Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Optimal Tracking Control for Unknown Nonlinear Systems Based on Locally Weighted Learning

Wenjie Dong and Jay A. Farrell

Abstract— **This paper considers the optimal tracking control of unknown nonlinear systems. To deal with the uncertainties in the system, a locally weighted learning observer (LWLO) is first proposed. Based on the proposed LWLO, analytic optimal controllers are proposed in the sense of pointwise min-norm. To show effectiveness of the proposed controllers, numerical simulations are presented.**

I. INTRODUCTION

Optimal control theory was formally developed about fifty years ago in the seminal works of L. S. Pontryagin [1] in the former Soviet Union and R. Bellman [2] in the United States. While Pontryagin introduced the minimum principle, which gave necessary conditions for the existence of optimal trajectories, Bellman introduced the concept of dynamic programming. The development of dynamic programming led to the notion of the celebrated Hamilton-Jacobi Bellman (HJB) partial differential equation, which had the value function as its solution.

For the Linear Quadratic Gaussian (LQG) problem [3], i.e., the $\mathcal{H}_2$ optimal control problem, the HJB partial differential equation becomes two separate Riccati equations, which could be solved very efficiently. However, LQG regulators can have arbitrarily small robustness margins [4]. To improve the robustness of the closed-loop optimal control, for linear systems the $\mathcal{H}_\infty$ control problem was proposed and was solved at the end of the 1980s [5].

For optimal control of general nonlinear systems, it is hard to obtain the optimal controllers efficiently. One reason is that the HJB equation is extremely hard to solve for nonlinear systems. To solve the optimal control of nonlinear systems, the receding horizon control method was proposed and is often used in industry. Receding horizon control is also known as moving horizon control or model predictive control. In receding horizon control, a finite horizon open-loop optimal control problem is solved online with the current state as an initial state; the optimization yields an optimal sequence and the first control in this sequence is applied to the plant [6]. Early results on receding horizon control did not consider the stability of the closed-loop system. To guarantee the stability, different terminal constraints may be introduced in solving for the optimal controller, such as the terminal equality constraint [7], [8], the terminal cost function [9], [10], the terminal constraint set [11], [12], the terminal cost and the constraint set [13], [14], [15], [16], and so on. In the terminal cost and constraint set methods, the stability of the closed-loop system is guaranteed by first finding a global control Lyapunov function (CLF) and then solving the receding horizon control by introducing additional state constraints that require the derivative of the CLF along the trajectory of the closed-loop system to be negative [17], [18], [19]. For the receding horizon control of nonlinear systems, analytic controllers are generally not available. The controllers are obtained by numeric approximation [20].

Wenjie Dong and Jay A. Farrell are with Department of Electrical Engineering, University of California, Riverside, USA

For the optimal control of uncertain nonlinear systems, one approach is to learn the unknown system model offline and then design optimal controllers based on the estimated models. Another way is to apply nonlinear $\mathcal{H}_\infty$ control theory and the optimal controllers are obtained based on the Hamilton-Jacobi-Isaac (HJI) equations [21], [22], which are hard to solve. neural Network based algorithms were proposed in [23] to optimize both $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms of performances for uncertain nonlinear systems, .

In this paper, we consider the optimal control of the uncertain nonlinear system shown in (1). To deal with the uncertain term, we first propose a locally weighted learning observer (LWLO) to estimate the unknown nonlinear system. Based on the approximation model, a pointwise min-norm problem is defined. For the defined optimal problem, analytic controllers are proposed based on a selected Lyapunov function. To show the effectiveness of the proposed optimal controller, a numeric example is presented.

II. PROBLEM STATEMENT

Consider an n-th order nonlinear system

$$\begin{aligned} \dot{x}_i &= x_{i+1}, \quad 1 \le i \le n-1 \\ \dot{x}_n &= f_0(x) + f(x) + g_0(x)u \end{aligned} \tag{1}$$

where $x = [x_1 \ldots, x_n]^\top \in R^n$ is the state, and $u \in R$ is the input. Functions f_0 and g_0 are known continuous functions. Function $f(x)$ is continuous in x and unknown. Furthermore, function $g_0(x)$ satisfies the following assumption.

Assumption 1: Function $g_0(x)$ is bounded below, i.e.,

$$g_0(x) > g_l(x) > c_g > 0$$

where c_g is a positive constant.

Given a desired bounded trajectory $x^d = [x_1^d, x_2^d, \ldots, x_n^d]^\top$ which satisfies

$$\dot{x}_1^d = x_2^d, \quad \dot{x}_2^d = x_3^d, \quad \cdots, \quad \dot{x}_{n-1}^d = x_n^d. \tag{2}$$

The problem discussed in this article is to design an optimal controller u such that the cost function

$$J_\infty = \int_0^\infty [(x - x^d)^\top Q(x - x^d) + u^2] d\tau \tag{3}$$

achieves its minimum, where Q is a positive definite matrix.

If $f(x)$ is known, a standard dynamic programming argument reduces the above optimal control problem to finding the value function V^* solving the Hamilton-Jacobi-Bellman partial differential equation (HJB)

$$V_x^* f_e - \frac{1}{4}\left(V_x^* g_e g_e^\top V_x^{*\top}\right) + (x-x^d)^\top Q(x-x^d) = 0 \quad (4)$$

where V_x^* denotes $\frac{\partial V^*}{\partial x}$, $f_e = [x_2, x_3, \ldots, x_n, f_0(x) + f(x)]^\top$, and $g_e = [0, \ldots, 0, g_0(x)]^\top$. If there exists a continuously differentiable positive definite solution to eqn. (4), then the optimal controller is

$$u = -\frac{1}{2} g_e^\top V_x^{*\top}. \quad (5)$$

In this article, there are two obstacles which prevent us from finding the optimal controller. The first one is that $f(x)$ is unknown. The second one is that it is extremely difficult to solve the HJB partial differential equation (4) even if $f(x)$ is known. To overcome the first obstacle, we propose a locally weighted learning observer (LWLO) to estimate $f(x)$. To deal with the second obstacle, we modify the optimal problem to a new one such that analytic controllers can be proposed.

III. LOCALLY WEIGHTED LEARNING OBSERVER

Let the observer be defined as follows.

$$\begin{aligned} \dot{\hat{x}}_i &= \hat{x}_{i+1}, \quad 1 \le i \le n-1 \\ \dot{\hat{x}}_n &= f_0(x) + \hat{f}(x) + g_0(x)u + v \end{aligned} \quad (6)$$

where $\hat{x} = [\hat{x}_1, \ldots, \hat{x}_n]^\top$ is the estimate of x, v is a stabilizing observer signal, $\hat{f}$ is the estimate of $f(x)$ based on a locally weighted learning (LWL) algorithm [24], [25], [26], [27], [28].

In LWL, the approximation of $f(x)$ at a point x is formed from the normalized weighted average of local approximators $\hat{f}_k(x)$ such that

$$\hat{f}(x) = \frac{\sum_k \omega_k(x)\hat{f}_k(x)}{\sum_k \omega_k(x)} \quad (7)$$

where each ω_k is nonzero only on a set denoted by S_k (defined below in eqn. (11)) over which the $\hat{f}_k$ will be adapted to improve their accuracy relative to f.

For $z = [z_1, \ldots, z_n]^\top = x - \hat{x}$, we have

$$\begin{aligned} \dot{z}_i &= z_{i+1}, \quad 1 \le i \le n-1 \\ \dot{z}_n &= f(x) - \hat{f}(x) - v. \end{aligned} \quad (8)$$

Let

$$e(t) = L^\top z(t) \quad (9)$$

where

$$\begin{aligned} L &= [l_1, l_2, \ldots, l_{n-1}, 1]^\top \\ &= [\lambda^{n-1}, C_{n-1}^1 \lambda^{n-2}, \ldots, C_{n-1}^{n-2}\lambda, 1]^\top \end{aligned} \quad (10)$$

λ is a positive constant, and $C_n^m = \frac{n!}{i!(n-i)!}$. We have the following lemma.

Lemma 1: ([29]) If $\lim_{t\to\infty} |e(t)| \le \mu_e$ where μ_e is a positive constant, then $\lim_{t\to\infty} |z_i| \le \frac{2^{i-1}}{\lambda^{n-i}}\mu_e$ for $1 \le i \le n$. Furthermore, if $\lim_{t\to\infty} e(t) = 0$, then $\lim_{t\to\infty} z_i = 0$ for $1 \le i \le n$.

By Lemma 1, to make the estimate $\hat{x}$ asymptotically converge to x, it is sufficient to choose suitable v and $\hat{f}_k$ such that e converges to zero.

A. Weighting Functions

For a given bounded compact operational region $\mathcal{D}^n \in R^n$, we define a continuous, non-negative and locally supported weighting function $\omega_k(x)$ for the k-th local approximator. Denote the support of $\omega_k(x)$ by

$$S_k = \Big\{ x \in \mathcal{D}^n \mid \omega_k(x) \ne 0 \Big\}. \quad (11)$$

Let $\bar{S}_k$ denote the closure of S_k. Note that $\bar{S}_k$ is a compact set. In this article, we choose ω_k as follows.

$$\omega_k(x) = \begin{cases} \left(1 - \left(\frac{||x-c_k||}{\mu}\right)^2\right)^2, & \text{if } ||x-c_k|| < \mu \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

where c_k is the center location of the k-th weighting function and μ is a constant which represents the radius of the region of support. The region of support is

$$S_k = \Big\{ x \in \mathcal{D}^n \mid \|x - c_k\| < \mu \Big\}.$$

We choose c_k such that $\|c_i - c_j\| = \frac{3}{2}\mu$ and $c_i \notin \bar{S}_j$ for any $i \ne j$. Since $\mathcal{D}^n$ is compact, there are finite local regions S_k $(1 \le k \le N)$. The centers c_j are selected such that

$$\mathcal{D}^n = \bigcup_{1 \le k \le N} S_k.$$

When $x(t) \in \mathcal{D}^n$, there exists at least one k such that $\omega_k(x) \ne 0$. For $x(t) \in \mathcal{D}^n$ the normalized weighting functions are defined as

$$\bar{\omega}_k(x) = \frac{\omega_k(x)}{\sum_{k=1}^N \omega_k(x)}.$$

The set of non-negative functions $\{\bar{\omega}_k(x)\}_{k=1}^N$ forms a *partition of unity* on $\mathcal{D}^n$:

$$\sum_{k=1}^N \bar{\omega}_k(x) = 1, \quad \text{for all } x \in \mathcal{D}^n.$$

Note that the support of $\omega_k(x)$ is exactly the same as the support of $\bar{\omega}_k(x)$.

When $x(t) \notin \mathcal{D}^n$, all $\omega_k(x)$ for $1 \le k \le N(t)$ are zero. Therefore, to complete the approximator definition of eqn. (7) to be valid for any $x \in \Re^n$:

$$\hat{f}(x) = \begin{cases} \sum_{k=1}^N \bar{\omega}_k(x)\hat{f}_k(x) & \text{if } x \in \mathcal{D}^n \\ 0 & \text{if } x \in \Re^n - \mathcal{D}^n. \end{cases} \quad (13)$$

In the reminder of this section, we will only consider the case when $x(t) \in \mathcal{D}^n$ to give all definitions for the LWL algorithm. Approaches that ensure $\mathcal{D}^n$ is attractive and invariant are discussed in [**?**].

B. Local Approximators

We define

$$f_k(x) = \bar{x}_k^T \theta_{f_k}^* \tag{14}$$

where

$$\bar{x}_k = \begin{bmatrix} 1 \\ x - c_k \end{bmatrix}$$

are the basis functions. For the function $f(x)$ in (1), the vectors $\theta_{f_k}^*$ denote the unknown optimal parameter estimates for $x \in \bar{S}_k$:

$$\theta_{f_k}^* = \arg\min_{\theta_{f_k}} \left(\int_{\bar{S}_k} \omega_k(x) \left| f(x) - \hat{f}_k(x) \right| dx \right) \tag{15}$$

where

$$\hat{f}_k(x) = \bar{x}_k^T \theta_{f_k} \tag{16}$$

Note that $\theta_{f_k}^*$ are well defined for each k because f and f_k are smooth on compact $\bar{S}_k$. Therefore, f_k will be referred to as the optimal local approximator to f on $\bar{S}_k$.

Let the approximation error on $\bar{S}_k$ be denoted as ϵ_{f_k}:

$$\epsilon_{f_k}(x) \quad = \quad f(x) - f_k(x). \tag{17}$$

In order for ϵ_{f_k} to be defined everywhere, let

$$\epsilon_{f_k}(x) = \begin{cases} f(x) - f_k(x), & x \in \bar{S}_k, \\ 0, & \text{otherwise.} \end{cases}$$

Since $f_k(x)$ is an approximation of $f(x)$ on S_k, the approximation error between $f(x)$ and $f_k(x)$ outside $\bar{S}_k$ is in fact not defined. Because f_k are multiplied by $\bar{\omega}_k$ or ω_k in all expressions and $\omega_k = \bar{\omega}_k = 0$ for $x \notin S_k$, the value of ϵ_{f_k} and f_k for $x \notin S_k$ is irrelevant. For simplicity, we choose the value of ϵ_{f_k} to be zero for $x \notin \bar{S}_k$. The controller will use a known design constant $\epsilon_f > 0$. We choose μ sufficiently small such that $|\epsilon_{f_k}(x)| \le \bar{\epsilon}_f$ for $x \in \bar{S}_k$ for some (unknown) positive constant $\bar{\epsilon}_f < \epsilon_f$. The existence of such μ is guaranteed by the continuity of the function $f(x)$. Note that the boundedness of $\max_{x \in \bar{S}_k}(|\epsilon_{f_k}(x)|)$ comes from the fact that $|\epsilon_{f_k}|$ are continuous on compact $\bar{S}_k$.

For any $x \in \mathcal{D}^n$, $f(x)$ can be represented as the weighted sum of the local approximators:

$$f(x) \quad = \quad \sum_k \bar{\omega}_k(x) f_k(x) + \delta_f(x). \tag{18}$$

This expression defines the approximation error $\delta_f(x)$ on $\mathcal{D}^n$ which satisfy $|\delta_f(x)| \le \bar{\epsilon}_f$ [29]. Therefore, if each local model $f_k(x)$ has accuracy $\bar{\epsilon}_f$ on $\bar{S}_k$, then the global accuracy of $\sum_k \bar{\omega}_k(x) f_k(x)$ on $\mathcal{D}^n$ also achieves at least accuracy $\bar{\epsilon}_f$ due to $\{\bar{\omega}_k\}_{k=1}^N$ forming a partition of unity on $\mathcal{D}^n$. The δ_f term in (18) is the *inherent approximation error* of $\hat{f}(x)$ for $f(x)$.

C. Update Laws

Since we assume that f is unknown, the parameter vector $\theta_{f_k}^*$ is unknown for each k. We update θ_{f_k} using the following adaptive laws

$$\dot{\theta}_{f_k} \quad = \quad \Gamma_{f_k} \bar{\omega}_k e \bar{x}_k \tag{19}$$

where Γ_{f_k} are positive constant matrices.

D. Stabilizing Observer Signal

To make the state of the locally weighted learning observer (6) asymptotically converge to the state x, the stabilizing observer signal is chosen as

$$v = l_1 z_2 + \cdots + l_{n-1} z_n + Ke + \frac{\epsilon_f e}{\sqrt{e^2 + \exp(-t)}} \tag{20}$$

where L is defined in (10), K is a positive constant, ϵ_f is a constant, and $\epsilon_f > |\epsilon_{f_k}|$.

Lemma 2: For system (6), with the stabilizing observer signal v defined in (20), locally weighted learning (13) and (16), update algorithm (19), then $(x - \hat{x})$ converges to zero and θ_{f_k} are bounded.

Proof: By eqn. (9), we have

$$\begin{aligned} \dot{e} &= l_1 z_2 + \cdots + l_{n-1} z_n + f(x) - \hat{f}(x) - v \\ &= l_1 z_2 + \cdots + l_{n-1} z_n + \sum_k \bar{\omega}_k(x)(f_k(x) - \hat{f}_k(x)) \\ &\quad + \delta_f(x) - v \\ &= l_1 z_2 + \cdots + l_{n-1} z_n + \sum_k \bar{\omega}_k(x) \bar{x}_k^T (\theta_{f_k}^* - \theta_{f_k}) \\ &\quad + \delta_f(x) - v \\ &= l_1 z_2 + \cdots + l_{n-1} z_n + \sum_k \bar{\omega}_k(x) \bar{x}_k^T \tilde{\theta}_{f_k} \\ &\quad + \delta_f(x) - v \end{aligned} \tag{21}$$

where $\tilde{\theta}_{f_k} = \theta_{f_k}^* - \theta_{f_k}$. Define the positive Lyapunov function

$$V_1 = \frac{1}{2} e^2 + \sum_k \tilde{\theta}_{f_k}^\top \Gamma_{f_k}^{-1} \tilde{\theta}_{f_k}.$$

Differentiating it along the solution of (21), we have

$$\begin{aligned} \dot{V}_1 &= -Ke^2 + e\delta_f - \frac{\epsilon_f e^2}{\sqrt{e^2 + \exp(-t)}} \\ &\le -Ke^2 + \epsilon_f \exp(-t/2). \end{aligned}$$

Therefore, V_1 is bounded by integrating both sides, which means that θ_{f_k} and e are bounded. By integrating both sides, it can be shown that e^2 is integrable. Therefore, e converges to zero. By Lemma 1, we can prove that $(x - \hat{x})$ converges to zero. ■

Remark 1: In the observer, we apply the locally weighted learning idea. The advantages of the locally weighted learning are two fold. First of all, the approximation errors are functions of local approximators. Secondly, the burden of the computation for learning is relieved.

Remark 2: In the observer, μ in (12) is a control parameter. It affects the number of local regions (N) and the magnitude of v through ϵ_f. If μ is large, in general N will be small but the magnitude of the last term in v may be large. Alternatively, as N is increased, the magnitude of the last term in v will decrease. So, the choice of μ involves a trade-off between the control magnitude and computation burden.

With the aid of Lemma 2, we can design optimal controllers for system (6), i.e., we can design an optimal controller u such that

$$\bar{J}_\infty = \int_0^\infty (\hat{x} - x^d)^\top Q(\hat{x} - x^d) + u^2)d\tau \tag{22}$$

achieves its minimum. Since $\hat{x}$ is close to x as time converges to infinity, x will converge to a small neighborhood of x^d if $\hat{x}$ converges to a small neighborhood of x^d.

IV. Pointwise Min-norm Controller

With the aid of the locally weighted learning observer, it may seem that the optimal control problem of (22) can be solved by using the dynamic programming technique. In fact, the optimal control problem (22) is not generally solvable because the dynamics of $\hat{x}$ is nonlinear. To obtain an analytical control law, we consider the pointwise min-norm problem proposed in [30], [16] instead. Before defining the problem, we need some preparation. Let $q = [q_1, \ldots, q_n]^\top$ and

$$q_i = \hat{x}_i - x_i^d, \quad 1 \le i \le n,$$

then

$$\begin{aligned} \dot{q}_i &= q_{i+1}, \quad 1 \le i \le n-1 \\ \dot{q}_n &= f_0(x) + \hat{f}(x) + g_0(x)u + v - \dot{x}_n^d. \end{aligned} \tag{23}$$

A control Lyapunov function (CLF) of system (23) is a continuously differentiable, positive definite function $V(q)$: $R^n \to R^+$ such that

$$\inf_u [V_q \bar{f} + V_q \bar{g} u] < 0 \tag{24}$$

for all $q \ne 0$ [31], [32], where

$$\bar{f} = [q_2, \ldots, q_n, f_0(x) + \hat{f}(x) + v - \dot{x}_n^d]^\top$$
$$\bar{g} = [0, \ldots, 0, g_0(x)]^\top.$$

If there is a CLF such that eqn. (24) is satisfied, the control input u obtained at each point from eqn. (24) can make the state of system (23) converge to zero. This can be seen when we choose V as a Lyapunov function under those control actions. For a general nonlinear system, it may be difficult to find a CLF or even to determine whether one exists. However, for system (23) there exists a CLF. In fact, the function

$$V = q^\top P q \tag{25}$$

is one of CLFs of system (23), where P is a positive definite matrix satisfying

$$P\Lambda + \Lambda^\top P = -Q \tag{26}$$

where Q is a positive definite matrix,

$$\Lambda = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_n \end{bmatrix}$$

and the constants α_i $(1 \le i \le n)$ are chosen such that matrix Λ is Hurwitz.

Given a control Lyapunov function $V(q)$ for system (23), the pointwise min-norm problem is defined as follows.

Pointwise Min-norm Problem:

$$\min_u \quad u^2 \tag{27}$$

such that

$$V_q(\bar{f} + \bar{g}u) \le -\sigma(q) \tag{28}$$

where σ is a positive definite function of q which is chosen by the designer.

Remark 3: In the pointwise min-norm problem, V can be any CLF of system (23). With different CLF V and σ, different optimal control u can be obtained. The function σ is a design parameter to be chosen according to the trade-off between the magnitude of the control effort and the closed-loop system performance.

For the pointwise min-norm problem, we have the following closed form solution.

Lemma 3: For the pointwise min-norm problem (27)-(28), the optimal control is

$$u = \begin{cases} -\dfrac{V_q \bar{f} + \sigma}{V_q \bar{g}}, & V_q \bar{g} \ne 0 \\ 0, & V_q \bar{g} = 0 \end{cases} \tag{29}$$

Proof: If $V_q \bar{g} = 0$, the constraint (28) holds automatically by (24). So $u = 0$ is the optimal control. If $V_q \bar{g} \ne 0$, the constraint (28) is active. We solve the optimal problem: $\min_u u^2$ such that $V_q(\bar{f} + \bar{g}u) + \sigma = 0$. By the Lagrange multiplier method, we obtain $u = -\frac{V_q \bar{f} + \sigma}{V_q \bar{g}}$. ♢

From Lemma 3, we can see that the optimal controller u depends on V and σ. An example min-norm optimal controller is given by Sontag's formula [32] as follows.

Lemma 4: For the pointwise min-norm problem (27)-(28), if

$$\sigma = \sqrt{(V_q \bar{f})^2 + q^\top Q q (V_q \bar{g} \bar{g}^\top V_q^\top)} \tag{30}$$

then the optimal control law is

$$u = \begin{cases} -\left[\frac{V_q \bar{f} + \sqrt{(V_q \bar{f})^2 + q^\top Q q (V_q \bar{g} \bar{g}^\top V_q^\top)}}{V_q \bar{g} \bar{g}^\top V_q^\top}\right] \bar{g}^\top V_q^\top, & V_q \bar{g} \ne 0 \\ 0, & V_q \bar{g} = 0 \end{cases} \tag{31}$$

Proof: Substitute σ into the optimal controller in Lemma 3, the lemma can be proved. ♢

It should be noted that if the control Lyapunov function V is the value function of the HJB equation corresponding to the cost function (22) when σ is chosen as (30). The optimal control (31) would be the solution to the optimal control problem (22). This fact leads us to choose σ as in (30). In [30], it was shown that every CLF is the value function of some meaningful cost function which means that it solves the HJB equation associated with a meaningful cost. This is referred to as "inverse optimal".

Combining the results in this subsection and the last subsection, we have the following result.

Theorem 1: For system (1) with the locally weighted learning observer defined in (6) with update laws (19) and stabilizing observer signal (20), the optimal control (31) solves the pointwise min-norm problem (27)-(28) with a given CLF $V(q)$ and make $(x - x^d)$ converges to zero.

Proof: By Lemma 4, the optimal control (31) solves the pointwise min-norm problem (27)-(28). Choose V in (28) as a Lyapunov function, $(\hat{x} - x^d)$ converges to zero since eqn. (28) holds for every point $\hat{x}$. By Lemma 2, $(x-x^d)$ converges to zero. ■

In Theorem 1, there are several control parameters. Constant μ determines the number of the local regions and the magnitude of the control input (see Remark 2). The control Lyapunov function V is an important control parameter. By suitably choosing V the performance of the closed-loop system with the controller (31) will be close to the performance of the closed-loop system with the optimal controller of the optimal control problem (22). There is no general approach for choosing CLF V. In practice, we can choose Lyapunov function V as in (25) by choice of Q, which also specifies J_∞ of (22).

V. Numerical Example

We consider for illustrative purpose a second order system given by

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= \sin(0.4(x_1+x_2)) + \Big(2+\sin(0.4(x_1+x_2))\Big)u.\end{aligned}$$

For the example, $x \in \Re^2, u \in \Re$ and we assume that there is only partial priori knowledge of the system nonlinearities. The known 'design model' has $f_0(x_1,x_2) = 0.4(x_1+x_2)$ and $g_0(x_1,x_2) = 2+\sin(0.4(x_1+x_2)$; therefore, the unknown design model error is

$$f(x) = \sin(0.4(x_1+x_2)) - 0.4(x_1+x_2).$$

Given a desired trajectory x^d, we want to design an optimal control in the sense of pointwise min-norm such that $(x-x^d)$ converges to zero.

The locally weighted learning observer is

$$\begin{aligned}\dot{\hat{x}}_1 &= \hat{x}_2\\ \dot{\hat{x}}_2 &= f_0(x) + \hat{f} + g(x)u + v\end{aligned}$$

where $\hat{f}$ is an online approximation to f with locally weighted learning algorithm (19) and v is defined in (20). In the locally weighted learning observer, we choose $\mu = 0.5$, $\lambda = 1$. The function approximation accuracies are specified as $\epsilon_f = 0.03$.

The weighting function is the biquadratic kernel of the form as

$$\omega_k(x) = \begin{cases} \left(1-R^2\right)^2, & \text{if } R<1\\ 0, & \text{otherwise.}\end{cases} \quad (32)$$

where

$$R = \left\| \frac{|x_1 - c_{k,1}|}{\mu}, \frac{|x_2 - c_{k,2}|}{\mu} \right\|_2.$$

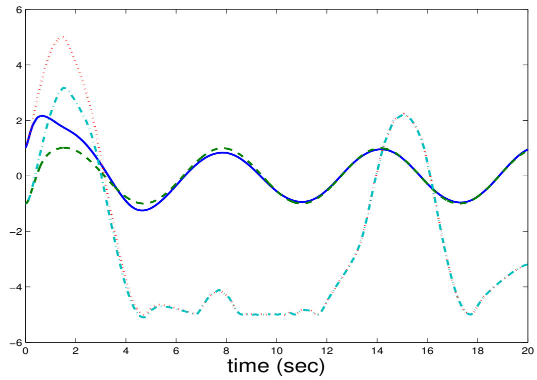

Fig. 1. Response of $\hat{x}_1$ and x_1 w/o learning (x_1 with learning: solid, $\hat{x}_1$ with learning: dashed, x_1 without learning: dotted, $\hat{x}_1$ without learning: dashdot)

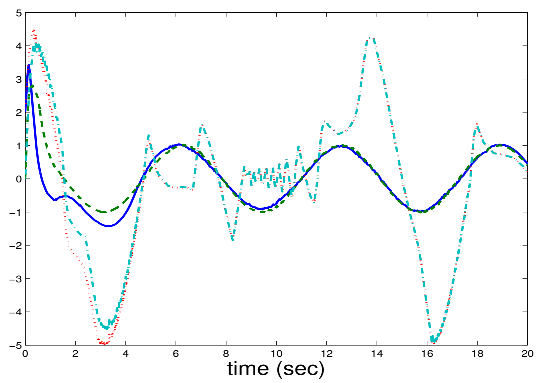

Fig. 2. Response of $\hat{x}_2$ and x_2 w/o learning (x_2 with learning: solid, $\hat{x}_2$ with learning: dashed, x_2 without learning: dotted, $\hat{x}_2$ without learning: dashdot)

The local basis function is

$$\bar{x}_k = \begin{bmatrix} 1 \\ x_1 - c_{k,1} \\ x_2 - c_{k,2}\end{bmatrix}$$

with c_k being the center of the $\bar{S}_k$. Therefore, f_k is the optimal local affine approximation to f on $\bar{S}_k$. We select $c_{k,1} = c_{k,2} = \frac{k\mu}{2}$ for $-20 \le k \le 20$. For simplicity, we choose the initial conditions of $\theta_{f_k}(0) = 0$ The adaptation rate matrices are set to $\Gamma_{f_k} = diag([1,1,1])$ where $diag(v)$ is the square diagonal matrix with diagonal component equal to the vector v.

For $x^d = [\sin t, \cos t]^\top$, Fig. 1 and Fig. 2 show the responses of $\hat{x}$ and x with/without learning. Fig. 3 shows the tracking errors $x - x^d$ with and without learning.

VI. Conclusion

This paper considers the optimal control of uncertain nonlinear systems. By applying locally weighted learning algorithms, asymptotical observer of the uncertain nonlinear system is proposed. Optimal controllers are proposed for the observer in the sense of poitwise min-norm. The advantage of the proposed method is that analytic optimal controllers are proposed and the stability of the closed-loop system is

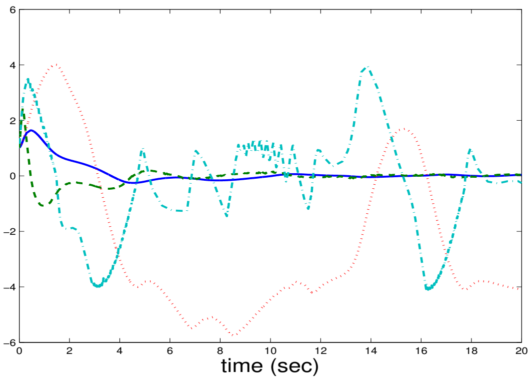

Fig. 3. Response of the tracking errors $x - x^d$ w/o learning ($x_1 - x_1^d$ with learning: solid, $x_2 - x_2^d$ with learning: dashed, $x_1 - x_1^d$ without learning: dotted, $x_2 - x_2^d$ without learning: dashdot)

guaranteed. Furthermore, if the control Lyapunov function V is the value function of the HJB equation corresponding to the cost function (22) when σ is chosen as (30). The optimal control (31) would be the solution to the optimal control problem (22).

Acknowledgements: This material is based upon work supported by the National Science Foundation under Grant No. ECCS-0701791. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

[1] L. S. Pontryagin, "Optimal control processes," *Uspehi Mat. Nauk (in Russian)*, vol. 14, pp. 3–20, 1959.
[2] R. Bellman, "The theory of dynamic programming," *Proc of Bat. Acad. Sci. USA*, vol. 38, 1952.
[3] M. Athans, "The role and use of the stochastic linear-quadratic-gaussian problem in control system design," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 529–552, 1971.
[4] J. C. Doyle, "Guaranteed margins for LQG regulators," *IEEE Transactions on Automatic Control*, vol. 23, no. 4, pp. 756–757, 1978.
[5] J. C. Doyle, K. Glover, P. Khargonekar, and B. Francis, "State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems," *IEEE Transactions on Automatic Control*, vol. 34, no. 8, pp. 831–847, 1989.
[6] D. Q. Mayne, J. B. Rawlings, C. V. Rao, and P. O. M. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, pp. 789–814, 2000.
[7] S. S. Keerthi and E. G. Gilbert, "Optimal, infinite horizon feedback laws for a general class of constrained discrete time systems: Stability and moving-horizon approximations," *Journal of Optimization Theory and Application*, vol. 57, pp. 265–293, 1988.
[8] L. Magni and R. Sepulchre, "Stability margins of nonlinear receding-horizon control via inverse optimality," *Systems and Control Letters*, vol. 32, pp. 241–245, 1997.
[9] R. R. Bitmead, M. Gevers, and V. Wertz, *Adaptive optimal control – The thinking man's GPC*. Englewood Cliffs, NJ: Prentice-Hall, 1990.
[10] M. Alamir and G. Bornard, "Stability of a truncated infinite constrained receding horizon scheme: The general discrete nonlinear case," *Automatica*, vol. 31, no. 9, pp. 1353–1356, 1995.
[11] H. Michalska and D. Q. Mayne, "Robust receding horizon control of constrained nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 38, pp. 1623–1632, 1993.
[12] L. Chisci, A. Lombardi, and E. Mosca, "Dual receding horizon control of constrained discrete-time systems," *European Journal of Control*, vol. 2, pp. 278–285, 1996.
[13] A. Jadbabaie, J. Yu, and J. Hauser, "Unconstrained receding-horizon control of nonlinear systems," *IEEE Transactions on Automatic Control*, pp. 776–783, 2001.
[14] H. Chen and F. Allgower, "A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability," *Automatica*, vol. 14, no. 10, pp. 1205–1217, 1998.
[15] T. Parisini and R. Zoppoli, "A receding horizon regulator for nonlinear systems and a neural approximation," *Automatica*, vol. 31, no. 10, pp. 1443–1451, 1995.
[16] J. A. Primbs, V. Nevistic, and J. C. Doyle, "Nonlinear optimal control: A control Lyapunov function and receding horizon perspective," *Asian Journal of Control*, vol. 1, no. 1, pp. 14–24, 1999.
[17] J. A. Primbs, "Nonlinear optimal control: A receding horizon approach," *PhD thesis, California Institute of Technology, Pasadena, CA*, 1998.
[18] A. Jadbabaie, *Receding Horizon Control of Nonlinear Systems: A Control Lyapunov Function Approach*. PhD thesis, California Institute of Technology, Pasadena, CA, 2000.
[19] P. Mhaskar, N. H. Ei-Farra, and P. D. Cristofides, "Stabilization of nonlinear systems with state and control constraints using Lyapunov based predictive control," *Systems and Control Letters*, vol. 55, pp. 650–659, 2006.
[20] J. M. Maciejowski and J. Marian, *Predictive Control with Constraints*. Harlow, England: Prentice-Hall, 2002.
[21] A. J. V. der Schaft, "On a state-space approach to nonlinear H_∞ control," *System and Control Letters*, vol. 16, pp. 1–8, 1991.
[22] A. Isidori and A. Astolfi, "Nonlinear H_∞ control via measurement feedback," *Journal of Mathmatics, System, Estimation, and Control*, vol. 2, pp. 31–34, 1992.
[23] M. Abu-Khalaf, J. Huang, and F. L. Lewis, *Nonlinear $\mathcal{H}_2/\mathcal{H}_\infty$ Constrained Feedback Control: A Practical Design Approach Using Neural Networks*. Springer Verlag, 2006.
[24] C. G. Atkeson, A. W. Moore, and S. Schaal, "Locally weighted learning for control," *Artifical Intelligence Review*, vol. 11, pp. 75–113, 1996.
[25] J. A. Farrell and Y. Zhao, "Self-organizing approximation based control," *Submitted to 2006 American Control Conference*, pp. 3378–3384, 2006.
[26] S. Schaal and C. G. Atkeson, "Constructive incremental learning from only local information," *Neural Computation*, vol. 10, no. 8, pp. 2047–2084, 1998.
[27] J. Nakanishi, J. A. Farrell, and S. Schall, "A locally weighted learning composite adaptive controller with structure adaptation," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 882–889, 2002.
[28] J. Nakanishi, J. A. Farrell, and S. Schaal, "Composite adaptive control with locally weighted statistical learning," *Neural Networks*, vol. 18, no. 1, pp. 71–90, 2005.
[29] Y. Zhao and J. A. Farrell, "Self-organizing approximation-based control for higher order systems," *IEEE Transactions on Neural Networks*, vol. 18, no. 4, pp. 1220–1231, 2007.
[30] R. Freeman and P. V. Kokotovic, "Inverse optimality in robust stabilization," *SIAM J. Control and Optimization*, vol. 34, pp. 1365–1391, 1996.
[31] E. D. Sontag, "A Lyapunov-like characterization of asymptotic controllability," *SIAM J. Control and Optimization*, vol. 21, no. 3, pp. 462–471, 1983.
[32] ——, "A 'universal' construction of Artstein's theorem on nonlinear stabilization," *System and Control Letters*, vol. 13, no. 2, pp. 117–123, 1989.