2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Optimal Control of Uncertain Nonlinear Systems using a Neural Network and RISE Feedback

K. Dupree, P. M. Patre, Z. D. Wilcox, and W. E. Dixon
email: {kdupree, pmpatre, zibrus, wdixon}@ufl.edu
Department of Mechanical and Aerospace Engineering
University of Florida, Gainesville, FL 32611

***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 systems is challenging. 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 current effort builds on our previous efforts to illustrate how a NN can be combined with a recent robust feedback method to asymptotically minimize a given quadratic performance index as the generalized coordinates of a nonlinear Euler-Lagrange system asymptotically track a desired time-varying trajectory despite general uncertainty in the dynamics. A Lyapunov analysis is provided to examine the stability of the developed optimal controller.**

I. INTRODUCTION[1]

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 can not be solved explicitly. If the nonlinear dynamics are exactly known, then the problem can be reduced to solving an ARE through feedback linearization methods (cf. [1]–[5]).

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), [6] 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 [7], 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 [7] exploits the implicit learning characteristic [8] of the RISE controller to asymptotically cancel LP and non-LP uncertain dynamics so that the overall control structure converges to an optimal controller.

[1]The authors would like to gratefully acknowledge the support of the Department of Energy, grant number DE-FG04-86NE37967. This work was accomplished as part of the DOE University Research Program in Robotics (URPR).

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 [9]–[14] 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 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 efforts in this paper investigate the amalgam of the robust RISE feedback method with NN methods to yield a direct optimal controller. The utility of combining these feedforward and feedback methods are twofold. Our previous efforts in [15] indicate that modifying the RISE feedback with a feedforward term can reduce the control effort and improve the transient and steady state response of the RISE controller. Hence, the combined results should converge to the optimal controller faster. Moreover, combining NN feedforward controllers with RISE feedback yields asymptotic results [16]. Hence, the efforts in this paper provide a modification to the results in [9]–[14] that allows for asymptotic stability and convergence to the optimal controller rather than to approximate the optimal controller.

As is typical with previous nonlinear direct optimal controllers the unknown LP and non-LP dynamics are temporarily assumed to be known so that a controller can be developed for a residual system based on the HJB optimization method for a given quadratic performance index. The original uncertain nonlinear system is then examined, where the optimal controller is augmented to include the RISE feedback and NN feedforward terms to asymptotically cancel the uncertainties. A Lyapunov-based stability analysis is included to show that the RISE and NN components asymptotically identify the unknown dynamics (yielding semi-global asymptotic tracking) provided upper bounds on the disturbances are known and the control gains are selected appropriately. Moreover, the controller converges to the optimal controller for the a priori given quadratic performance index.

II. DYNAMIC MODEL AND PROPERTIES

The class of nonlinear dynamic systems considered in this paper is assumed to be modeled by the following Euler-

Lagrange [17] formulation:

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

In (1), $M(q) \in \mathbb{R}^{n\times n}$ denotes the inertia matrix, $V_m(q,\dot{q}) \in \mathbb{R}^{n\times n}$ denotes the centripetal-Coriolis matrix, $G(q) \in \mathbb{R}^n$ denotes the gravity vector, $F(\dot{q}) \in \mathbb{R}^n$ denotes friction, $\tau_d(t) \in \mathbb{R}^n$ denotes a general nonlinear disturbance (e.g., unmodeled effects), $\tau(t) \in \mathbb{R}^n$ represents the input, and $q(t)$, $\dot{q}(t)$, $\ddot{q}(t) \in \mathbb{R}^n$ denote the 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 properties and assumptions will be exploited in the subsequent development.

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

$$m_1 \|y\|^2 \leq y^T M(q) y \leq \bar{m}(q) \|y\|^2, \quad (2)$$

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.

Property 2: The following skew-symmetric relationship is satisfied:

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

Property 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.

Property 4: 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.

Property 5: 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.

III. 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. To quantify this objective, a position tracking error, denoted by $e_1(t) \in \mathbb{R}^n$, is defined as

$$e_1 \triangleq q_d - q. \quad (4)$$

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

$$e_2 \triangleq \dot{e}_1 + \alpha_1 e_1 \quad (5)$$

$$r \triangleq \dot{e}_2 + \alpha_2 e_2, \quad (6)$$

where $\alpha_1 \in \mathbb{R}^{n\times n}$, denotes a positive, constant, gain matrix, and $\alpha_2 \in \mathbb{R}$ is a positive constant. The filtered tracking error $r(t)$ is not measurable since the expression in (6) depends on $\ddot{q}(t)$.

IV. Optimal Computed Controller 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. This development motivates the control design in Section V, where a NN and a robust controller are developed to identify the unknown dynamics and additive disturbance.

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

$$M(q)\dot{e}_2 = -V_m e_2 - \tau + h + \tau_d, \quad (7)$$

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

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

Under the (temporary) assumption that the dynamics in (1) are known, the control input can be designed as [7]

$$\tau \triangleq h + \tau_d - u, \quad (9)$$

to yield the state-space model

$$\dot{z} = A(q,\dot{q})z + B(q)u, \quad (10)$$

where $u(t) \in \mathbb{R}^n$ is an auxiliary control input that will be designed to minimize a subsequent performance index, and $A(q,\dot{q}) \in \mathbb{R}^{2n\times 2n}$, $B(q) \in \mathbb{R}^{2n\times n}$, and $z(t) \in \mathbb{R}^{2n}$ are defined as

$$\begin{aligned} A(q,\dot{q}) &\triangleq \begin{bmatrix} -\alpha_1 & I_{n\times n} \\ 0_{n\times n} & -M^{-1}V_m \end{bmatrix}, \\ B(q) &\triangleq \begin{bmatrix} 0_{n\times n} & M^{-1} \end{bmatrix}^T, \\ z(t) &\triangleq \begin{bmatrix} e_1 & e_2 \end{bmatrix}^T, \end{aligned}$$

where $I_{n\times n}$ and $0_{n\times n}$ denote a $n \times n$ identity matrix and matrix of zeros, respectively. The quadratic performance index $J(u) \in \mathbb{R}$ to be minimized subject to the constraints in (10) is

$$J(u) \triangleq \int_0^\infty \frac{1}{2} z^T Q z + \frac{1}{2} u^T R u \quad dt. \quad (11)$$

In (11), $Q \in \mathbb{R}^{2n\times 2n}$ 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. As stated in [9], [10], 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) can not be modified by the optimal design phase.

To facilitate the subsequent development, let $\Omega(q) \in \mathbb{R}^{2n\times 2n}$ be defined as

$$\Omega(q) = \begin{bmatrix} K & 0_{n\times n} \\ 0_{n\times n} & M \end{bmatrix} \quad (12)$$

where $K \in \mathbb{R}^{n \times n}$ denotes a gain matrix. If α_1, R, and K, introduced in (5), (11), and (12), satisfy the algebraic relationships

$$K = K^T = -\frac{1}{2}\left(Q_{12} + Q_{12}^T\right) > 0 \quad (13)$$

$$Q_{11} = \alpha_1^T K + K\alpha_1, \quad (14)$$

$$R^{-1} = Q_{22}, \quad (15)$$

where $Q_{ij} \in \mathbb{R}^{n \times n}$ denotes a block of Q, then Theorem 1 of [9] and [10] can be invoked to prove that $\Omega(q)$ satisfies the Riccati differential equation, and the value function $V_a(z,t) \in \mathbb{R}$

$$V_a = \frac{1}{2} z^T \Omega z$$

satisfies the HJB equation. Lemma 1 of [9] and [10] can be used to conclude that the optimal control $u(t)$ that minimizes (11) subject to (10) is

$$u(t) = -R^{-1} B^T \left(\frac{\partial V_a(z,t)}{\partial z}\right)^T = -R^{-1} e_2. \quad (16)$$

V. Control Development

In general, the bounded disturbance, so the controller given in (9) can not 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 (10) so that $u(t)$ minimizes the respective performance index. In this section, a controller is developed that exploits the universal approximation property of NNs and the implicit learning characteristics of the RISE feedback to identify the nonlinear effects and bounded disturbances to enable $z(t)$ to asymptotically converge to the state space model.

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 [18], [19]

$$f(x) = W^T \sigma\left(V^T x\right) + \varepsilon(x). \quad (17)$$

In (17), $V \in \mathbb{R}^{(N_1+1)\times N_2}$ 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[2] in (17) 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 (17), the typical three-layer NN approximation for $f(x)$ is given as [18], [19]

$$\hat{f}(x) \triangleq \hat{W}^T \sigma\left(\hat{V}^T x\right), \quad (18)$$

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

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

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

$$\tilde{\sigma} \triangleq \sigma - \hat{\sigma} = \sigma\left(V^T x\right) - \sigma\left(\hat{V}^T x\right). \quad (19)$$

One of the NN estimate properties that facilitate the subsequent development is described as follows.

Property 6: (*Boundedness of the Ideal Weights*) The ideal weights are assumed to exist and be bounded by known positive values so that

$$\|V\|_F^2 = tr(V^T V) \le \bar{V}_B \quad (20)$$

$$\|W\|_F^2 = tr(W^T W) \le \bar{W}_B \quad (21)$$

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

$$Mr = -V_m e_2 + h + \tau_d + \alpha_2 M e_2 - \tau. \quad (22)$$

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

$$f_d \triangleq M(q_d)\ddot{q}_d + V_m(q_d, \dot{q}_d)\dot{q}_d + G(q_d) + F(\dot{q}_d), \quad (23)$$

is added and subtracted to (22) to yield

$$Mr = -V_m e_2 + \bar{h} + f_d + \tau_d + \alpha_2 M e_2 - \tau, \quad (24)$$

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

$$\bar{h} \triangleq h - f_d. \quad (25)$$

The expression in (23) can be represented by a three-layer NN as

$$f_d = W^T \sigma\left(V^T x_d\right) + \varepsilon(x_d). \quad (26)$$

In (26), the input $x_d(t) \in \mathbb{R}^{3n+1}$ is defined as $x_d(t) \triangleq [1 \;\; q_d^T(t) \;\; \dot{q}_d^T(t) \;\; \ddot{q}_d^T(t)]^T$ so that $N_1 = 3n$ where N_1 was introduced in (17). Based on the assumption that the desired trajectory is bounded, the following inequalities hold

$$\|\varepsilon(x_d)\| \le \varepsilon_{b_1} \qquad \|\dot{\varepsilon}(x_d, \dot{x}_d)\| \le \varepsilon_{b_2} \quad (27)$$
$$\|\ddot{\varepsilon}(x_d, \dot{x}_d, \ddot{x}_d)\| \le \varepsilon_{b_3},$$

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

Based on the open-loop error system in (22), the control input is composed of the optimal control developed in (16), a three-layer NN feedforward term, plus the RISE feedback term as

$$\tau \triangleq \hat{f}_d + \mu - u. \quad (28)$$

Specifically, $\mu(t) \in \mathbb{R}^n$ denotes the RISE feedback control term defined as [20]

$$\mu(t) \triangleq (k_s+1)e_2(t) - (k_s+1)e_2(0) + \int_0^t [(k_s+1)\alpha_2 e_2(\sigma) + \beta_1 sgn(e_2(\sigma))]d\sigma, \quad (29)$$

[2] A variety of activation functions (e.g., sigmoid, hyperbolic tangent or radial basis) could be used for the control development.

where k_s, $\beta_1 \in \mathbb{R}$ are positive constant control gains. The feedforward NN component in (28), denoted by $\hat{f}_d(t) \in \mathbb{R}^n$, is generated as

$$\hat{f}_d \triangleq \hat{W}^T \sigma\left(\hat{V}^T x_d\right). \quad (30)$$

The estimates for the NN weights in (30) are generated on-line (there is no off-line learning phase) as

$$\begin{aligned} \dot{\hat{W}} &= proj(\Gamma_1 \hat{\sigma}' \hat{V}^T \dot{x}_d e_2^T) \quad (31)\\ \dot{\hat{V}} &= proj(\Gamma_2 \dot{x}_d (\hat{\sigma}'^T \hat{W} e_2)^T) \end{aligned}$$

where $\sigma'(\hat{V}^T x) \equiv d\sigma\left(V^T x\right)/d\left(V^T x\right)|_{V^T x = \hat{V}^T x}$, and $\Gamma_1 \in \mathbb{R}^{(N_2+1)\times(N_2+1)}$, $\Gamma_2 \in \mathbb{R}^{(3n+1)\times(3n+1)}$ are constant, positive definite, symmetric matrices. In (31), $proj(\cdot)$ 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 [21] for further details.

The closed-loop tracking error system is obtained by substituting (28) into (22) as

$$Mr = -V_m e_2 + \alpha_2 M e_2 + f_d - \hat{f}_d + \bar{h} + \tau_d + u - \mu. \quad (32)$$

To facilitate the subsequent stability analysis, the time derivative of (32) is determined as

$$\begin{aligned} M\dot{r} &= -\dot{M}r - \dot{V}_m e_2 - V_m \dot{e}_2 + \alpha_2 \dot{M} e_2 \quad (33)\\ &\quad + \alpha_2 M \dot{e}_2 + \dot{f}_d - \dot{\hat{f}}_d + \dot{\bar{h}} + \dot{\tau}_d + \dot{u} - \dot{\mu}. \end{aligned}$$

Using (17) and (30), the closed-loop error system in (33) can be expressed as

$$\begin{aligned} M\dot{r} &= -\dot{M}r - \dot{V}_m e_2 - V_m \dot{e}_2 + \alpha_2 \dot{M} e_2 \quad (34)\\ &\quad + \alpha_2 M \dot{e}_2 + W^T \sigma' V^T \dot{x}_d - \dot{\hat{W}}^T \hat{\sigma} \\ &\quad - \hat{W}^T \hat{\sigma}' \dot{\hat{V}}^T x_d - \hat{W}^T \hat{\sigma}' \hat{V}^T \dot{x}_d \\ &\quad + \dot{\varepsilon} + \dot{\bar{h}} + \dot{\tau}_d + \dot{u} - \dot{\mu}, \end{aligned}$$

where the notations $\hat{\sigma}$ and $\tilde{\sigma}$ are introduced in (19). 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 (34), yields

$$\begin{aligned} M\dot{r} &= -\dot{M}r - \dot{V}_m e_2 - V_m \dot{e}_2 + \alpha_2 \dot{M} e_2 \quad (35)\\ &\quad + \alpha_2 M \dot{e}_2 + \hat{W}^T \hat{\sigma}' \tilde{V}^T \dot{x}_d + \tilde{W}^T \hat{\sigma}' \hat{V}^T \dot{x}_d - \dot{\hat{W}}^T \hat{\sigma} \\ &\quad - \hat{W}^T \hat{\sigma}' \dot{\hat{V}}^T x_d + W^T \sigma' V^T \dot{x}_d - W^T \hat{\sigma}' \hat{V}^T \dot{x}_d \\ &\quad - \hat{W}^T \hat{\sigma}' \tilde{V}^T \dot{x}_d + \dot{\varepsilon} + \dot{\bar{h}} + \dot{\tau}_d + \dot{u} - \dot{\mu}. \end{aligned}$$

Using (16) and the NN weight tuning laws in (31), the expression in (35) can be rewritten as

$$\begin{aligned} M\dot{r} &= -\frac{1}{2}\dot{M}(q)r + \tilde{N} + N - e_2 - R^{-1}r \quad (36)\\ &\quad -(k_s+1)r - \beta_1 sgn(e_2), \end{aligned}$$

where the fact that the time derivative of (29) is given as

$$\dot{\mu} = (k_s+1)r + \beta_1 sgn(e_2) \quad (37)$$

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

$$\begin{aligned} \tilde{N} &\triangleq -\frac{1}{2}\dot{M}(q)r + \dot{\bar{h}} + e_2 + \alpha_2 R^{-1} e_2 \quad (38)\\ &\quad - \dot{V}_m e_2 - V_m \dot{e}_2 + \alpha_2 \dot{M} e_2 + \alpha_2 M \dot{e}_2 \\ &\quad - proj(\Gamma_1 \hat{\sigma}' \hat{V}^T \dot{x}_d e_2^T)^T \hat{\sigma} \\ &\quad - \hat{W}^T \hat{\sigma}' proj(\Gamma_2 \dot{x}_d (\hat{\sigma}'^T \hat{W} e_2)^T)^T x_d \\ N &\triangleq N_D + N_B. \quad (39) \end{aligned}$$

In (39), $N_D(t) \in \mathbb{R}^n$ is defined as

$$N_D = W^T \sigma' V^T \dot{x}_d + \dot{\varepsilon} + \dot{\tau}_d, \quad (40)$$

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

$$N_B = N_{B_1} + N_{B_2}, \quad (41)$$

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

$$N_{B_1} = -W^T \hat{\sigma}' \hat{V}^T \dot{x}_d - \hat{W}^T \hat{\sigma}' \tilde{V}^T \dot{x}_d, \quad (42)$$

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

$$N_{B_2} = \hat{W}^T \hat{\sigma}' \tilde{V}^T \dot{x}_d + \tilde{W}^T \hat{\sigma}' \hat{V}^T \dot{x}_d. \quad (43)$$

Segregating the terms as in (40)-(43) facilitates the development of the NN weight update laws and the subsequent stability analysis. For example, the terms in (40) 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 (41) can be upper bounded by a constant but their derivatives are state dependent. The terms in (41) are further segregated because $N_{B_1}\left(\hat{W}, \hat{V}, x_d\right)$ will be rejected by the RISE feedback, whereas $N_{B_2}\left(\hat{W}, \hat{V}, x_d\right)$ 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 [20], the Mean Value Theorem can be used to develop the following upper bound[3]

$$\left\|\tilde{N}(t)\right\| \le \rho\left(\|y\|\right)\|y\|, \quad (44)$$

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

$$y(t) \triangleq [e_1^T \quad e_2^T \quad r^T]^T, \quad (45)$$

and the bounding function $\rho(\|y\|) \in \mathbb{R}$ is a positive globally invertible nondecreasing function. The following inequalities can be developed based on Property 4, (20), (21), (27), (31) and (41)-(43):

$$\|N_D\| \le \zeta_1 \qquad \|N_B\| \le \zeta_2 \qquad \left\|\dot{N}_D\right\| \le \zeta_3 \quad (46)$$

$$\left\|\dot{N}_B\right\| \le \zeta_4 + \zeta_5 \|e_2\|. \quad (47)$$

In (46) and (47) $\zeta_i \in \mathbb{R}$ $(i = 1, 2, ..., 5)$ are known positive constants.

[3]Details of the bound in (44) are available on request.

VI. Stability Analysis

Theorem: The nonlinear optimal controller given in (28)-(31) ensures that all system signals are bounded under closed-loop operation and that the position tracking error is regulated in the sense that

$$\|e_1(t)\| \to 0 \qquad as \qquad t \to \infty. \tag{48}$$

The result in (48) can be achieved provided the control gain k_s introduced in (29) is selected sufficiently large, and α_1, α_2 are selected according to the following sufficient conditions:

$$\lambda_{\min}(\alpha_1) > \frac{1}{2} \qquad \alpha_2 > \beta_2 + 1, \tag{49}$$

where $\lambda_{\min}(\cdot) \in \mathbb{R}$ denotes the minimum eigenvalue, and β_i $(i = 1, 2)$ are selected according to the following sufficient conditions:

$$\beta_1 > \zeta_1 + \zeta_2 + \frac{1}{\alpha_2}\zeta_3 + \frac{1}{\alpha_2}\zeta_4 \qquad \beta_2 > \zeta_5, \tag{50}$$

where $\zeta_i \in \mathbb{R}$, $i = 1, 2,..., 5$ are introduced in (46)-(47), β_1 was introduced in (29), and β_2 is introduced in (53). Furthermore, $u(t)$ converges to an optimal controller that minimizes (11) subject to (10) provided the gain conditions given in (13)-(15) are satisfied.

Remark 1: The control gain α_1 can not 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 (49) is 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

$$\Phi(t) \triangleq [y^T(t) \quad \sqrt{P(t)} \quad \sqrt{G(t)}]^T. \tag{51}$$

In (51), the auxiliary function $P(t) \in \mathbb{R}$ is defined as

$$P(t) \triangleq \beta_1 \sum_{i=1}^{n} |e_{2_i}(0)| - e_2(0)^T N(0) - \int_0^t L(\tau)d\tau, \tag{52}$$

where $e_{2_i}(0)$ is equal to the i^{th} element of $e_2(0)$ and the auxiliary function $L(t) \in \mathbb{R}$ is defined as

$$\begin{aligned} L(t) \quad \triangleq \quad & r^T(N_{B_1}(t) + N_D(t) - \beta_1 sgn(e_2)) \\ & + \dot{e}_2^T(t) N_{B_2}(t) - \beta_2 \|e_2(t)\|^2, \end{aligned} \tag{53}$$

where $\beta_i \in \mathbb{R}$ $(i = 1, 2)$ are positive constants chosen according to the sufficient conditions in (50). Provided the sufficient conditions introduced in (50) are satisfied[4]

$$\int_0^t L(\tau)d\tau \leq \beta_1 \sum_{i=1}^{n} |e_{2_i}(0)| - e_2(0)^T N_B(0). \tag{54}$$

Hence, (54) can be used to conclude that $P(t) \geq 0$. The auxiliary function $G(t) \in \mathbb{R}$ in (51) is defined as

$$G(t) = \frac{\alpha_2}{2} tr\left(\tilde{W}^T \Gamma_1^{-1} \tilde{W}\right) + \frac{\alpha_2}{2} tr\left(\tilde{V}^T \Gamma_2^{-1} \tilde{V}\right) \tag{55}$$

[4]Details of the bound in (54) are available on request.

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

$$V_L(\Phi, t) \triangleq e_1^T e_1 + \frac{1}{2} e_2^T e_2 + \frac{1}{2} r^T M(q) r + P + G, \tag{56}$$

which satisfies the following inequalities:

$$U_1(\Phi) \leq V_L(\Phi, t) \leq U_2(\Phi) \tag{57}$$

provided the sufficient conditions introduced in (50) are satisfied. In (57), 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 λ_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, $\bar{m}(q)$ are introduced in (2). After taking the time derivative of (56), $\dot{V}_L(\Phi, t)$ can be expressed as

$$\dot{V}_L(\Phi, t) = 2e_1^T \dot{e}_1 + e_2^T \dot{e}_2 + \frac{1}{2} r^T \dot{M}(q) r + r^T M(q) \dot{r} + \dot{P} + \dot{G}.$$

By utilizing (5), (6), (36), and substituting in for the time derivative of $P(t)$ and $G(t)$, $\dot{V}(\Phi, t)$ can be simplified as

$$\begin{aligned} \dot{V}_L(\Phi, t) \quad = \quad & -2e_1^T \alpha_1 e_1 - (k_s + 1)\|r\|^2 - r^T R^{-1} r \\ & + 2e_2^T e_1 + r^T \tilde{N}(t) - \alpha_2 \|e_2\|^2 + \beta_2 \|e_2(t)\|^2 \\ & + \alpha_2 e_2^T \left[\hat{W}^T \hat{\sigma}' \tilde{V}^T \dot{x}_d + \tilde{W}^T \hat{\sigma}' \hat{V}^T \dot{x}_d\right] \\ & + tr\left(\alpha_2 \tilde{W}^T \Gamma_1^{-1} \dot{\tilde{W}}\right) + tr\left(\alpha_2 \tilde{V}^T \Gamma_2^{-1} \dot{\tilde{V}}\right). \end{aligned} \tag{58}$$

Based on the fact that

$$e_2^T e_1 \leq \frac{1}{2}\|e_1\|^2 + \frac{1}{2}\|e_2\|^2$$

and using (31), the expression in (58) can be simplified as

$$\begin{aligned} \dot{V}_L(\Phi, t) \quad \leq \quad & r^T \tilde{N}(t) - (k_s + 1 + \lambda_{\min}(R^{-1}))\|r\|^2 \\ & - (2\lambda_{\min}(\alpha_1) - 1)\|e_1\|^2 \\ & - (\alpha_2 - 1 - \beta_2)\|e_2\|^2. \end{aligned} \tag{59}$$

By using (44), the expression in (59) can be rewritten as

$$\dot{V}_L(\Phi, t) \leq -\lambda_3 \|y\|^2 - \left[k_s \|r\|^2 - \rho(\|y\|)\|r\|\|y\|\right], \tag{60}$$

where $\lambda_3 \triangleq \min\{2\lambda_{\min}(\alpha_1) - 1, \alpha_2 - 1 - \beta_2, 1 + \lambda_{\min}(R^{-1})\}$; hence, α_1, and α_2 must be chosen according to the sufficient condition in (49). After completing the squares for the terms inside the brackets in (60), the following expression can be obtained:

$$\dot{V}_L(\Phi, t) \leq -\lambda_3 \|y\|^2 + \frac{\rho^2(\|y\|)\|y\|^2}{4k_s} \leq -U(\Phi), \tag{61}$$

where $U(\Phi) = c\|y\|^2$, for some positive constant c, is a continuous, positive semi-definite function that is defined on the following domain:

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

The inequalities in (57) and (61) 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 $\mathcal{D}$, Property 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 Property 4, we can show that $\tau(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Given that $r(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, (37) can be used to show that $\dot{\mu}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$. Since $\dot{q}(t)$, $\ddot{q}(t) \in \mathcal{L}_\infty$ in $\mathcal{D}$, Property 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, (36) 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 follows:

$$\mathcal{S} \triangleq \left\{ \Phi(t) \subset \mathcal{D} \mid U_2(\Phi(t)) < \lambda_1 \left(\rho^{-1} \left(2\sqrt{\lambda_3 k_s} \right) \right)^2 \right\}. \tag{62}$$

The region of attraction in (62) 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) [20]. Theorem 8.4 of [22] can now be invoked to state that

$$c \left\| y(t) \right\|^2 \to 0 \qquad as \qquad t \to \infty \qquad \forall y(0) \in \mathcal{S}. \tag{63}$$

Based on the definition of $y(t)$, (63) can be used to show that

$$\left\| e_1(t) \right\| \to 0 \qquad as \qquad t \to \infty \qquad \forall y(0) \in \mathcal{S}. \tag{64}$$

The result in (63) indicates that as $t \to \infty$, (32) reduces to

$$\hat{f}_d + \mu = h + \tau_d. \tag{65}$$

Therefore, dynamics in (7) converge to the state-space system in (10). Hence, $u(t)$ converges to an optimal controller that minimizes (11) subject to (10) provided the gain conditions given in (13)-(15), (49), and (50) are satisfied.

VII. 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 a feedforward NN and RISE feedback method is augmented with an auxiliary control term that minimizes a quadratic performance index based on a HJB optimization scheme. Like the influential work in [9]–[14], [23], [24] 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. A preliminary numerical simulation is included to support these results.

References

[1] R. Freeman and P. Kokotovic, “Optimal nonlinear controllers for feedback linearizable systems,” in *Proc. of the American Controls Conf.*, Seattle, WA, 1995, pp. 2722–2726.

[2] Q. Lu, Y. Sun, Z. Xu, and T. Mochizuki, “Decentralized nonlinear optimal excitation control,” *IEEE Trans. Power Syst.*, vol. 11, no. 4, pp. 1957–1962, Nov. 1996.

[3] V. Nevistic and J. A. Primbs, “Constrained nonlinear optimal control: a converse HJB approach,” California Institute of Technology, Pasadena, CA 91125, Tech. Rep. CIT-CDS 96-021, 1996.

[4] J. A. Primbs and V. Nevistic, “Optimality of nonlinear design techniques: A converse hjb approach,” California Institute of Technology, Pasadena, CA 91125, Tech. Rep. CIT-CDS 96-022, 1996.

[5] M. Sekoguchi, H. Konishi, M. Goto, A. Yokoyama, and Q. Lu, “Nonlinear optimal control applied to STATCOM for power system stabilization,” in *Proc. IEEE/PES Transmission and Distribution Conference and Exhibition*, Yokohama, Japan, 2002, pp. 342–347.

[6] R. Johansson, “Quadratic optimization of motion coordination and control,” *IEEE Trans. Autom. Control*, vol. 35, no. 11, pp. 1197–1208, Nov. 1990.

[7] K. Dupree, P. M. Patre, Z. D. Wilcox, and W. E. Dixon, “Optimal control of uncertain nonlinear systems using RISE feedback,” in *Proc. IEEE Conf. on Decision and Control*, Dec. 2008, pp. 2154–2159.

[8] Z. Qu and J. Xu, “Model-based learning controls and their comparisons using Lyapunov direct method,” *Asian Journal of Control*, vol. 4, no. 1, pp. 99–110, Mar. 2002.

[9] Y. Kim, F. Lewis, and D. Dawson, “Intelligent optimal control of robotic manipulator using neural networks,” *Automatica*, vol. 36, pp. 1355–1364, 2000.

[10] Y. Kim and F. Lewis, “Optimal design of CMAC neural-network controller for robot manipulators,” *IEEE Trans. Syst., Man, Cybern. C*, vol. 30, no. 1, pp. 22–31, Feb. 2000.

[11] M. Abu-Khalaf and F. Lewis, “Nearly optimal HJB solution for constrained input systems using a neural network least-squares approach,” in *Proc. IEEE Conf. on Decision and Control*, Las Vegas, NV, 2002, pp. 943–948.

[12] T. Cheng and F. Lewis, “Fixed-final time constrained optimal control of nonlinear systems using neural network HJB approach,” in *Proc. IEEE Conf. on Decision and Control*, San Diego, CA, 2006, pp. 3016–3021.

[13] ——, “Neural network solution for finite-horizon H-infinity constrained optimal control of nonlinear systems,” in *Proc. IEEE Intl. Conf. on Control and Automation*, Guangzhou, China, 2007, pp. 1966–1972.

[14] T. Cheng, F. Lewis, and M. Abu-Khalaf, “A neural network solution for fixed-final time optimal control of nonlinear systems,” *Automatica*, vol. 43, pp. 482–490, 2007.

[15] P. M. Patre, W. MacKunis, C. Makkar, and W. E. Dixon, “Asymptotic tracking for systems with structured and unstructured uncertainties,” *IEEE Trans. Contr. Syst. Technol.*, vol. 16, no. 2, pp. 373–379, 2008.

[16] P. M. Patre, W. MacKunis, K. Kaiser, and W. E. Dixon, “Asymptotic tracking for uncertain dynamic systems via a multilayer neural network feedforward and rise feedback control structure,” *IEEE Trans. Autom. Control*, 2008, to appear.

[17] R. Ortega, A. Loría, P. J. Nicklasson, and H. J. Sira-Ramirez, *Passivity-based Control of Euler-Lagrange Systems: Mechanical, Electrical and Electromechanical Applications*. Springer, 1998.

[18] F. L. Lewis, “Nonlinear network structures for feedback control,” *Asian Journal of Control*, vol. 1, no. 4, pp. 205–228, 1999.

[19] F. Lewis, J. Campos, and R. Selmic, *Neuro-Fuzzy Control of Industrial Systems with Actuator Nonlinearities*. SIAM, PA, 2002.

[20] B. Xian, D. M. Dawson, M. S. de Queiroz, and J. Chen, “A continuous asymptotic tracking control strategy for uncertain nonlinear systems,” *IEEE Trans. Automat. Contr.*, vol. 49, no. 7, pp. 1206–1211, Jul. 2004.

[21] W. E. Dixon, A. Behal, D. M. Dawson, and S. P. Nagarkatti, *Nonlinear Control of Engineering Systems: a Lyapunov-Based Approach*. Boston: Birkhäuser, 2003.

[22] H. K. Khalil, *Nonlinear Systems*. New Jersey: Prentice-Hall, 2002.

[23] M. Abu-Khalaf, J. Huang, and F. Lewis, *Nonlinear H-2/H-Infinity Constrained Feedback Control*. Springer, 2006.

[24] F. L. Lewis, *Optimal Control*. John Wiley & Sons, 1986.