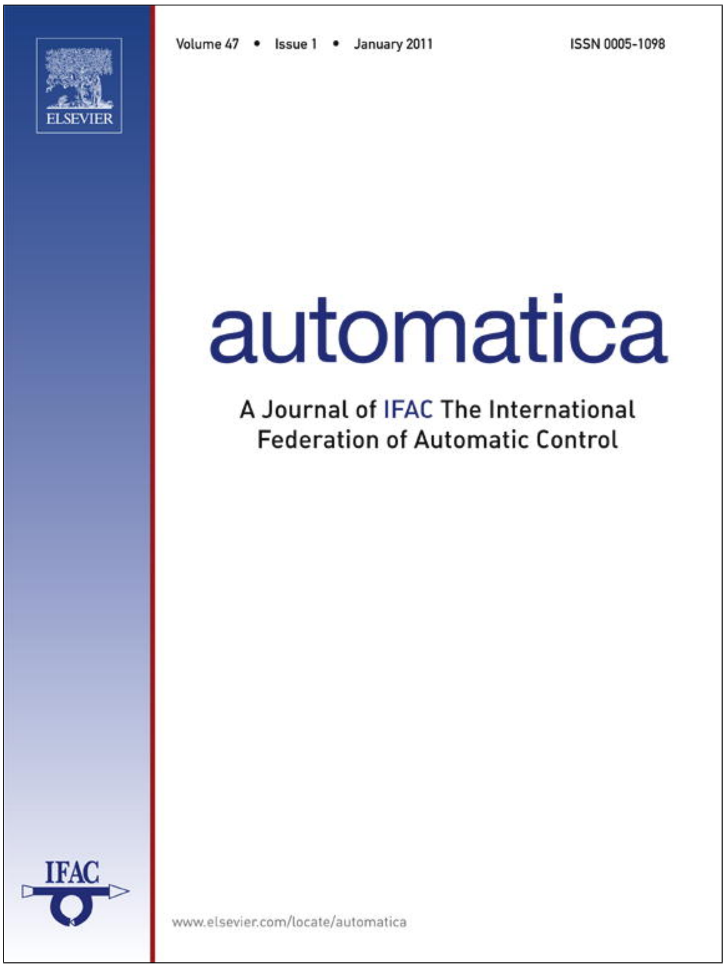
Volume 47 • Issue 1 • January 2011
ISSN 0005-1098
ELSEVIER
automatica
A Journal of IFAC The International Federation of Automatic Control
IFAC
www.elsevier.com/locate/automatica

Automatica 47 (2011) 99–107

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Brief paper

Asymptotic optimal control of uncertain nonlinear Euler–Lagrange systems☆

Keith Dupree, Parag M. Patre, Zachary D. Wilcox, Warren E. Dixon*

Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:
Received 5 June 2009
Received in revised form
15 July 2010
Accepted 16 August 2010
Available online 18 November 2010

Keywords:
Optimal control
Neural networks
RISE
Nonlinear control
Lyapunov-based control

ABSTRACT

A sufficient condition to solve an optimal control problem is to solve the Hamilton–Jacobi–Bellman (HJB) equation. However, finding a value function that satisfies the HJB equation for a nonlinear system is challenging. For an optimal control problem when a cost function is provided a priori, previous efforts have utilized feedback linearization methods which assume exact model knowledge, or have developed neural network (NN) approximations of the HJB value function. The result in this paper uses the implicit learning capabilities of the RISE control structure to learn the dynamics asymptotically. Specifically, a Lyapunov stability analysis is performed to show that the RISE feedback term asymptotically identifies the unknown dynamics, yielding semi-global asymptotic tracking. In addition, it is shown that the system converges to a state space system that has a quadratic performance index which has been optimized by an additional control element. An extension is included to illustrate how a NN can be combined with the previous results. Experimental results are given to demonstrate the proposed controllers.

1. Introduction

Optimal control theory involves the design of controllers that can satisfy some objective while simultaneously minimizing some performance metric. A sufficient condition to solve an optimal control problem is to solve the Hamilton–Jacobi–Bellman (HJB) equation. For the special case of linear time-invariant systems, the HJB equation reduces to an algebraic Riccati equation (ARE); however, for nonlinear systems, finding a value function that satisfies the HJB equation is challenging because it requires the solution of a partial differential equation that may not have an explicit solution. If the nonlinear dynamics are exactly known, then the problem can be reduced to solving an ARE through feedback linearization methods (cf. Freeman & Kokotovic, 1995; Lu, Sun, Xu, & Mochizuki, 1996; Nevistic & Primbs, 1996; Primbs & Nevistic, 1996; Sekoguchi, Konishi, Goto, Yokoyama, & Lu, 2002).

Inverse optimal control (see Freeman & Kokotovi'c, 1996; Krstic & Li, 1998; Krstic & Tsiotras, 1999) is an alternative method to solve the nonlinear optimal control problem by circumventing the need to solve the HJB equation. By finding a control Lyapunov function, which can be shown to also be a value function, an optimal controller can be developed that optimizes a derived cost. In most cases inverse optimal control requires exact knowledge of the nonlinear dynamics, however inverse optimal adaptive control (cf. Fausz, Chellaboina, & Haddad, 1997; Lewis, Syrmos, Li, & Krstic, 1995; Luo, Chu, & Ling, 2005) techniques have been developed for systems with linear in the constant parameters (LP) uncertainty.

Motivated by the desire to eliminate the requirement for exact knowledge of the dynamics for a direct optimal controller (i.e., where the cost function is given a priori), (Johansson, 1990) developed a self-optimizing adaptive controller to yield global asymptotic tracking despite LP uncertainty provided the parameter estimation error could somehow converge to zero. In the preliminary work in Dupree, Patre, Wilcox, and Dixon (2008), we illustrated how a Robust Integral of the Sign of the Error (RISE) feedback controller could be modified to yield a direct optimal controller that achieves semi-global asymptotic tracking. The result in Dupree et al. (2008) exploits the implicit learning characteristic (Qu & Xu, 2002) of the RISE controller to asymptotically cancel LP, non-LP uncertainties and additive disturbances in the dynamics so that the overall control structure converges to an optimal controller.

Researchers have also investigated the use of the universal approximation property of neural networks (NNs) to approximate the LP and non-LP unknown dynamics as a means to develop direct optimal controllers. Specifically, results such as Abu-Khalaf and Lewis (2002), Cheng, Li, and Zhang (2006), Cheng and Lewis (2007), Cheng, Lewis, and Abu-Khalaf (2007), Kim and Lewis (2000) and Kim, Lewis, and Dawson (2000) find an optimal controller for a given cost function for a partially feedback linearized system, and then modify the optimal controller with a NN to approximate the unknown dynamics. Specifically the tracking errors for the

☆ This research is supported in part by the NSF CAREER award number 0547448, NSF award number 0901491. The material in this paper was partially presented at the 47th IEEE CDC, December 9–11, 2008, Cancun, Mexico. This paper was recommended for publication in revised form by Associate Editor Derong Liu under the direction of Editor Miroslav Krstic.

* Corresponding author. Tel.: +1 352 846 1463; fax: +1 352 392 7303.
E-mail addresses: kdupree@gmail.com (K. Dupree), parag.patre@gmail.com (P.M. Patre), zibrus@gmail.com (Z.D. Wilcox), wdixon@ufl.edu (W.E. Dixon).

0005-1098/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.automatica.2010.10.007

NN methods are proven to be uniformly ultimately bounded (UUB) and the resulting state space system, for which the HJB optimal controller is developed, is only approximated.

The contribution in this work, and our preliminary efforts in Dupree et al. (2008), arise from incorporating optimal control elements with an implicit learning feedback control strategy coined the Robust Integral of the Sign of the Error (RISE) method in Patre, MacKunis, Makkar, and Dixon (2008). The RISE method is used to identify the system and reject disturbances, while achieving asymptotic tracking and the convergence of a control term to the optimal controller. Inspired by the previous work in Abu-Khalaf and Lewis (2002), Abu-Khalaf, Huang, and Lewis (2006), Cheng et al. (2006), Cheng and Lewis (2007), Cheng et al. (2007), Johansson (1990), Kim and Lewis (2000), Kim et al. (2000) and Lewis (1986) a system in which all terms are assumed known (temporarily) is feedback linearized and a control law is developed based on the HJB optimization method for a given quadratic performance index. Under the assumption that parametric uncertainty and unknown bounded disturbances are present in the dynamics, the control law is modified to contain the RISE feedback term which is used to identify the uncertainty. A Lyapunov stability analysis indicates that the RISE feedback term asymptotically identifies the unknown dynamics (yielding semi-global asymptotic tracking) provided upper bounds on the disturbances are known and the control gains are selected appropriately. A distinction in this work is that the uncertain nonlinear disturbances are asymptotically identified (rather than UUB), allowing the developed controller to asymptotically converge to an optimal controller for the residual uncertain nonlinear system.

An extension is included that investigates the amalgam of the robust RISE feedback method with NN methods to yield a direct optimal controller. Combining a NN feedforward controller with the RISE feedback method yields an asymptotic result (Patre, MacKunis, Kaiser, & Dixon, 2008), that can also be proven to converge to an optimal controller through the efforts in this paper. Hence, a modification to the results in Abu-Khalaf and Lewis (2002), Cheng et al. (2006), Cheng and Lewis (2007), Cheng et al. (2007), Kim and Lewis (2000) and Kim et al. (2000) and is provided that allows for asymptotic tracking and convergence to the optimal controller.

2. Dynamic model

The class of considered nonlinear dynamic systems is assumed to be modeled by the following Euler–Lagrange formulation:

$$M(q)\ddot{q} + V_m(q,\dot{q})\dot{q} + G(q) + F(\dot{q}) + \tau_d(t) = \tau(t). \tag{1}$$

In (1), $M(q) \in \mathbb{R}^{n\times n}$ denotes the generalized inertia matrix, $V_m(q,\dot{q}) \in \mathbb{R}^{n\times n}$ denotes the generalized centripetal-Coriolis matrix, $G(q) \in \mathbb{R}^n$ denotes the generalized gravity vector, $F(\dot{q}) \in \mathbb{R}^n$ denotes generalized friction, $\tau_d(t) \in \mathbb{R}^n$ denotes a general disturbance (e.g., unmodeled effects), $\tau(t) \in \mathbb{R}^n$ represents the input control vector, and $q(t), \dot{q}(t), \ddot{q}(t) \in \mathbb{R}^n$ denote the generalized position, velocity, and acceleration vectors, respectively. The subsequent development is based on the assumption that $q(t)$ and $\dot{q}(t)$ are measurable and that $M(q)$, $V_m(q,\dot{q})$, $G(q)$, $F(\dot{q})$ and $\tau_d(t)$ are unknown. Moreover, the following assumptions will be exploited in the subsequent development.

Assumption 1. The inertia matrix $M(q)$ is symmetric, positive definite, and satisfies the following inequality $\forall \xi(t) \in \mathbb{R}^n$:

$$m_1\|\xi\|^2 \le \xi^T M(q)\xi \le \bar{m}(q)\|\xi\|^2, \tag{2}$$

where $m_1 \in \mathbb{R}$ is a known positive constant, $\bar{m}(q) \in \mathbb{R}$ is a known positive function, and $\|\cdot\|$ denotes the standard Euclidean norm.

Assumption 2. The following skew-symmetric relationship is satisfied:

$$\xi^T(\dot{M}(q) - 2V_m(q,\dot{q}))\xi = 0 \quad \forall \xi \in \mathbb{R}^n. \tag{3}$$

Assumption 3. If $q(t), \dot{q}(t) \in \mathcal{L}_\infty$, then $V_m(q,\dot{q})$, $F(\dot{q})$ and $G(q)$ are bounded. Moreover, if $q(t), \dot{q}(t) \in \mathcal{L}_\infty$, then the first and second partial derivatives of the elements of $M(q)$, $V_m(q,\dot{q})$, $G(q)$ with respect to $q(t)$ exist and are bounded, and the first and second partial derivatives of the elements of $V_m(q,\dot{q})$, $F(\dot{q})$ with respect to $\dot{q}(t)$ exist and are bounded.

Assumption 4. The desired trajectory is assumed to be designed such that $q_d(t), \dot{q}_d(t), \ddot{q}_d(t), \dddot{q}_d(t), \ddddot{q}_d(t) \in \mathbb{R}^n$ exist, and are bounded.

Assumption 5. The nonlinear disturbance term and its first two time derivatives, i.e. $\tau_d(t), \dot{\tau}_d(t), \ddot{\tau}_d(t)$ are bounded by known constants.

3. Control objective

The control objective is to ensure that the system tracks a desired time-varying trajectory, denoted by $q_d(t) \in \mathbb{R}^n$, despite uncertainties in the dynamic model, while minimizing a given performance index. To quantify the tracking objective, a position tracking error, denoted by $e_1(t) \in \mathbb{R}^n$, is defined as

$$e_1 \triangleq q_d - q. \tag{4}$$

To facilitate the subsequent analysis, filtered tracking errors, denoted by $e_2(t), r(t) \in \mathbb{R}^n$, are also defined as

$$e_2 \triangleq \dot{e}_1 + \alpha_1 e_1 \tag{5}$$

$$r \triangleq \dot{e}_2 + \alpha_2 e_2, \tag{6}$$

where $\alpha_1 \in \mathbb{R}^{n\times n}$, denotes a subsequently defined positive definite, constant, gain matrix, and $\alpha_2 \in \mathbb{R}$ is a positive constant. The subsequent development is based on the assumption that $q(t)$ and $\dot{q}(t)$ are measurable, so the filtered tracking error $r(t)$ is not measurable since the expression in (6) depends on $\ddot{q}(t)$. The error systems are based on the assumption that the generalized coordinates of the Euler–Lagrange dynamics allow additive and not multiplicative errors.

4. Optimal control design

In this section, a state-space model is developed based on the tracking errors in (4) and (5). Based on this model, a controller is developed that minimizes a quadratic performance index under the (temporary) assumption that the dynamics in (1), including the additive disturbance, are known. The development in this section motivates the control design in Section 5, where a robust controller is developed to identify the unknown dynamics and additive disturbance.

To develop a state-space model for the tracking errors in (4) and (5), the inertia matrix is premultiplied by the time derivative of (5), and substitutions are made from (1) and (4) to obtain

$$M\dot{e}_2 = -V_m e_2 - \tau + h + \tau_d, \tag{7}$$

where the nonlinear function $h(q, \dot{q}, q_d, \dot{q}_d, \ddot{q}_d) \in \mathbb{R}^n$ is defined as

$$h \triangleq M(\ddot{q}_d + \alpha_1\dot{e}_1) + V_m(\dot{q}_d + \alpha_1 e_1) + G + F. \tag{8}$$

Under the (temporary) assumption that the dynamics in (1) are known, the control input can be designed as

$$\tau \triangleq h + \tau_d - u, \tag{9}$$

where $u(t) \in \mathbb{R}^n$ is an auxiliary control input that will be designed to minimize a subsequent performance index. By substituting (9) into (7) the closed-loop error system for $e_2(t)$ can be obtained as

$$M\dot{e}_2 = -V_m e_2 + u. \quad (10)$$

A time-varying state-space model for (5) and (10) can now be developed as

$$\dot{z} = A(q,\dot{q})z + B(q)u, \quad (11)$$

where $A(q,\dot{q}) \in \mathbb{R}^{2n\times 2n}$, $B(q) \in \mathbb{R}^{2n\times n}$, and $z(t) \in \mathbb{R}^{2n}$ are defined as

$$A(q,\dot{q}) \triangleq \begin{bmatrix} -\alpha_1 & I_{n\times n} \\ 0_{n\times n} & -M^{-1}V_m \end{bmatrix}, \qquad B(q) \triangleq \begin{bmatrix} 0_{n\times n} & M^{-1} \end{bmatrix}^T$$

$$z(t) \triangleq \begin{bmatrix} e_1^T & e_2^T \end{bmatrix}^T$$

where $I_{n\times n}$ and $0_{n\times n}$ denote an $n \times n$ identity matrix and matrix of zeros, respectively. The quadratic performance index $J(u) \in \mathbb{R}$ to be minimized subject to the constraints in (11) is

$$J(u) = \int_0^\infty \left(\frac{1}{2}z^TQz + \frac{1}{2}u^TRu\right) dt. \quad (12)$$

In (12), $Q \in \mathbb{R}^{2n\times 2n}$ and $R \in \mathbb{R}^{n\times n}$ are positive definite symmetric matrices to weight the influence of the states and (partial) control effort, respectively. Furthermore, the matrix Q can be broken into blocks as:

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^T & Q_{22} \end{bmatrix}.$$

Given the performance index $J(u)$, the control objective is to find the auxiliary control input $u(t)$ that minimizes (12) subject to the differential constraints imposed by (11). The optimal control that achieves this objective is denoted by $u^*(t)$. As stated in Kim and Lewis (2000) and Kim et al. (2000), the fact that the performance index is only penalized for the auxiliary control $u(t)$ is practical since the gravity, Coriolis, and friction compensation terms in (8) cannot be modified by the optimal design phase.

The controller $u^*(t)$ minimizes (12) subject to (11) if and only if there exists a value function $V(z,t)$ where

$$-\frac{\partial V}{\partial t} = \frac{1}{2}z^TQz + \frac{1}{2}u^{*T}Ru^* + \frac{\partial V}{\partial z}\dot{z}$$

that satisfies the HJB equation (Lewis, Syrmos, Li, & Krstic, 1995)

$$\frac{\partial V}{\partial t} + \min_u \left[H\left(z,u,\frac{\partial V}{\partial z},t\right)\right] = 0 \quad (13)$$

where the Hamiltonian of optimization $H(z,u,\frac{\partial V}{\partial z},t) \in \mathbb{R}$ is defined as

$$H\left(z,u,\frac{\partial V}{\partial z},t\right) = \frac{1}{2}z^TQz + \frac{1}{2}u^TRu + \frac{\partial V}{\partial z}\dot{z}.$$

The minimum of (12) is obtained for the optimal controller $u(t) = u^*(t)$ where the respective Hamiltonian is

$$H^* = \min_u \left[H\left(z,u,\frac{\partial V}{\partial z},t\right)\right] = -\frac{\partial V}{\partial t}. \quad (14)$$

To facilitate the subsequent development, let $P(q) \in \mathbb{R}^{2n\times 2n}$ be a positive definite symmetric matrix defined as

$$P(q) = \begin{bmatrix} K & 0_{n\times n} \\ 0_{n\times n} & M \end{bmatrix}, \quad (15)$$

where $K \in \mathbb{R}^{n\times n}$ denotes a positive definite symmetric gain matrix. If α_1, R, and K, introduced in (5), (12) and (15), satisfy the following algebraic relationships

$$K = -Q_{12} = -Q_{12}^T > 0 \quad (16)$$

$$Q_{11} = \alpha_1^TK + K\alpha_1, \quad (17)$$

$$R^{-1} = Q_{22}, \quad (18)$$

then Theorem 1 of Kim and Lewis (2000) and Kim et al. (2000) can be invoked to prove that $P(q)$ satisfies

$$z^T(PA + A^TP^T - PBR^{-1}B^TP + \dot{P} + Q)z = 0, \quad (19)$$

and the value function $V(z,t) \in \mathbb{R}$

$$V = \frac{1}{2}z^TPz \quad (20)$$

satisfies the HJB equation in (14). Lemma 1 of Kim and Lewis (2000) and Kim et al. (2000) can be used to conclude that the optimal control $u^*(t)$ that minimizes (12) subject to (11) is

$$u^*(t) = -R^{-1}B^T\left(\frac{\partial V(z,t)}{\partial z}\right)^T = -R^{-1}e_2. \quad (21)$$

Remark 1. An infinite horizon optimal controller for a linear time-invariant system can be developed by solving an ARE. The constants that make up the ARE are a state weighting matrix Q and a control weighting matrix R which come from the typical quadratic cost functional and can be independently selected. Solving the infinite horizon optimal controller for the nonlinear time-varying system in (11) (sufficiently) involves finding the solution to a more general HJB equation. As described in Kim et al. (2000) and Kim and Lewis (2000), a solution to the HJB equation in (13) involves solving the matrix equation in (19). To solve (19), a strategic choice of $P(q)$ is selected and (3) is used to reduce the problem to a set of algebraic constraints/gain conditions. The resulting gain conditions indicate that Q_{11} can be selected to weight the output state $e_1(t)$, Q_{12} can be independently selected to weight the state vector cross terms, and that R can be independently selected to weight the control vector in the cost functional. However, for this approach, the weighting matrix Q_{22} for the auxiliary state $e_2(t)$ is dependent on the control vector weight matrix (i.e., $Q_{22} = R^{-1}$). For alternative methods to solve the infinite horizon optimal control problem for time-varying systems see Kirk (2004) and Lewis et al. (1995).

5. RISE feedback control development

In general, the bounded disturbance $\tau_d(t)$ and the nonlinear dynamics given in (8) are unknown, so the controller given in (9) cannot be implemented. However, if the control input contains some method to identify and cancel these effects, then $z(t)$ will converge to the state space model in (11) so that $u(t)$ minimizes the respective performance index. As stated in the introduction, several results have explored this strategy using function approximation methods such as neural networks, where the tracking control errors converge to a neighborhood near the state space model yielding a type of approximate optimal controller. In this section, a control input is developed that exploits RISE feedback to identify the nonlinear effects and bounded disturbances to enable asymptotic convergence to the state space model.

To develop the control input, the error system in (6) is premultiplied by $M(q)$ and the expressions in (1), (4) and (5) are utilized to obtain

$$Mr = -V_me_2 + h + \tau_d + \alpha_2Me_2 - \tau. \quad (22)$$

Based on the open-loop error system in (22), the control input is composed of the optimal control developed in (21), plus a subsequently designed auxiliary control term $\mu(t) \in \mathbb{R}^n$ as

$$\tau \triangleq \mu - u. \tag{23}$$

The closed-loop tracking error system can be developed by substituting (23) into (22) as

$$Mr = -V_m e_2 + h + \tau_d + \alpha_2 M e_2 + u - \mu. \tag{24}$$

To facilitate the subsequent stability analysis the auxiliary function $f_d(q_d, \dot{q}_d, \ddot{q}_d) \in \mathbb{R}^n$, which is defined as

$$f_d \triangleq M(q_d)\ddot{q}_d + V_m(q_d, \dot{q}_d)\dot{q}_d + G(q_d) + F(\dot{q}_d), \tag{25}$$

is added and subtracted to (24) to yield

$$Mr = -V_m e_2 + \bar{h} + f_d + \tau_d + u - \mu + \alpha_2 M e_2, \tag{26}$$

where $\bar{h}(q, \dot{q}, q_d, \dot{q}_d, \ddot{q}_d) \in \mathbb{R}^n$ is defined as

$$\bar{h} \triangleq h - f_d. \tag{27}$$

The time derivative of (26) can be written as

$$M\dot{r} = -\frac{1}{2}\dot{M}r + \tilde{N} + N_D - e_2 - R^{-1}r - \dot{\mu} \tag{28}$$

after strategically grouping specific terms. In (28), the unmeasurable auxiliary terms $\tilde{N}(e_1, e_2, r, t), N_D(t) \in \mathbb{R}^n$ are defined as

$$\begin{aligned} \tilde{N} \triangleq\ & -\dot{V}_m e_2 - V_m \dot{e}_2 - \frac{1}{2}\dot{M}r + \dot{\bar{h}} + \alpha_2 \dot{M} e_2 \\ & + \alpha_2 M \dot{e}_2 + e_2 + \alpha_2 R^{-1} e_2 \end{aligned}$$

$$N_D \triangleq \dot{f}_d + \dot{\tau}_d.$$

Motivation for grouping terms into $\tilde{N}(e_1, e_2, r, t)$ and $N_D(t)$ comes from the subsequent stability analysis and the fact that the Mean Value Theorem and Assumptions 3–5 can be used to upper bound the auxiliary terms as

$$\|\tilde{N}(t)\| \le \rho(\|y\|)\|y\|, \tag{29}$$

$$\|N_D\| \le \zeta_1, \qquad \|\dot{N}_D\| \le \zeta_2, \tag{30}$$

where $y(t) \in \mathbb{R}^{3n}$ is defined as

$$y(t) \triangleq [e_1^T \ e_2^T \ r^T]^T, \tag{31}$$

the bounding function $\rho(\|y\|) \in \mathbb{R}$ is a positive globally invertible nondecreasing function, and $\zeta_i \in \mathbb{R}$ $(i = 1, 2)$ denote known positive constants. Based on (28), the control term $\mu(t)$ is designed based on the RISE framework (see Patre et al., 2008; Qu & Xu, 2002; Xian, Dawson, de Queiroz, & Chen, 2004) as

$$\mu(t) \triangleq (k_s + 1)e_2(t) - (k_s + 1)e_2(0) + \nu(t) \tag{32}$$

where $\nu(t) \in \mathbb{R}^n$ is the generalized solution to

$$\dot{\nu} = (k_s + 1)\alpha_2 e_2 + \beta_1 sgn(e_2),$$

$k_s \in \mathbb{R}$ is a positive constant control gain, and $\beta_1 \in \mathbb{R}$ is a positive control gain selected according to the following sufficient condition

$$\beta_1 > \zeta_1 + \frac{1}{\alpha_2}\zeta_2. \tag{33}$$

The closed loop error systems for $r(t)$ can now be obtained by substituting the time derivative of (32) into (28) as

$$\begin{aligned} M\dot{r} =\ & -\frac{1}{2}\dot{M}r + \tilde{N} + N_D - e_2 - R^{-1}r - (k_s + 1)r \\ & - \beta_1 sgn(e_2). \end{aligned} \tag{34}$$

To facilitate the subsequently stability analysis, the control gains α_1 and α_2 introduced in (5) and (6), respectively, are selected according to the sufficient conditions:

$$\lambda_{\min}(\alpha_1) > \frac{1}{2} \qquad \alpha_2 > 1, \tag{35}$$

where $\lambda_{\min}(\alpha_1)$ is the minimum eigenvalue of α_1. The control gain α_1 cannot be arbitrarily selected, rather it is calculated using a Lyapunov equation solver. Its value is determined based on the value of Q and R. Therefore Q and R must be chosen such that (35) is satisfied.

6. Stability and optimality analysis

Theorem 1. *The controller given in (21) and (23) ensures that all system signals are bounded under closed-loop operation, and the tracking errors are regulated in the sense that*

$$\|y(t)\| \to 0 \quad \text{as } t \to \infty \quad \forall y(0) \in \mathcal{S} \tag{36}$$

where the set $\mathcal{S}$ can be made arbitrarily large by selecting k_s based on the initial conditions of the system (i.e., a semi-global result). The boundedness of the closed loop signals and the result in (36) can be obtained provided the sufficient conditions in (33) and (35) are satisfied. Furthermore, $u(t)$ converges to an optimal controller that minimizes (12) subject to (11) provided the gain conditions given in (16)–(18) are satisfied.

Proof. Let $\mathcal{D} \subset \mathbb{R}^{3n+1}$ be a domain containing $\Phi(t) = 0$, where $\Phi(t) \in \mathbb{R}^{3n+1}$ is defined as

$$\Phi(t) \triangleq [y^T(t) \ \sqrt{O(t)}]^T. \tag{37}$$

In (37), the auxiliary function $O(t) \in \mathbb{R}$ is defined as

$$O(t) \triangleq \beta_1 \sum_{i=1}^{n} |e_{2_i}(0)| - e_2(0)^T N_D(0) - L(t), \tag{38}$$

where $e_{2_i}(0)$ is equal to the ith element of $e_2(0)$ and the auxiliary function $L(t) \in \mathbb{R}$ is the generalized solution to

$$\dot{L}(t) \triangleq r^T(N_D(t) - \beta_1 sgn(e_2)), \tag{39}$$

where $\beta_1 \in \mathbb{R}$ is a positive constant chosen according to the sufficient conditions in (33). Provided the sufficient conditions introduced in (33) are satisfied, the following inequality can be obtained (Xian et al., 2004):

$$L(t) \le \beta_1 \sum_{i=1}^{n} |e_{2_i}(0)| - e_2(0)^T N_D(0). \tag{40}$$

Hence, (40) can be used to conclude that $O(t) \ge 0$.

Let $V_L(\Phi, t) : \mathcal{D} \times [0, \infty) \to \mathbb{R}$ be a continuously differentiable positive definite function defined as

$$V_L(\Phi, t) \triangleq e_1^T e_1 + \frac{1}{2}e_2^T e_2 + \frac{1}{2}r^T M(q) r + O, \tag{41}$$

which satisfies the following inequalities:

$$U_1(\Phi) \le V_L(\Phi, t) \le U_2(\Phi). \tag{42}$$

In (42), the continuous positive definite functions $U_1(\Phi)$, and $U_2(\Phi) \in \mathbb{R}$ are defined as $U_1(\Phi) \triangleq \lambda_1\|\Phi\|^2$, and $U_2(\Phi) \triangleq \lambda_2(q)\|\Phi\|^2$, where $\lambda_1, \lambda_2(q) \in \mathbb{R}$ are defined as

$$\lambda_1 \triangleq \frac{1}{2}\min\{1, m_1\} \qquad \lambda_2(q) \triangleq \max\left\{\frac{1}{2}\bar{m}(q), 1\right\},$$

where m_1 and $\bar{m}(q)$ are introduced in (2). After taking the time derivative of (41), and using (5), (6), (34), and substituting for the time derivative of $O(t)$

$$\dot{V}_L(\Phi, t) = -2e_1^T\alpha_1 e_1 + 2e_2^T e_1 + r^T\tilde{N}(t) - (k_s + 1)r^T r - R^{-1}r^T r - \alpha_2 e_2^T e_2. \quad (43)$$

Based on the fact that

$$e_2^T e_1 \le \frac{1}{2}\|e_1\|^2 + \frac{1}{2}\|e_2\|^2, \quad (44)$$

the expression in (43) can be upper bounded as

$$\dot{V}_L(\Phi, t) \le r^T\tilde{N}(t) - (k_s + 1 + \lambda_{\min}(R^{-1}))\|r\|^2 - (2\lambda_{\min}(\alpha_1) - 1)\|e_1\|^2 - (\alpha_2 - 1)\|e_2\|^2. \quad (45)$$

By using (29), the expression in (45) can be rewritten as

$$\dot{V}_L(\Phi, t) \le -\lambda_3\|y\|^2 - [k_s\|r\|^2 - \rho(\|y\|)\|r\|\|y\|], \quad (46)$$

where $\lambda_3 \triangleq \min\{2\lambda_{\min}(\alpha_1) - 1, \alpha_2 - 1, 1 + \lambda_{\min}(R^{-1})\}$, and α_1 and α_2 are chosen according to the sufficient condition in (35). After completing the squares for the terms inside the brackets in (46), the following expression can be obtained

$$\dot{V}_L(\Phi, t) \le -\lambda_3\|y\|^2 + \frac{\rho^2(\|y\|)\|y\|^2}{4k_s}. \quad (47)$$

The expression in (47) can be further upper bounded by a continuous, positive semi-definite function

$$\dot{V}_L(\Phi, t) \le -U(\Phi) = c\|y\|^2 \quad \forall \Phi \in \mathcal{D}$$

for some positive constant c, where

$$\mathcal{D} \triangleq \{\Phi \in \mathbb{R}^{3n+1} \mid \|\Phi\| \le \rho^{-1}(2\sqrt{\lambda_3 k_s})\}.$$

The inequalities in (42) and (47) can be used to show that $V_L(\Phi, t) \in \mathcal{L}_\infty$ in $\mathcal{D}$; hence, $e_1(t)$, $e_2(t)$, and $r(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Given that $e_1(t)$, $e_2(t)$, and $r(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, standard linear analysis methods can be used to prove that $\dot{e}_1(t), \dot{e}_2(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$ from (5) and (6). Since $e_1(t), e_2(t), r(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, the assumption that $q_d(t), \dot{q}_d(t), \ddot{q}_d(t)$ exist and are bounded can be used along with (4)–(6) to conclude that $q(t), \dot{q}(t), \ddot{q}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Since $q(t), \dot{q}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, Assumption 3 can be used to conclude that $M(q), V_m(q, \dot{q}), G(q)$, and $F(\dot{q}) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Thus from (1) and Assumption 4, we can show that $\tau(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Given that $r(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, it can be shown that $\dot{\mu}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Since $\dot{q}(t), \ddot{q}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, Assumption 3 can be used to show that $\dot{V}_m(q, \dot{q}), \dot{G}(q), \dot{F}(q)$ and $\dot{M}(q) \in \mathcal{L}_\infty$ in $\mathcal{D}$; hence, (34) can be used to show that $\dot{r}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Since $\dot{e}_1(t), \dot{e}_2(t), \dot{r}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, the definitions for $U(y)$ and $z(t)$ can be used to prove that $U(y)$ is uniformly continuous in $\mathcal{D}$.

Let $\mathcal{S} \subset \mathcal{D}$ denote a set defined as:

$$\mathcal{S} \triangleq \{\Phi(t) \in \mathcal{D} \mid U_2(\Phi(t)) < \lambda_1(\rho^{-1}(2\sqrt{\lambda_3 k_s}))^2\}. \quad (48)$$

The region of attraction in (48) can be made arbitrarily large to include any initial conditions by increasing the control gain k_s (i.e., a semi-global type of stability result) (Xian et al., 2004). The LaSalle–Yoshizawa Theorem (see Theorem 8.4 of Khalil, 2002) can now be invoked to state that

$$c\|y(t)\|^2 \to 0 \quad \text{as } t \to \infty \quad \forall y(0) \in \mathcal{S}. \quad (49)$$

Based on the definition of $y(t)$, (49) can be used to conclude the results in (36) $\forall y(0) \in \mathcal{S}$. From (21), $u(t) \to 0$ as $e_2(t) \to 0$, therefore (26) can be used to conclude that

$$\mu \to \bar{h} + f_d + \tau_d \quad \text{as } r(t), \quad e_2(t) \to 0. \quad (50)$$

From (27), (49) and (50)

$$\mu \to h + \tau_d \quad \text{as } t \to \infty. \quad (51)$$

Using (51), and comparing (9) to (23) indicates that the dynamics in (1) converge to the state-space system in (11). Hence, $u(t)$ converges to an optimal controller that minimizes (12) subject to (11) provided the gain conditions given in (16)–(18), (33) and (35) are satisfied.

7. Neural network extension

The efforts in this section investigate the amalgam of the robust RISE feedback method with NN methods to yield a direct optimal controller. These efforts provide a modification to the results in Abu-Khalaf and Lewis (2002), Cheng et al. (2006), Cheng and Lewis (2007), Cheng et al. (2007), Kim and Lewis (2000) and Kim et al. (2000) that allows for asymptotic stability and convergence to the optimal controller rather than to approximate the optimal controller.

7.1. Feedforward NN estimation

The universal approximation property indicates that weights and thresholds exist such that some continuous function $f(x) \in \mathbb{R}^{N_1+1}$ can be represented by a three-layer NN as (Ge, Hang, & Zhang, 1999; Lewis, 1999; Lewis, Selmic, & Campos, 2002)

$$f(x) = W^T\sigma(V^T x) + \varepsilon(x). \quad (52)$$

In (52), $V \in \mathbb{R}^{(N_1+1)\times N_2}$ and $W \in \mathbb{R}^{(N_2+1)\times n}$ are bounded constant ideal weight matrices for the first-to-second and second-to-third layers respectively, where N_1 is the number of neurons in the input layer, N_2 is the number of neurons in the hidden layer, and n is the number of neurons in the third layer. The activation function in (52) is denoted by $\sigma(\cdot) : \mathbb{R}^{N_1+1} \to \mathbb{R}^{N_2+1}$, and $\varepsilon(x) : \mathbb{R}^{N_1+1} \to \mathbb{R}^n$ is the functional reconstruction error. Based on (52), the typical three-layer NN approximation for $f(x)$ is given as (Ge et al., 1999; Lewis, 1999; Lewis et al., 2002)

$$\hat{f}(x) \triangleq \hat{W}^T\sigma(\hat{V}^T x), \quad (53)$$

where $\hat{V}(t) \in \mathbb{R}^{(N_1+1)\times N_2}$ and $\hat{W}(t) \in \mathbb{R}^{(N_2+1)\times n}$ are subsequently designed estimates of the ideal weight matrices. The estimate mismatches for the ideal weight matrices, denoted by $\tilde{V}(t) \in \mathbb{R}^{(N_1+1)\times N_2}$ and $\tilde{W}(t) \in \mathbb{R}^{(N_2+1)\times n}$, are defined as

$$\tilde{V} \triangleq V - \hat{V}, \qquad \tilde{W} \triangleq W - \hat{W},$$

and the mismatch for the hidden-layer output error for a given $x(t)$, denoted by $\tilde{\sigma}(x) \in \mathbb{R}^{N_2+1}$, is defined as

$$\tilde{\sigma} \triangleq \sigma - \hat{\sigma} = \sigma(V^T x) - \sigma(\hat{V}^T x). \quad (54)$$

Assumption 6. Ideal weights are assumed to exist and be bounded by known positive values so that

$$\|V\|_F^2 = tr(V^T V) \le \bar{V}_B \quad (55)$$

$$\|W\|_F^2 = tr(W^T W) \le \bar{W}_B \quad (56)$$

where $\|\cdot\|_F$ is the Frobenius norm of a matrix, and $tr(\cdot)$ is the trace of a matrix.

To develop the control input, the error system in (6) is premultiplied by $M(q)$ and the expressions in (1), (4) and (5) are utilized to obtain

$$Mr = -V_m e_2 + \bar{h} + f_d + \tau_d + \alpha_2 M e_2 - \tau, \quad (57)$$

where (25) and (27) were used. The auxiliary function $f_d(q_d, \dot{q}_d, \ddot{q}_d) \in \mathbb{R}^n$ in (25) can be represented by a three-layer NN as (Ge et al., 1999; Lewis, 1999; Lewis et al., 2002)

$$f_d = W^T\sigma(V^T x_d) + \varepsilon(x_d). \quad (58)$$

In the above equation, the input $x_d(t) \in \mathbb{R}^{3n+1}$ is defined as $x_d(t) \triangleq [\,1 \quad q_d^T(t) \quad \dot{q}_d^T(t) \quad \ddot{q}_d^T(t)\,]^T$ so that $N_1 = 3n$ where N_1 was introduced in (52). Based on the assumption that the desired trajectory is bounded, the following inequalities hold

$$\|\varepsilon\| \le \varepsilon_{b_1}, \qquad \|\dot{\varepsilon}\| \le \varepsilon_{b_2}, \qquad \|\ddot{\varepsilon}\| \le \varepsilon_{b_3}, \quad (59)$$

where $\varepsilon_{b_1}, \varepsilon_{b_2}, \varepsilon_{b_3} \in \mathbb{R}$ are known positive constants.

7.2. Closed-loop error system

Based on the open-loop error system in (57), the control input is composed of the optimal control developed in (21), a three-layer NN feedforward term, plus the RISE feedback term in (32) as

$$\tau \triangleq \hat{f}_d + \mu - u. \quad (60)$$

The feedforward NN component in (60), denoted by $\hat{f}_d(t) \in \mathbb{R}^n$, is defined as

$$\hat{f}_d \triangleq \hat{W}^T\sigma(\hat{V}^T x_d). \quad (61)$$

The estimates for the NN weights in (61) are generated on-line (there is no off-line learning phase) as

$$\dot{\hat{W}} = proj(\Gamma_1\hat{\sigma}'\hat{V}^T\dot{x}_d e_2^T) \quad (62)$$

$$\dot{\hat{V}} = proj(\Gamma_2\dot{x}_d(\hat{\sigma}'^T\hat{W}e_2)^T)$$

where $\sigma'(\hat{V}^T x) \equiv d\sigma(V^T x)/d(V^T x)|_{V^T x = \hat{V}^T x}$; $\Gamma_1 \in \mathbb{R}^{(N_2+1)\times(N_2+1)}$ and $\Gamma_2 \in \mathbb{R}^{(3n+1)\times(3n+1)}$ are constant, positive definite, symmetric matrices. In (62), $proj(\cdot)$ denotes a smooth convex projection algorithm that ensures $\hat{W}(t)$ and $\hat{V}(t)$ remain bounded inside known bounded convex regions. See Section 4.3 in Dixon, Fang, Dawson, and Flynn (2003) for further details.

The closed-loop tracking error system is obtained by substituting (60) into (57) as

$$Mr = -V_m e_2 + \alpha_2 M e_2 + f_d - \hat{f}_d + \bar{h} + \tau_d + u - \mu. \quad (63)$$

Taking the time derivative of (63) and using (58) and (61) yields

$$\begin{aligned} M\dot{r} = & -\dot{M}r - \dot{V}_m e_2 - V_m\dot{e}_2 + \alpha_2\dot{M}e_2 + \alpha_2 M\dot{e}_2 \\ & + W^T\sigma'V^T\dot{x}_d - \dot{\hat{W}}^T\hat{\sigma} - \hat{W}^T\hat{\sigma}'\dot{\hat{V}}^T x_d \\ & - \hat{W}^T\hat{\sigma}'\hat{V}^T\dot{x}_d + \dot{\varepsilon} + \dot{\bar{h}} + \dot{\tau}_d + \dot{u} - \dot{\mu}, \end{aligned} \quad (64)$$

where the notations $\hat{\sigma}$ and $\tilde{\sigma}$ are introduced in (54). Adding and subtracting the terms $W^T\hat{\sigma}'\hat{V}^T\dot{x}_d + \hat{W}^T\hat{\sigma}'\tilde{V}^T\dot{x}_d$ to (64), yields

$$\begin{aligned} M\dot{r} = & -\dot{M}r - \dot{V}_m e_2 - V_m\dot{e}_2 + \alpha_2\dot{M}e_2 + \alpha_2 M\dot{e}_2 \\ & + \hat{W}^T\hat{\sigma}'\tilde{V}^T\dot{x}_d + \tilde{W}^T\hat{\sigma}'\hat{V}^T\dot{x}_d - \dot{\hat{W}}^T\hat{\sigma} \\ & - \hat{W}^T\hat{\sigma}'\dot{\hat{V}}^T x_d + W^T\sigma'V^T\dot{x}_d - W^T\hat{\sigma}'\hat{V}^T\dot{x}_d \\ & - \hat{W}^T\hat{\sigma}'\tilde{V}^T\dot{x}_d + \dot{\varepsilon} + \dot{\bar{h}} + \dot{\tau}_d + \dot{u} - \dot{\mu}. \end{aligned} \quad (65)$$

By using (21) and the NN weight tuning laws in (62), the expression in (65) can be rewritten as

$$\begin{aligned} M\dot{r} = & -\frac{1}{2}\dot{M}(q)r + \tilde{N} + N - e_2 - R^{-1}r - (k_s+1)r \\ & - \beta_1 sgn(e_2), \end{aligned} \quad (66)$$

where the unmeasurable auxiliary terms $\tilde{N}(e_1, e_2, r, t)$, $N(\hat{W}, \hat{V}, x_d, t) \in \mathbb{R}^n$ are defined as

$$\begin{aligned} \tilde{N} \triangleq & -\frac{1}{2}\dot{M}r + \dot{\bar{h}} + e_2 + \alpha_2 R^{-1}e_2 - \dot{V}_m e_2 - V_m\dot{e}_2 \\ & + \alpha_2\dot{M}e_2 + \alpha_2 M\dot{e}_2 - proj(\Gamma_1\hat{\sigma}'\hat{V}^T\dot{x}_d e_2^T)^T\hat{\sigma} \\ & - \hat{W}^T\hat{\sigma}' proj(\Gamma_2\dot{x}_d(\hat{\sigma}'^T\hat{W}e_2)^T)^T x_d \end{aligned} \quad (67)$$

$$N \triangleq N_D + N_B. \quad (68)$$

In (68), $N_D(t) \in \mathbb{R}^n$ is defined as

$$N_D = W^T\sigma'V^T\dot{x}_d + \dot{\varepsilon} + \dot{\tau}_d, \quad (69)$$

while $N_B(\hat{W}, \hat{V}, x_d) \in \mathbb{R}^n$ is further segregated as

$$N_B = N_{B_1} + N_{B_2}, \quad (70)$$

where $N_{B_1}(\hat{W}, \hat{V}, x_d) \in \mathbb{R}^n$ is defined as

$$N_{B_1} = -W^T\hat{\sigma}'\hat{V}^T\dot{x}_d - \hat{W}^T\hat{\sigma}'\tilde{V}^T\dot{x}_d, \quad (71)$$

and the term $N_{B_2}(\hat{W}, \hat{V}, x_d) \in \mathbb{R}^n$ is defined as

$$N_{B_2} = \hat{W}^T\hat{\sigma}'\tilde{V}^T\dot{x}_d + \tilde{W}^T\hat{\sigma}'\hat{V}^T\dot{x}_d. \quad (72)$$

Segregating the terms as in (69)–(72) facilitates the development of the NN weight update laws and the subsequent stability analysis. For example, the terms in (69) are grouped together because the terms and their time derivatives can be upper bounded by a constant and rejected by the RISE feedback, whereas the terms grouped in (70) can be upper bounded by a constant but their derivatives are state dependent. The terms in (70) are further segregated because $N_{B_1}(\hat{W}, \hat{V}, x_d)$ will be rejected by the RISE feedback, whereas $N_{B_2}(\hat{W}, \hat{V}, x_d)$ will be partially rejected by the RISE feedback and partially canceled by the adaptive update law for the NN weight estimates.

In a similar manner as in Xian et al. (2004), the Mean Value Theorem can be used to develop the following upper bound

$$\|\tilde{N}(t)\| \le \rho(\|y\|)\|y\|, \quad (73)$$

where $y(t) \in \mathbb{R}^{3n}$ was defined in (31), and the bounding function $\rho(\|y\|) \in \mathbb{R}$ is a positive globally invertible nondecreasing function. The following inequalities can be developed based on Assumption 5, (55), (56), (59), (62) and (70)–(72):

$$\|N_D\| \le \zeta_1 \qquad \|N_B\| \le \zeta_2 \qquad \|\dot{N}_D\| \le \zeta_3 \quad (74)$$

$$\|\dot{N}_B\| \le \zeta_4 + \zeta_5\|e_2\|, \quad (75)$$

where $\zeta_i \in \mathbb{R}$ $(i = 1, 2, \ldots, 5)$ are known positive constants. To facilitate the subsequent stability analysis, the control gains α_1 and α_2 introduced in (5) and (6), respectively, are selected according to the sufficient conditions:

$$\lambda_{\min}(\alpha_1) > \frac{1}{2} \qquad \alpha_2 > \zeta_5 + 1, \quad (76)$$

and β_1 is selected according to the following sufficient conditions:

$$\beta_1 > \zeta_1 + \zeta_2 + \frac{1}{\alpha_2}\zeta_3 + \frac{1}{\alpha_2}\zeta_4, \quad (77)$$

where $\zeta_i \in \mathbb{R}$, $i = 1, 2, \ldots, 5$ are introduced in (74) and (75), and β_1 was introduced in (32).

7.3. Stability and optimality analysis

Theorem 2. *The nonlinear optimal controller given in (60)–(62) ensures that all system signals are bounded under closed-loop operation and that the position tracking error is regulated in the sense that*

$$\|y(t)\| \to 0 \quad \text{as } t \to \infty \quad \forall y(0) \in \mathcal{S} \tag{78}$$

where the set $\mathcal{S}$ can be made arbitrarily large by selecting k_s based on the initial conditions of the system (i.e., a semi-global result). The boundedness of the closed loop signals and the result in (78) can be obtained provided the sufficient conditions in (76) and (77) are satisfied. Furthermore, $u(t)$ converges to an optimal controller that minimizes (12) subject to (11) provided the gain conditions given in (16)–(18) are satisfied.

Proof. Let $\mathcal{D} \subset \mathbb{R}^{3n+2}$ be a domain containing $\Phi(t) = 0$, where $\Phi(t) \in \mathbb{R}^{3n+2}$ is defined as

$$\Phi(t) \triangleq [y^T(t)\ \sqrt{P(t)}\ \sqrt{G(t)}]^T. \tag{79}$$

In (79), the auxiliary function $P(t) \in \mathbb{R}$ is defined as

$$P(t) \triangleq \beta_1 \sum_{i=1}^{n} |e_{2_i}(0)| - e_2(0)^T N(0) - L(t), \tag{80}$$

where $e_{2_i}(0)$ is equal to the ith element of $e_2(0)$ and the auxiliary function $L(t) \in \mathbb{R}$ is the generalized solution to

$$\dot{L}(t) \triangleq r^T(N_{B_1}(t) + N_D(t) - \beta_1 sgn(e_2)) + \dot{e}_2^T(t)N_{B_2}(t) - \zeta_5\|e_2(t)\|^2. \tag{81}$$

Provided the sufficient conditions introduced in (77) are satisfied (Xian et al., 2004)

$$L(t) \leq \beta_1 \sum_{i=1}^{n} |e_{2_i}(0)| - e_2(0)^T N_B(0). \tag{82}$$

Hence, (82) can be used to conclude that $P(t) \geq 0$. The auxiliary function $G(t) \in \mathbb{R}$ in (79) is defined as

$$G(t) = \frac{\alpha_2}{2} tr(\tilde{W}^T \Gamma_1^{-1} \tilde{W}) + \frac{\alpha_2}{2} tr(\tilde{V}^T \Gamma_2^{-1} \tilde{V}). \tag{83}$$

Since Γ_1 and Γ_2 are constant, symmetric, and positive definite matrices and $\alpha_2 > 0$, it is straightforward that $G(t) \geq 0$.

Let $V_L(\Phi, t) : \mathcal{D} \times [0, \infty) \to \mathbb{R}$ be a continuously differentiable positive definite function defined as

$$V_L(\Phi, t) \triangleq e_1^T e_1 + \frac{1}{2} e_2^T e_2 + \frac{1}{2} r^T M(q) r + P + G, \tag{84}$$

provided the sufficient conditions introduced in (77) are satisfied. After taking the time derivative of (84), utilizing (5), (6) and (66), and substituting for the time derivative of $P(t)$ and $G(t)$, $\dot{V}_L(\Phi, t)$ can be simplified as

$$\begin{aligned}\dot{V}_L(\Phi, t) = {} & -2e_1^T\alpha_1 e_1 - (k_s + 1)\|r\|^2 - r^T R^{-1} r 2 e_2^T e_1 \\ & + r^T \tilde{N}(t) - \alpha_2\|e_2\|^2 + \zeta_5\|e_2(t)\|^2 \\ & + \alpha_2 e_2^T[\hat{W}^T \hat{\sigma}' \tilde{V}^T \dot{x}_d + \tilde{W}^T \hat{\sigma}' \hat{V}^T \dot{x}_d] \\ & + tr(\alpha_2 \tilde{W}^T \Gamma_1^{-1} \dot{\tilde{W}}) + tr(\alpha_2 \tilde{V}^T \Gamma_2^{-1} \dot{\tilde{V}}).\end{aligned} \tag{85}$$

Based on (44) and (62), the expression in (85) can be simplified as

$$\begin{aligned}\dot{V}_L(\Phi, t) \leq {} & r^T \tilde{N}(t) - (k_s + 1 + \lambda_{\min}(R^{-1}))\|r\|^2 \\ & - (2\lambda_{\min}(\alpha_1) - 1)\|e_1\|^2 - (\alpha_2 - 1 - \zeta_5)\|e_2\|^2.\end{aligned} \tag{86}$$

By using (73), the expression in (86) can be rewritten as

$$\dot{V}_L(\Phi, t) \leq -\lambda_3\|y\|^2 - [k_s\|r\|^2 - \rho(\|y\|)\|r\|\|y\|], \tag{87}$$

where $\lambda_3 \triangleq \min\{2\lambda_{\min}(\alpha_1) - 1, \alpha_2 - 1 - \zeta_5, 1 + \lambda_{\min}(R^{-1})\}$; hence, α_1, and α_2 must be chosen according to the sufficient condition in (76). After completing the squares for the terms inside the brackets in (87), the following expression can be obtained:

$$\dot{V}_L(\Phi, t) \leq -\lambda_3\|y\|^2 + \frac{\rho^2(\|y\|)\|y\|^2}{4k_s}. \tag{88}$$

Based on (84) and (88), the same stability arguments from Theorem 1 can be used to conclude the result in (78). Furthermore, the result in (78) can be used along with (63) to indicate that

$$\hat{f}_d + \mu = h + \tau_d \quad \text{as } t \to \infty. \tag{89}$$

Using (89), and comparing (9) to (60) indicates that the dynamics in (7) converge to the state-space system in (11). Hence, $u(t)$ converges to an optimal controller that minimizes (12) subject to (11) provided the gain conditions given in (16)–(18), (76) and (77) are satisfied.

8. Experimental results

To examine the performance of the controllers developed in (23) and (60) an experiment was performed on a two-link robot testbed. The testbed is composed of a two-link direct drive revolute robot consisting of two aluminum links, mounted on a 240.0 [Nm] (base joint) and 20.0 [Nm] (second joint) switched reluctance motors. The control objective is to track the desired time-varying trajectory

$$q_{d_1} = q_{d_2} = 60\sin(2t)(1 - \exp(-0.01t^3)). \tag{90}$$

To achieve the control objective, the control gains α_2, k_s, and β_1 defined as scalars in (6) and (32), were implemented (with non-consequential implications to the stability result) as diagonal gain matrices. The weighting matrixes for the controllers were chosen as

$$Q_{11} = \begin{bmatrix} 40 & 2 \\ 2 & 40 \end{bmatrix}, \qquad Q_{12} = \begin{bmatrix} -4 & 4 \\ 4 & -6 \end{bmatrix}$$

$$Q_{22} = diag\left\{4, \quad 4\right\},$$

which using (16)–(18) yielded the following values for K, α_1, and R

$$K = \begin{bmatrix} 4 & -4 \\ -4 & 6 \end{bmatrix}, \qquad \alpha_1 = \begin{bmatrix} 15.6 & 10.6 \\ 10.6 & 10.4 \end{bmatrix}$$

$$R = diag\left\{0.25, \quad 0.25\right\}.$$

The remaining control gains for both controllers were selected as

$$\alpha_2 = diag\left\{60, \quad 35\right\}, \qquad \beta_1 = diag\left\{5, \quad 0.1\right\}$$

$$k_s = diag\left\{140, \quad 20\right\}.$$

The neural network update law weights were selected as

$$\Gamma_1 = 25I_{11} \qquad \Gamma_2 = 25I_7.$$

For all experiments, the rotor's velocity signal is obtained by applying a standard backwards difference algorithm to the position signal. The integral structure for the RISE term in (32) was computed on-line via a standard trapezoidal algorithm. In addition, all the states were initialized to zero. Each experiment was performed ten times, and data from the experiments is displayed in Table 1. Figs. 1 and 2 depict the tracking errors and control torques for one experimental trial for the optimal RISE controller. Figs. 3 and 4 depict the tracking errors and control torques for one experimental trial for the optimal NN+RISE controller.

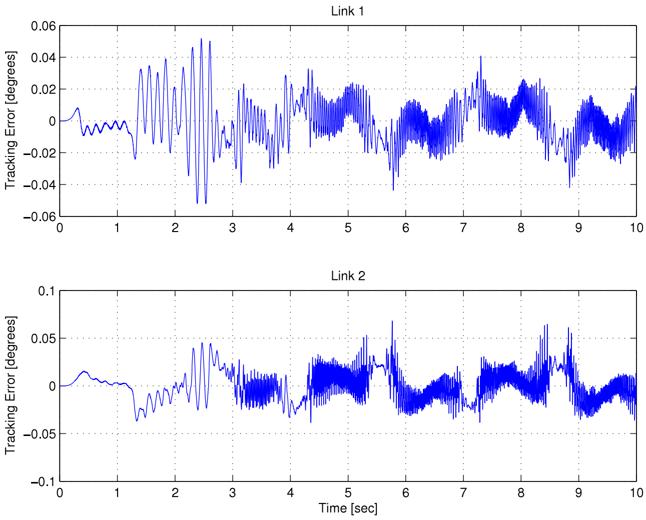

Fig. 1. Tracking errors for the optimal RISE controller.

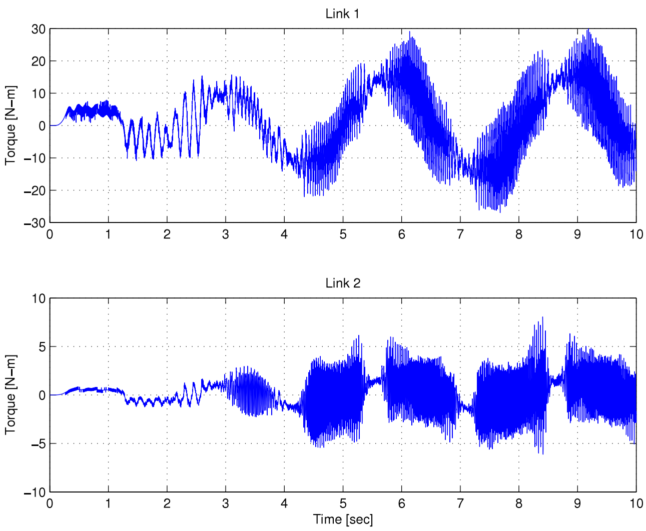

Fig. 2. Torques for the optimal RISE controller.

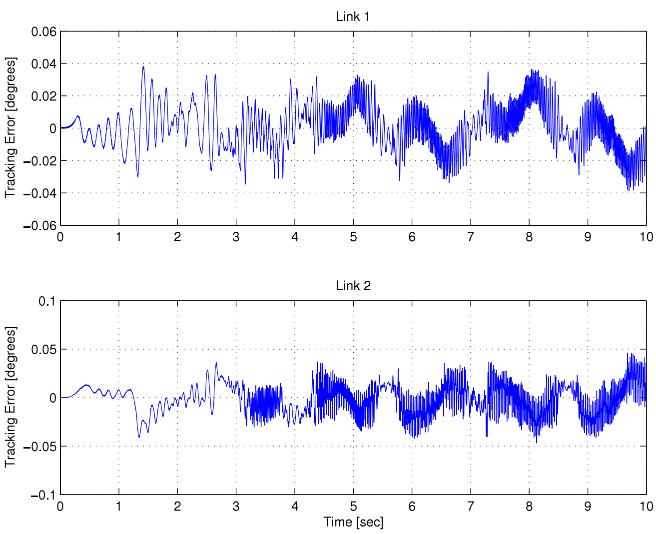

Fig. 3. Tracking errors for the optimal NN+RISE controller.

Table 1
Tabulated values for the 10 runs for the developed controllers.

	RISE	NN+RISE
Avg. Max SS Error (deg)– Link 1	0.0416	0.0416
Avg. Max SS Error (deg)– Link 2	0.0573	0.0550
Avg. RMS Error (deg) – Link 1	0.0128	0.0139
Avg. RMS Error (deg) – Link 2	0.0139	0.0143
Avg. RMS Torque (Nm) – Link 1	9.4217	9.4000
Avg. RMS Torque (Nm) – Link 2	1.7375	1.6825
Error std dev (deg) – Link 1	0.0016	0.0011
Error std dev (deg) – Link 2	0.0019	0.0015
Torque std dev (Nm) – Link 1	0.2775	0.3092
Torque std dev (Nm) – Link 2	0.0734	0.1746

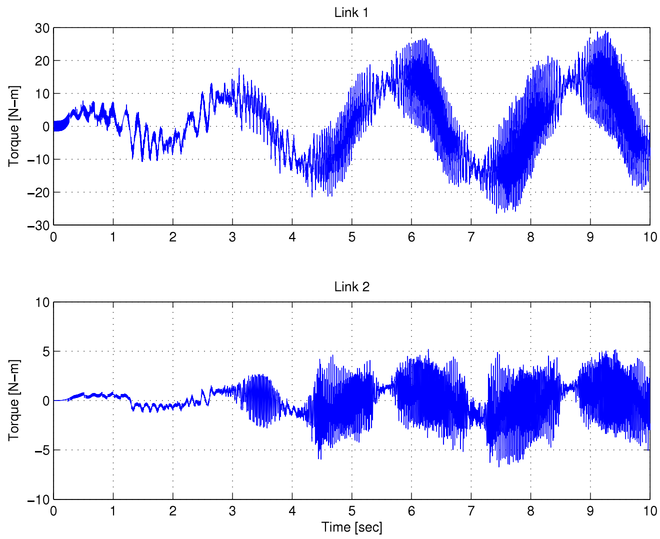

Fig. 4. Torques for the optimal NN+RISE controller.

8.1. Discussion

The experiments show that both controllers stabilize the system. Both controllers keep the average maximum steady state (defined as the last 5 s of the experiment) error under 0.05 degrees for the first link and under 0.06 degrees for the second link. The data in Table 1 indicates that the NN+RISE controller resulted in slightly more RMS error for each link, although a reduced or equal maximum steady state error, with a slightly reduced torque. The reduced standard deviation of the NN+RISE controller show that the results from each run were more alike then the RISE controller alone, but there was greater variance in the torque.

9. Conclusion

A control scheme is developed for a class of nonlinear Euler–Lagrange systems that enables the generalized coordinates to asymptotically track a desired time-varying trajectory despite general uncertainty in the dynamics such as additive bounded disturbances and parametric uncertainty that do not have to satisfy a LP assumption. The main contribution of this work is that the RISE feedback method is augmented with and without a feedforward NN and an auxiliary control term that minimizes a quadratic performance index based on a HJB optimization scheme. Like the influential work in Abu-Khalaf and Lewis (2002), Abu-Khalaf et al. (2006), Cheng et al. (2006), Cheng and Lewis (2007), Cheng et al. (2007), Johansson (1990), Kim and Lewis (2000), Kim et al. (2000) and Lewis (1986) the result in this effort initially develops an optimal controller based on a partially feedback linearized state-space model assuming exact knowledge of the dynamics. The optimal controller is then combined with a feedforward NN and

RISE feedback. A Lyapunov stability analysis is included to show that the NN and RISE identify the uncertainties, therefore the dynamics asymptotically converge to the state-space system that the HJB optimization scheme is based on. Experiments show that both controllers stabilize the system and the optimal NN+RISE controller yields similar steady state error compared to the optimal RISE controller while requiring slightly reduced torque.

References

Abu-Khalaf, M., Huang, J., & Lewis, F. L. (2006). *Nonlinear H-2/H-infinity constrained feedback control*. Springer.

Abu-Khalaf, M., & Lewis, F. L. (2002) Nearly optimal HJB solution for constrained input systems using a neural network least-squares approach. In *Proc. IEEE conf. decis. control* (pp. 943–948), Las Vegas, NV, 2002.

Cheng, J., Li, H., & Zhang, Y. (2006) Robust low-cost sliding mode overload control for uncertain agile missile model. In *Proc. world congr. intell. control autom.* (pp. 2185–2188), Dalian, China, June 2006.

Cheng, T., & Lewis, F. (2007) Neural network solution for finite-horizon H-infinity constrained optimal control of nonlinear systems. In *Proc. IEEE int. conf. control autom.* (pp. 1966–1972), June 2007.

Cheng, T., Lewis, F. L., & Abu-Khalaf, M. (2007). A neural network solution for fixed-final time optimal control of nonlinear systems. *Automatica*, *43*, 482–490.

Dixon, W. E., Fang, Y., Dawson, D. M., & Flynn, T. J. (2003). Range identification for perspective vision systems. *IEEE Transactions on Automatic Control*, *48*(12), 2232–2238.

Dupree, K., Patre, P. M., Wilcox, Z. D., & Dixon, W. E. (2008) Optimal control of uncertain nonlinear systems using RISE feedback. In *Proc. IEEE conf. decis. control*. (pp. 2154–2159), December 2008.

Fausz, J. L., Chellaboina, V. -S., & Haddad, W. M. (1997) Inverse optimal adaptive control for nonlinear uncertain systems with exogenous disturbances. In *Proc. IEEE Conf. Decis. Control* (pp. 2654–2659), December 1997.

Freeman, R. A., & Kokotovi'c, P. V. (1996). *Robust nonlinear control design: state-space and lyapunov techniques*. Boston, MA: BirkÍauser.

Freeman, R. A., & Kokotovic, P. V. (1995) Optimal nonlinear controllers for feedback linearizable systems. In *Proc. IEEE am. control conf.* (pp. 2722–2726), June 1995.

Ge, S. S., Hang, C. C., & Zhang, T. (1999). Nonlinear adaptive control using neural networks and its application to CSTR systems. *Journal of Process Control*, *9*(4), 313–323.

Johansson, R. (1990). Quadratic optimization of motion coordination and control. *IEEE Transactions on Automatic Control*, *35*(11), 1197–1208.

Khalil, H. K. (2002). *Nonlinear systems* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Kim, Y. H., & Lewis, F. L. (2000). Optimal design of CMAC neural-network controller for robot manipulators. *IEEE Transactions on Systems, Man and Cybernetics Part C Applications and Review*, *30*(1), 22–31.

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

Kirk, D. (2004). *Optimal control theory: an introduction*. Dover Pubns.

Krstic, M., & Li, Z.-H. (1998). Inverse optimal design of input-to-state stabilizing nonlinear controllers. *IEEE Transactions on Automatic Control*, *43*(3), 336–350.

Krstic, M., & Tsiotras, P. (1999). Inverse optimal stabilization of a rigid spacecraft. *IEEE Transactions on Automatic Control*, *44*(5), 1042–1049.

Lewis, F. L. (1986). *Optimal control*. John Wiley & Sons.

Lewis, F. L. (1999). Nonlinear network structures for feedback control. *Asian Journal of Control*, *1*(4), 205–228.

Lewis, F. L., Selmic, R., & Campos, J. (2002). *Neuro-fuzzy control of industrial systems with actuator nonlinearities*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.

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

Li, Z. H., & Krstic, M. (1997). Optimal design of adaptive tracking controllers for nonlinear systems. *Automatica*, *33*, 1459–1473.

Lu, Q., Sun, Y., Xu, Z., & Mochizuki, T. (1996). Decentralized nonlinear optimal excitation control. *IEEE Transactions on Power Systems*, *11*(4), 1957–1962.

Luo, W., Chu, Y.-C., & Ling, K.-V. (2005). Inverse optimal adaptive control for attitude tracking of spacecraft. *IEEE Transactions on Automatic Control*, *50*(11), 1639–1654.

Nevistic, V., & Primbs, J.A. (1996) Constrained nonlinear optimal control: a converse HJB approach. Technical report CIT-CDS 96–021, California Institute of Technology, Pasadena, CA 91125.

Patre, P. M., MacKunis, W., Kaiser, K., & Dixon, W. E. (2008). Asymptotic tracking for uncertain dynamic systems via a multilayer neural network feedforward and RISE feedback control structure. *IEEE Transactions on Automatic Control*, *53*(9), 2180–2185.

Patre, P. M., MacKunis, W., Makkar, C., & Dixon, W. E. (2008). Asymptotic tracking for systems with structured and unstructured uncertainties. *IEEE Transactions on Control Systems Technology*, *16*(2), 373–379.

Primbs, J.A., & Nevistic, V. (1996) Optimality of nonlinear design techniques: A converse HJB approach. Technical report CIT-CDS 96–022, California Institute of Technology, Pasadena, CA 91125.

Qu, Z., & Xu, J. X. (2002). Model-based learning controls and their comparisons using lyapunov direct method. *Asian Journal of Control*, *4*(1), 99–110.

Sekoguchi, M., Konishi, H., Goto, M., Yokoyama, A., & Lu, Q. (2002) Nonlinear optimal control applied to STATCOM for power system stabilization In *Proc. IEEE/PES Transm. Distrib. Conf. Exhib.* (pp. 342–347), October 2002.

Xian, B., Dawson, D. M., de Queiroz, M. S., & Chen, J. (2004). A continuous asymptotic tracking control strategy for uncertain nonlinear systems. *IEEE Transactions on Automatic Control*, *49*(7), 1206–1211.

Keith Dupree received his B.S. Degree in Aerospace Engineering with a minor in Physics in May of 2005, his M.S. in Mechanical Engineering in May of 2007, his Masters of Science in Electrical and Computer Engineering in August of 2008, and his Ph.D. in Mechanical Engineering; all from the University of Florida. He is currently employed as a Systems Engineer by Raytheon. His research interests include Nonlinear Controls, Image Processing, State Estimation, and Optimal Control.

Parag M. Patre received his B. Tech. degree in mechanical engineering from the Indian Institute of Technology Madras, India, in 2004. Following this he was with Larsen & Toubro Limited, India until 2005, when he joined the graduate school at the University of Florida. He received his M.S. and Ph.D. degrees in mechanical engineering in 2007 and 2009, respectively, and is currently a NASA postdoctoral fellow at the NASA Langley Research Center, Hampton, VA. His areas of research interest are Lyapunov-based design and analysis of control methods for uncertain nonlinear systems, robust and adaptive control, neural networks, and control of robots and aircraft.

Zachary D. Wilcox received his Ph.D. in 2010 from the Department of Mechanical and Aerospace Engineering at the University of Florida (UF). After completing his doctoral studies, he was selected as a Post-doctoral Research Associate at NASA Dryden, Edwards Air Force Base. Dr. Wilcox's main research interest has been the development and application of Lyapunov-based control techniques for aerospace systems with uncertain dynamic models.

Warren E. Dixon received his Ph.D. in 2000 from the Department of Electrical and Computer Engineering from Clemson University. After completing his doctoral studies he was selected as an Eugene P. Wigner Fellow at Oak Ridge National Laboratory (ORNL) where he worked in the Robotics and Energetic Systems Group. In 2004, Dr. Dixon joined the faculty of the University of Florida in the Mechanical and Aerospace Engineering Department. His main research interest is the development and application of Lyapunov-based control techniques for uncertain nonlinear systems. He is the co-author of four monographs, an edited collection, six chapters, and over 200 refereed journal and conference papers. His work has been recognized by the 2009 American Automatic Control Council (AACC) O. Hugo Schuck Award, 2006 IEEE Robotics and Automation Society (RAS) Early Academic Career Award, an NSF CAREER Award (2006–2011), 2004 DOE Outstanding Mentor Award, and the 2001 ORNL Early Career Award for Engineering Achievement. Dr. Dixon is a senior member of IEEE. He is a member of various technical committees, and conference program and organizing committees. He also serves on the conference editorial board for the IEEE CSS and RAS and the ASME DSC. He served as an appointed member to the IEEE CSS Board of Governors for 2008. He is currently an associate editor for *Automatica, IEEE Transactions on Systems Man and Cybernetics: Part B Cybernetics,International Journal of Robust and Nonlinear Control, and Journal of Robotics.*