpha_{j,L}(x) - \alpha_j(x)\|_{L_2(\Omega)} \to 0 \Rightarrow \|\phi(\alpha_{j,L}) - \phi(\alpha_j)\|_{L_2(\Omega)} \to 0,$$

and (R3) follows. □

Corollary 2. *If the results of Theorem* 2 *hold, then*

$$\sup_{x\in\Omega} |V_{x_L} - V_x| \to 0, \quad \sup_{x\in\Omega} |V_L - V| \to 0, \quad \sup_{x\in\Omega} |u_L - u| \to 0.$$

Proof. This follows by noticing that $\|\mathbf{w}_L - \mathbf{c}_L\|^2_{l2} \to 0$ and, the series with c_j is uniformly convergent, and (Hornik et al., 1990). □

Next, the admissibility of $u_L(x)$ is shown.

Corollary 3. *Admissibility of* $u_L(x)$:

$$\exists L_0 : L \geqslant L_0,\ u_L \in \Psi(\Omega).$$

Proof. Consider the following LE

$$\dot{V}^{(i)}(x, u_L^{(i+1)}) = -Q \overbrace{-2\int_{u^{(i+1)}}^{u^{(i)}} \boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v + 2\boldsymbol{\phi}^{-\mathrm{T}}(u^{(i+1)})R(u^{(i)} - u^{(i+1)})}^{\leqslant 0} -2\boldsymbol{\phi}^{-\mathrm{T}}(u^{(i+1)})R(u_L^{(i+1)} - u^{(i+1)}) - 2\int_0^{u^{(i+1)}} \boldsymbol{\phi}^{-\mathrm{T}}(v)R\,\mathrm{d}v.$$

Since $u_L^{(i+1)}$ is guaranteed to be within a tube around $u^{(i+1)}$ $u_L^{(i+1)} \to u^{(i+1)}$ uniformly. Therefore one can easily see that (31) with $\alpha > 0$ is satisfied $\forall x \in \Omega \not\cap \Omega_1(\varepsilon(L))$

where $\Omega_1(\varepsilon(L)) \subseteq \Omega$ containing the origin.

$$\phi^{-\mathrm{T}}(u^{(i+1)})Ru_L^{(i+1)} \geqslant \tfrac{1}{2}\cdot\phi^{-\mathrm{T}}(u^{(i+1)})Ru^{(i+1)} + \alpha\int_0^{u^{(i+1)}}\phi^{-\mathrm{T}}R\,\mathrm{d}v. \tag{31}$$

Hence, $\dot V^{(i)}(x,u_L^{(i+1)})<0\ \forall x\in\Omega\,\not\cap\,\Omega_1(\varepsilon(L))$. Given that $u_L^{(i+1)}(0)=0$, from the continuity of $u_L^{(i+1)}$, there exists $\Omega_2(\varepsilon(L))\subseteq\Omega_1(\varepsilon(L))$ containing the origin for which $\dot V^{(i)}(x,u_L^{(i+1)})<0$. As L increases, $\Omega_1(\varepsilon(L))$ gets smaller while $\Omega_2(\varepsilon(L))$ gets larger and (31) is satisfied $\forall x\in\Omega$. Therefore, $\exists L_0 : L\geqslant L_0, \dot V^{(i)}(x,u_L^{(i+1)})<0\ \forall x\in\Omega$ and hence $u_L\in\Psi(\Omega)$. □

Corollary 4. *It can be shown that* $\sup_{x\in\Omega}|u_L(x)-u(x)|\to 0$ *implies that* $\sup_{x\in\Omega}|J(x)-V(x)|\to 0$, *where* $LE(J,u_L)=0$, $LE(V,u)=0$.

4.3. Convergence of the method of least-squares to the solution of the HJB equation

In this section, we would like to have a theorem analogous to Theorem 1 that guarantees that the successive least-squares solution converge to the value function of the HJB equation (11).

Theorem 3. *Under the assumptions of Theorem* 2, *the following is satisfied* $\forall i\geqslant 0$:

(i) $\sup_{x\in\Omega}|V_L^{(i)}-V^{(i)}|\to 0$,
(ii) $\sup_{x\in\Omega}|u_L^{(i+1)}-u^{(i+1)}|\to 0$,
(iii) $\exists L_0 : L\geqslant L_0,\ u_L^{(i+1)}\in\Psi(\Omega)$.

Proof. The proof is by induction.

Basis step:

Using Corollaries 2 and 3, it follows that for any $u^{(0)}\in\Psi(\Omega)$, one has

(I) $\sup_{x\in\Omega}|V_L^{(0)}-V^{(0)}|\to 0$,
(II) $\sup_{x\in\Omega}|V_L^{(1)}-V^{(1)}|\to 0$,
(III) $\exists L_0 : L\geqslant L_0,\ u_L^{(1)}\in\Psi(\Omega)$.

Inductive step:

Assume that

(a) $\sup_{x\in\Omega}|V_L^{(i-1)}-V^{(i-1)}|\to 0$,
(b) $\sup_{x\in\Omega}|u_L^{(i)}-u^{(i)}|\to 0$,
(c) $\exists L_0 : L\geqslant L_0,\ u_L^{(i)}\in\Psi(\Omega)$.

If $J^{(i)}$ is such that $\mathrm{LE}(J^{(i)},\ u_L^{(i)})=0$. Then from Corollary 2, $J^{(i)}$ can be uniformly approximated by $V_L^{(i)}$. Moreover from assumption (b) and Corollary 4, it follows that as $u_L^{(i)}\to u^{(i)}$ uniformly then $J^{(i)}\to V^{(i)}$ uniformly. Therefore $V_L^{(i)}\to V^{(i)}$ uniformly.

Because $V_L^{(i)}\to V^{(i)}$ uniformly, $u_L^{(i+1)}\to u^{(i+1)}$ uniformly by Corollary 2. From Corollary 3, $\exists L_0 : L\geqslant L_0\Rightarrow u_L^{(i+1)}\in\Psi(\Omega)$.

Hence the proof by induction is complete. □

The next theorem is an important result upon which the algorithm proposed in Section 4.4 of this paper is justified.

Theorem 4. $\forall\varepsilon>0, \exists i_0, L_0 : i\geqslant i_0, L\geqslant L_0$ *one has*

(A) $\sup_{x\in\Omega}|V_L^{(i)}-V^*|<\varepsilon$,
(B) $\sup_{x\in\Omega}|u_L^{(i)}-u^*|<\varepsilon$,
(C) $u_L^{(i)}\in\Psi(\Omega)$.

Proof. This follows directly from Theorems 1 and 3. □

4.4. Algorithm for nearly optimal neurocontrol design with saturated controls: introducing a mesh in $\mathbb{R}^n$

Solving the integration in (29) is expensive computationally. The integrals can be well approximated by discretization. A mesh of points over the integration region can be introduced on Ω of size Δx. The terms of (29) can be rewritten as follows:

$$X=\lfloor\nabla\boldsymbol{\sigma}_L(f+gu)|_{x_1}\cdots\cdots\nabla\boldsymbol{\sigma}_L(f+gu)|_{x_p}\rfloor^{\mathrm{T}}, \tag{32}$$

$$Y=\left\lfloor Q+2\int_0^u\phi^{-\mathrm{T}}(v)R\,\mathrm{d}v\bigg|_{x_1}\cdots\cdots Q+2\int_0^u\phi^{-\mathrm{T}}(v)R\,\mathrm{d}v\bigg|_{x_p}\right\rfloor^{\mathrm{T}}, \tag{33}$$

where p in x_p represents the number of points of the mesh. Reducing the mesh size, we have

$$\langle\nabla\boldsymbol{\sigma}_L(f+gu),\nabla\boldsymbol{\sigma}_L(f+gu)\rangle=\lim_{\|\Delta x\|\to 0}(X^{\mathrm{T}}X)\cdot\Delta x,$$

$$\left\langle Q+2\int_0^u\phi^{-\mathrm{T}}(v)R\,\mathrm{d}v,\nabla\boldsymbol{\sigma}_L(f+gu)\right\rangle=\lim_{\|\Delta x\|\to 0}(X^{\mathrm{T}}Y)\cdot\Delta x. \tag{34}$$

This implies that we can calculate $\mathbf{w}_L$ as

$$\mathbf{w}_{L,p}=-(X^{\mathrm{T}}X)^{-1}(X^{\mathrm{T}}Y). \tag{35}$$

Various ways to efficiently approximate integrals exists. Monte Carlo integration techniques can be used. Here the mesh points are sampled stochastically instead of being selected in a deterministic fashion (Evans & Swartz, 2000). In any case however, the numerical algorithm at the end requires solving (35) which is a least-squares computation of the neural network weights.

Numerically stable routines that compute equations like (35) do exists in several software packages like MATLAB which is used the next section.

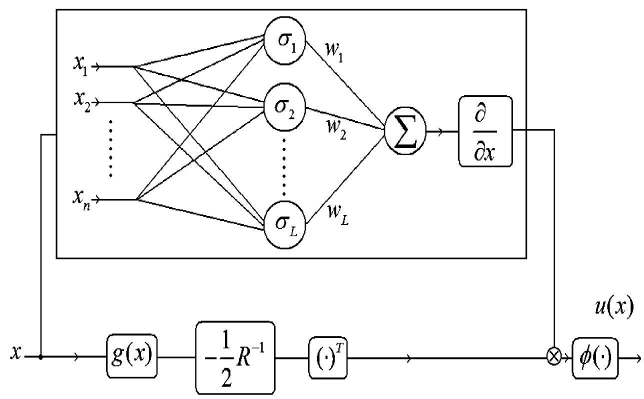

Fig. 1. Neural-network-based nearly optimal saturated control law.

The neurocontrol law structure is shown in Fig. 1. It is a neural network with activation functions given by $\boldsymbol{\sigma}$, multiplied by a function of the state variables.

A flowchart of the computational algorithm presented in this paper is shown in Fig. 2. This is an offline algorithm run a priori to obtain a neural network constrained state feedback controller that is nearly optimal.

5. Numerical examples

We now show the power of our neural network control technique of finding nearly optimal controllers for affine in input dynamical systems. Two examples are presented.

5.1. Multi input canonical form linear system with constrained inputs

We start by applying the algorithm obtained above for the linear system

$$\dot{x}_1 = 2x_1 + x_2 + x_3, \quad \dot{x}_2 = x_1 - x_2 + u_2, \quad \dot{x}_3 = x_3 + u_1.$$

It is desired to control the system with input constraints $|u_1| \leqslant 3$, $|u_2| \leqslant 20$. This system is null controllable. Therefore global asymptotic stabilization cannot be achieved (Sussmann et al., 1994).

To find a nearly optimal constrained state feedback controller, the following smooth function is used to approximate the value function of the system:

$$\begin{aligned} V_{21}(x_1, x_2, x_3) &= w_1 x_1^2 + w_2 x_2^2 + w_3 x_3^2 + w_4 x_1 x_2 \\ &+ w_5 x_1 x_3 + w_6 x_2 x_3 + w_7 x_1^4 + w_8 x_2^4 + w_9 x_3^4 \\ &+ w_{10} x_1^2 x_2^2 + w_{11} x_1^2 x_3^2 + w_{12} x_2^2 x_3^2 + w_{13} x_1^2 x_2 x_3 \\ &+ w_{14} x_1 x_2^2 x_3 + w_{15} x_1 x_2 x_3^2 + w_{16} x_1^3 x_2 + w_{17} x_1^3 x_3 \\ &+ w_{18} x_1 x_2^3 + w_{19} x_1 x_3^3 + w_{20} x_2 x_3^3 + w_{21} x_2^3 x_3. \end{aligned}$$

The selection of the neural network is usually a natural choice guided by engineering experience and intuition. This is a neural network with polynomial activation functions,

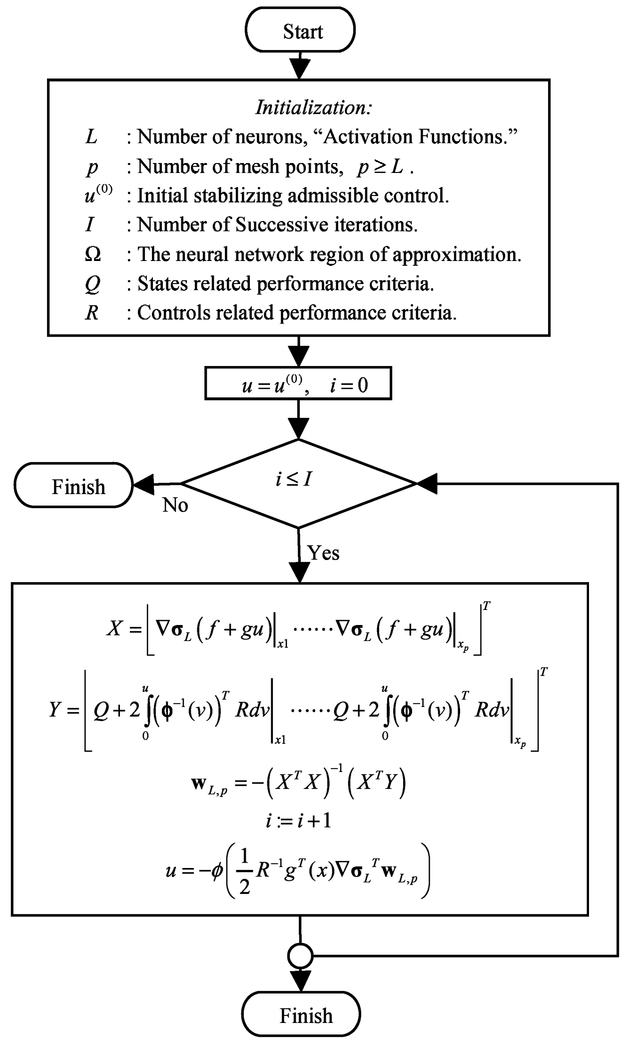

Fig. 2. Successive approximation algorithm for nearly optimal saturated neurocontrol.

and hence $V_{21}(0) = 0$. This is a power series neural network with 21 activation functions containing powers of the state variable of the system up to the fourth order. Convergence was not observed that for neurons with second-order power of the states. The number of neurons required is chosen to guarantee the uniform convergence of the algorithm. The activation functions selected in this example satisfy the properties of activation functions discussed in Section 4 of this paper, and in Lewis et al. (1999).

To initialize the algorithm, an LQR controller is derived assuming no input constraints. Then the control signal is passed through a saturation block. Note that the closed loop dynamics is not optimal anymore. The following controller is then found and its performance is shown in Fig. 3.

$$\begin{aligned} u_1 &= -8.31x_1 - 2.28x_2 - 4.66x_3, \quad |u_1| \leqslant 3, \\ u_2 &= -8.57x_1 - 2.27x_2 - 2.28x_3, \quad |u_2| \leqslant 20. \end{aligned}$$

In order to model the saturation of actuators, a nonquadratic cost performance term (8) is used. To show how to do this for the general case of $|u| \leqslant A$, we use $A^* \tanh(1/A \cdot \cdots)$ for $\phi(\cdot \cdot \cdot)$. Hence the nonquadratic cost

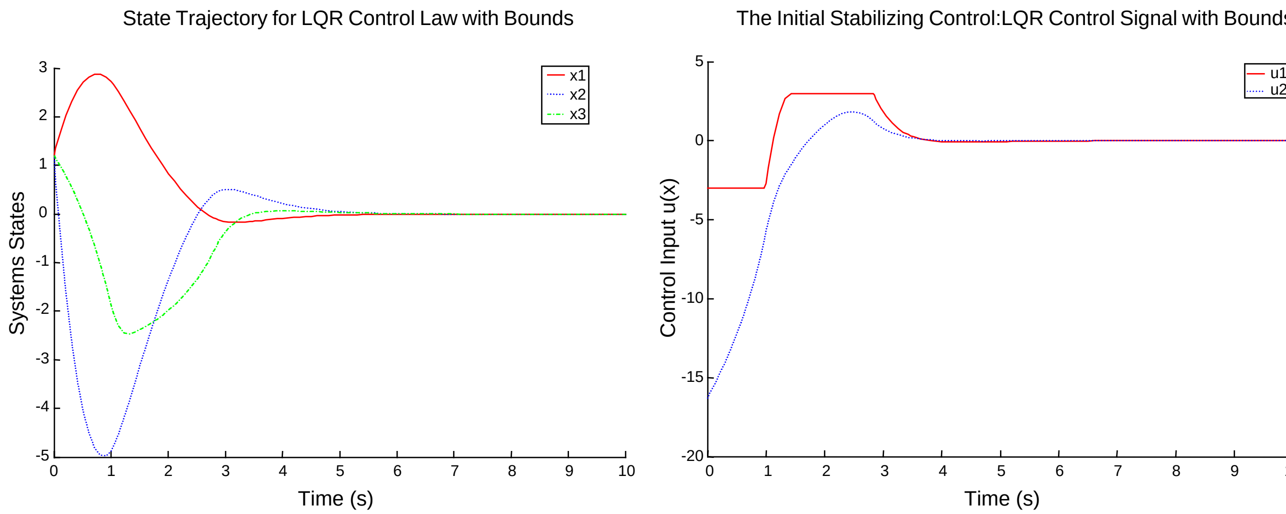

Fig. 3. LQR control with actuator saturation.

associated with controls in the optimization functional is given as

$$W(u) = 2\int_0^u (A\tanh^{-1}(v/A))^{\mathrm{T}} R\,\mathrm{d}v = 2ARu\tanh^{-1}(u/A) + A^2R\ln(1-u^2/A^2).$$

This nonquadratic cost performance is then used in the algorithm to calculate the optimal constrained controller. The algorithm of Fig. 2 is executed for the region $-1.2 \leqslant x_1 \leqslant 1.2,\ -1.2 \leqslant x_2 \leqslant 1.2,\ -1.2 \leqslant x_3 \leqslant 1.2$ with the design parameters $R = I_{2x2},\ Q = I_{3x3}$. This region falls within the RAS of the initial stabilizing controller. Methods to estimate the RAS are discussed in Khalil (2003).

After 20 successive iterations, the algorithm converges to the following nearly optimal saturated control,

$$u_1 = -3\tanh\left(\frac{1}{3}\left\{\begin{array}{l} 7.7x_1 + 2.44x_2 + 4.8x_3 \\ +2.45x_1^3 + 2.27x_1^2x_2 + 3.7x_1x_2x_3 \\ +0.71x_1x_2^2 + 5.8x_1^2x_3 + 4.8x_1x_3^2 \\ +0.08x_2^3 + 0.6x_2^2x_3 + 1.6x_2x_3^2 + 1.4x_3^3 \end{array}\right\}\right),$$

$$u_2 = -20\tanh\left(\frac{1}{20}\left\{\begin{array}{l} 9.8x_1 + 2.94x_2 + 2.44x_3 \\ -0.2x_1^3 - 0.02x_1^2x_2 + 1.42x_1x_2x_3 \\ +0.12x_1x_2^2 + 2.3x_1^2x_3 + 1.9x_1x_3^2 \\ +0.02x_2^3 + 0.23x_2^2x_3 + 0.57x_2x_3^2 \\ +0.52x_3^3 \end{array}\right\}\right)$$

This is a state feedback control law based on neural networks as shown in Fig. 1. The suitable performance of this saturated control law is revealed in Fig. 4. Note that the controller works fine even when the state trajectory exits the state space region for which training happened. In general however, the algorithm is guaranteed to work for those initial conditions for which the state trajectory remains within the training set.

5.2. *Nonlinear oscillator with constrained input*

We consider next the following nonlinear oscillator:

$$\dot{x}_1 = x_1 + x_2 - x_1(x_1^2 + x_2^2),$$

$$\dot{x}_2 = -x_1 + x_2 - x_2(x_1^2 + x_2^2) + u.$$

It is desired to control the system with control limits of $|u| \leqslant 1$. The following neural network is used:

$$\begin{aligned} V_{24}(x_1, x_2) = {} & w_1x_1^2 + w_2x_2^2 + w_3x_1x_2 + w_4x_1^4 + w_5x_2^4 \\ & + w_6x_1^3x_2 + w_7x_1^2x_2^2 + w_8x_1x_2^3 + w_9x_1^6 \\ & + w_{10}x_2^6 \\ & + w_{11}x_1^5x_2 + w_{12}x_1^4x_2^2 + w_{13}x_1^3x_2^3 \\ & + w_{14}x_1^2x_2^4 + w_{15}x_1x_2^5 \\ & + w_{16}x_1^8 + w_{17}x_2^8 + w_{18}x_1^7x_2 + w_{19}x_1^6x_2^2 \\ & + w_{20}x_1^5x_2^3 \\ & + w_{21}x_1^4x_2^4 + w_{22}x_1^3x_2^5 + w_{23}x_1^2x_2^6 \\ & + w_{24}x_1x_2^7. \end{aligned}$$

This is a power series neural network with 24 activation functions containing powers of the state variable of the system up to the 8th power. Note that the order of the neurons, 8th, required to guarantee convergence was higher than the one in the previous example.

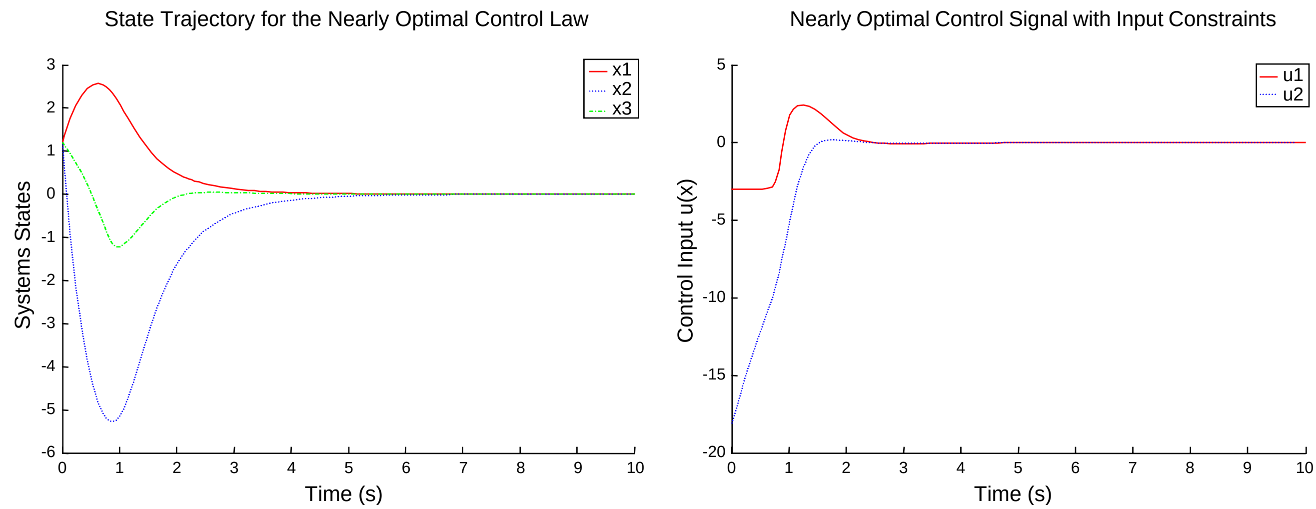

Fig. 4. Nearly optimal nonlinear neural control law considering actuator saturation.

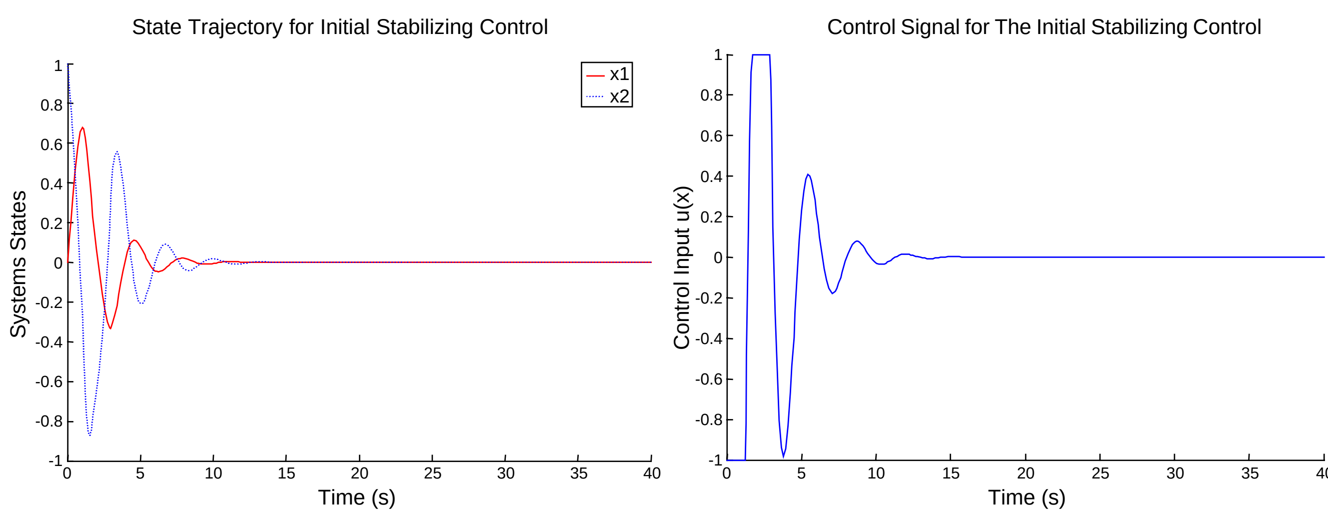

Fig. 5. Performance of the initial stabilizing control when saturated.

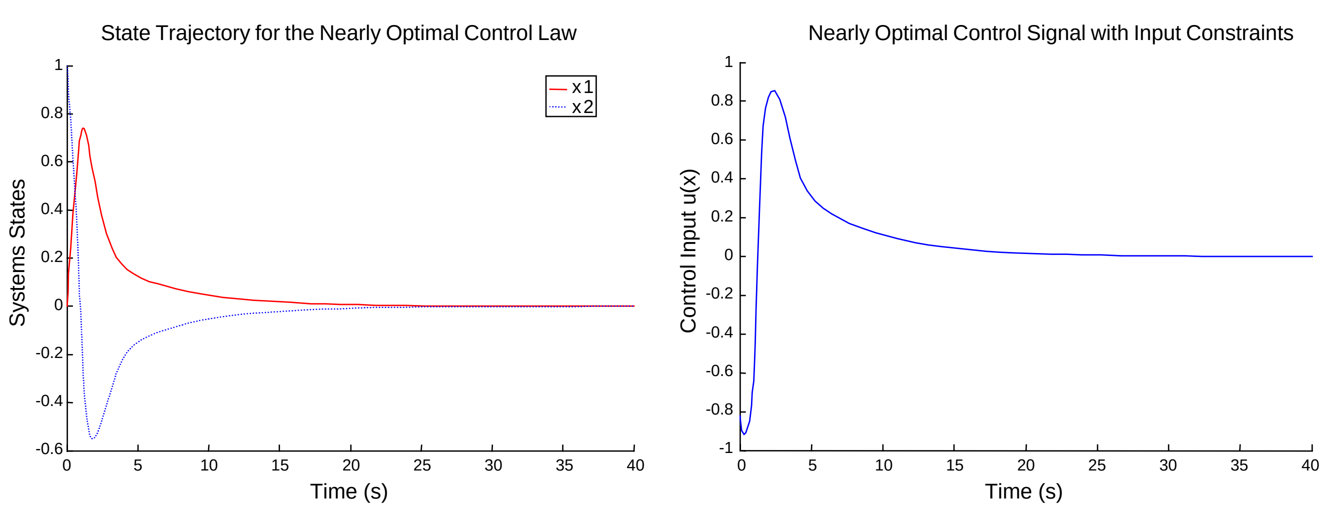

Fig. 6. Nearly optimal nonlinear control law considering actuator saturation.

Fig. 5 shows the performance of the bounded controller $u = sat_{-1}^{+1}(-5x_1 - 3x_2)$ for $x_1(0) = 0,\ x_2(0) = 1$. The algorithm is run over the region $-1 \leqslant x_1 \leqslant 1,\ -1 \leqslant x_2 \leqslant 1$, $R = 1,\ Q = I_{2x2}$. After 20 successive iterations, one has

$$u = -\tanh\begin{pmatrix} 2.6x_1 + 4.2x_2 + 0.4x_2^3 - 4.0x_1^3 - 8.7x_1^2x_2 \\ -8.9x_1x_2^2 - 5.5x_2^5 + 2.26x_1^5 + 5.8x_1^4x_2 \\ +11x_1^3x_2^2 + 2.6x_1^2x_2^3 + 2.00x_1x_2^4 + 2.1x_2^7 \\ -0.5x_1^7 - 1.7x_1^6x_2 - 2.71x_1^5x_2^2 - 2.19x_1^4x_2^3 \\ -0.8x_1^3x_2^4 + 1.8x_1^2x_2^5 + 0.9x_1x_2^6 \end{pmatrix}.$$

This is the control law in terms of a neural network following the structure shown in Fig. 1. Performance of this saturated control law is revealed in Fig. 6. Note that the states and the saturated input in Fig. 6 have fewer oscillations when compared to those of Fig. 5.

6. Conclusion

A rigorous computationally effective algorithm to find nearly optimal constrained control state feedback laws for general nonlinear systems with saturating actuators is presented. The control is given as the output of a neural network. This is an extension of the novel work in Saridis and Lee (1979), Beard (1995) and Lyshevski (2001). Conditions under which the theory of successive approximation applies were shown. Two numerical examples were presented.

Although it was not the focus of this paper, we believe that the results of this paper can be extended to handle constraints on states. Moreover, an issue of interest is how to increase RAS of an initial stabilizing controller. Finally, it would be interesting to show how one can solve for the HJB equation when the system dynamics f, g are unknown as discussed in Murray et al. (2002).

References

Adams, R., & Fournier, J. (2003). *Sobolev spaces*. (2nd ed.), New York: Academic Press.

Apostol, T. (1974). *Mathematical analysis*. Reading, MA: Addison-Wesley.

Bardi, M., & Capuzzo-Dolcetta, I. (1997). *Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations*. Boston, MA: Birkhauser.

Beard, R. (1995). *Improving the closed-loop performance of nonlinear systems*. Ph.D. Thesis, Rensselaer Polytechnic Institute, Troy, NY.

Beard, R., Saridis, G., & Wen, J. (1997). Galerkin approximations of the generalized Hamilton–Jacobi–Bellman equation. *Automatica*, *33*(12), 2159–2177.

Beard, R., Saridis, G., & Wen, J. (1998). Approximate solutions to the time-invariant Hamilton–Jacobi–Bellman equation. *Journal of Optimization Theory and Application*, *96*(3), 589–626.

Bernstein, D. S. (1995). Optimal nonlinear, but continuous, feedback control of systems with saturating actuators. *International Journal of Control*, *62*(5), 1209–1216.

Bertsekas, D. P., & Tsitsiklis, J. N. (1996). *Neuro-dynamic programming*. Belmont, MA: Athena Scientific.

Chen, F. C., & Liu, C. C. (1994). Adaptively controlling nonlinear continuous-time systems using multilayer neural networks. *IEEE Transactions on Automatic Control*, *39*(6), 1306–1310.

Evans, M., & Swartz, T. (2000). *Approximating integrals via Monte Carlo and deterministic methods*. Oxford: Oxford University Press.

Finlayson, B. A. (1972). *The method of weighted residuals and variational principles*. New York: Academic Press.

Han, D., Balakrishnan, S. N. (2000). State-constrained agile missile control with adaptive-critic-based neural networks. *Proceedings of the American control conference* (pp. 1929–1933).

Hornik, K., Stinchcombe, M., & White, H. (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. *Neural Networks*, *3*, 551–560.

Huang, C.-S., Wang, S., & Teo, K. L. (2000). Solving Hamilton–Jacobi–Bellman equations by a modified method of characteristics. *Nonlinear Analysis*, *40*, 279–293.

Huang, J., & Lin, C. F. (1995). Numerical approach to computing nonlinear H_∞ control laws. *Journal of Guidance, Control, and Dynamics*, *18*(5), 989–994.

Khalil, H. (2003). *Nonlinear systems*. (3rd ed.), Upper Saddle River, NJ: Prentice-Hall.

Kim, Y. H., Lewis, F. L., & Dawson, D. (2000). Intelligent optimal control of robotic manipulators using neural networks. *Automatica*, *36*, 1355–1364.

Kleinman, D. (1968). On an iterative technique for Riccati equation computations. *IEEE Transactions on Automatic Control*, 114–115.

Lee, H. W. J., Teo, K. L., Lee, W. R., & Wang, S. (2001). Construction of suboptimal feedback control for chaotic systems using B-splines with optimally chosen knot points. *International Journal of Bifurcation and Chaos*, *11*(9), 2375–2387.

Lewis, F. L., & Syrmos, V. L. (1995). *Optimal control*. New York: Wiley.

Lewis, F. L., Jagannathan, S., & Yesildirek, A. (1999). *Neural network control of robot manipulators and nonlinear systems*. London: Taylor & Francis.

Lio, F. D. (2000). On the Bellman equation for infinite horizon problems with unbounded cost functional. *Applied Mathematics and Optimization*, *41*, 171–197.

Liu, X., Balakrishnan, S. N. (2000). Convergence analysis of adaptive critic based optimal control. *Proceedings of American control conference* (pp. 1929–1933).

Lyshevski, S. E. (1996). Constrained optimization and control of nonlinear systems: new results in optimal control. *Proceedings of the IEEE conference on decision and control* (pp. 541–546).

Lyshevski, S. E. (1998). Optimal control of nonlinear continuous-time systems: design of bounded controllers via generalized nonquadratic functionals. *Proceedings of American control conference*, June 1998 (pp. 205–209).

Lyshevski, S. E. (2001a). *Control systems theory with engineering applications*. Boston, MA: Birkhauser.

Lyshevski, S. E. (2001b). Role of performance functionals in control laws design. *Proceedings of American control conference* (pp. 2400–2405).

Lyshevski, S. E., Meyer, A. U. (1995). Control system analysis and design upon the Lyapunov method.*Proceedings of American control conference* (pp. 3219–3223).

Mikhlin, S. G. (1964). *Variational methods in mathematical physics*. Oxford: Pergamon.

Miller, W. T., Sutton, R., & Werbos, P. (1990). *Neural networks for control*. Cambridge, Massachusetts: The MIT Press.

Munos, R., Baird, L. C., Moore, A. (1999). Gradient descent approaches to neural-net-based solutions of the Hamilton–Jacobi–Bellman equation. *International joint conference on neural networks IJCNN*, 3 (pp. 2152–2157).

Murray, J., Cox, C., Lendaris, G., & Saeks, R. (2002). Adaptive dynamic programming. *IEEE Transactions on Systems, MAN, and Cybernetics-Part C: Applications and Reviews*, *32*(2), 140–153.

Narendra, K. S., & Lewis, F. L. (2001). Special issue on neural network feedback control. *Automatica*, *37*(8), 1147–1148.

Parisini, T., & Zoppoli, R. (1998). Neural approximations for infinite-horizon optimal control of nonlinear stochastic systems. *IEEE Transactions on Neural Networks*, *9*(6), 1388–1408.

Polycarpou, M. M. (1996). Stable adaptive neural control scheme for nonlinear systems. *IEEE Transactions on Automatic Control*, *41*(3), 447–451.

Rovithakis, G. A., & Christodoulou, M. A. (1994). Adaptive control of unknown plants using dynamical neural networks. *IEEE Transactions on Systems, Man, and Cybernetics*, *24*(3), 400–412.

Saberi, A., Lin, Z., & Teel, A. (1996). Control of linear systems with saturating actuators. *IEEE Transactions on Automatic Control*, *41*(3), 368–378.

Sadegh, N. (1993). A perceptron network for functional identification and control of nonlinear systems. *IEEE Transactions on Neural Networks*, *4*(6), 982–988.

Sanner, R. M., Slotine, J. J. E. (1991). Stable adaptive control and recursive identification using radial gaussian networks. *Proceedings of IEEE conference on decision and control* (pp. 2116–2123) Brighton.

Saridis, G., & Lee, C. S. (1979). An approximation theory of optimal control for trainable manipulators. *IEEE Transactions on Systems, Man, Cybernetics*, *9*(3), 152–159.

Sussmann, H., Sontag, E. D., & Yang, Y. (1994). A general result on the stabilization of linear systems using bounded controls. *IEEE Transactions on Automatic Control*, *39*(12), 2411–2425.

Van Der Schaft, A. J. (1992). L_2-gain analysis of nonlinear systems and nonlinear state feedback H_∞ control. *IEEE Transactions on Automatic Control*, *37*(6), 770–784.

Murad Abu-Khalaf was born in Jerusalem, Palestine, in 1977. He did his high school studies at Ibrahimiah College in Jerusalem. He subsequently studied in Cyprus and received his Bachelor's Degree in Electronics and Electrical Engineering from Boğaziçi University in Istanbul, Turkey, in 1998. He then joined The University of Texas at Arlington from which he received the Master's of Science in Electrical Engineering in 2000. Currently he is working on his Ph.D. degree at The University of Texas at Arlington and is working as a research assistant at the Automation and Robotics Research Institute. He is an IEEE member, and a member of Etta Kappa Nu honor society. His interest is in the areas of nonlinear control, optimal control and neural network control.

Frank L. Lewis was born in Würzburg, Germany, subsequently studying in Chile and Gordonstoun School in Scotland. He obtained the Bachelor's Degree in Physics/Electrical Engineering and the Master's of Electrical Engineering Degree at Rice University in 1971. He spent six years in the U.S. Navy, serving as Navigator aboard the frigate USS Trippe (FF-1075), and Executive Officer and Acting Commanding Officer aboard USS Salinan (ATF-161). In 1977 he received the Master's of Science in Aeronautical Engineering from the University of West Florida. In 1981 he obtained the Ph.D. degree at The Georgia Institute of Technology in Atlanta, where he was employed as a Professor from 1981 to 1990 and is currently an Adjunct Professor. He is a Professor of Electrical Engineering at The University of Texas at Arlington, where he was awarded the Moncrief-O'Donnell Endowed Chair in 1990 at the Automation and Robotics Research Institute. He is a Fellow of the IEEE, a member of the New York Academy of Sciences, and a registered Professional Engineer in the State of Texas. He is a Charter Member (2004) of the UTA Academy of Distinguished Scholars. He has served as Visiting Professor at Democritus University in Greece, Hong Kong University of Science and Technology, Chinese University of Hong Kong, National University of Singapore. He is an elected Guest Consulting Professor at both Shanghai Jiao Tong University and South China University of Technology. Dr. Lewis' current interests include intelligent control, neural and fuzzy systems, microelectromechanical systems (MEMS), wireless sensor networks, nonlinear systems, robotics, condition-based maintenance, and manufacturing process control. He is the author/co-author of 3 US patents, 157 journal papers, 23 chapters and encyclopedia articles, 239 refereed conference papers, nine books, including *Optimal Control, Optimal Estimation, Applied Optimal Control and Estimation, Aircraft Control and Simulation, Control of Robot Manipulators*, *Neural Network Control, High-Level Feedback Control with Neural Networks* and the IEEE reprint volume *Robot Control*. He was selected to the Editorial Boards of *International Journal of Control*, *Neural Computing and Applications*, and *Int. J. Intelligent Control Systems*. He served as an Editor for the flagship journal *Automatica*. He is the recipient of an NSF Research Initiation Grant and has been continuously funded by NSF since 1982. Since 1991 he has received $4.8 million in funding from NSF and other government agencies, including significant DoD SBIR and industry funding. His SBIR program was instrumental in ARRI's receipt of the SBA Tibbets Award in 1996. He has received a Fulbright Research Award, the American Society of Engineering Education *F.E. Terman* Award, three Sigma Xi Research Awards, the UTA Halliburton Engineering Research Award, the UTA University-Wide Distinguished Research Award, the ARRI Patent Award, various Best Paper Awards, the IEEE Control Systems Society Best Chapter Award (as Founding Chairman), and the National Sigma Xi Award for Outstanding Chapter (as President). He was selected as Engineer of the year in 1994 by the Ft. Worth IEEE Section. He was appointed to the NAE Committee on Space Station in 1995 and to the IEEE Control Systems Society Board of Governors in 1996. In 1998 he was selected as an IEEE Control Systems Society *Distinguished Lecturer*. He is a Founding Member of the Board of Governors of the Mediterranean Control Association.