2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Model–Free Learning Control of Nonlinear Discrete–Time Systems

Nader Sadegh
The George W. Woodruff School of Mechanical Engineering
Georgia Institute of Technology, Atlanta Georgia 30332–0405
sadegh@gatech.edu

***Abstract*— A model–free learning controller for a general class of nonlinear discrete–time state–space systems is introduced. The learning component of the proposed controller can use an arbitrary function approximator such as a Polynomial, Radial Basis, or Neural Network to directly learn the inverse of the input-state mapping of the plant while forcing its state to track a prescribed desired trajectory. Unlike most of the existing direct adaptive or learning schemes, the nonlinear plant is *not* assumed to be feedback linearizable. The developed controller is subsequently applied to control the configuration of a nonholonomic differential drive robot. The simulation results of this application demonstrate a significant improvement in the tracking performance of the robot once the control input is fully learned.**

I. Introduction

Many practical control applications particularly those in the areas of motion control and robotics demand precise tracking of a specified desired trajectory without requiring a precise knowledge of the system's model. The learning control techniques evolved during the last two decades [1], [2], [3], [4], [5], [6] offer an effective solution for many applications ranging from robotics [7], [8] to fluid power applications [9], [10], [11] to disk drive actuation [6] and magnetic levitation [5].

The majority of the existing works on tracking control of nonlinear systems, particularly in the continuous–time domain, are based on feedback linearization; see for example [12], [13], [14], [5]. Although very powerful when applicable, they generally suffer from several drawbacks including stringent integrability conditions, in the case of input–state linearization, or the minimum phase requirement in the case of output feedback linearization. In spite of attempts to relax the minimum phase requirement in the continuous–time domain [15], [3], the resulting controllers are still either model based [15] or have a limited scope.

In this paper we formulate a learning control system composed of a nonlinear state feedback controller together with a learning mechanism for a fairly general class of nonlinear discrete–time systems. The learning component of the proposed control system is general enough to incorporate many commonly used function approximators including Polynomial, Radial Basis Functions [16], and Nodal Link Perceptron Networks (NLPN) [2], [17]. It enables the control system to learn the inverse of the input-state mapping of the plant while forcing its state to follow a prescribed desired trajectory. The main requirement that we impose here is for the plant to be controllable within the set of admissible states to be specified. Moreover, we provide a rigorous proof of stability for the resulting closed–loop control system. The learning controller of this paper has a similar structure to those in [1] and [2] but relaxes the key requirement of knowing the plant gradient information to implement the control law.

Another key approach exploited in formulating the state feedback part of our integrated controller is that of 'lifting' or 'block' modeling technique [18], [19], [20]. This approach, which is mainly applicable to discrete–time or sampled data systems, has proven highly effective for many problems arising from control, estimation, and even realization of linear and nonlinear [21], [22], [20] dynamical systems. In the block input–state realization considered here, the system is represented with a state-space realization that has at least as many inputs as the order of the system, even if the original system has fewer inputs. As a result, the inverse map of the system relating the input to the state variables [1], [2], [20], [14] can be determined by inverting an algebraic function.

Following our theoretical developments, we apply the proposed controller to a nonholonomic differential drive robot [23] for tracking control of its configuration. The novelty of the proposed control scheme lies in that it controls three degrees of freedom using only two inputs. The learning component of the control system takes the desired state of the robot as input and, once trained, generates the required actuation inputs.

A. Notations

The following notations are freely used in the paper: For functions f and g, $f \circ g$ denotes the function composition, i.e., $f \circ g(x) = f(g(x))$ for all x in the domain of g; $Df(x) = \frac{\partial f(x)}{\partial x}$; $D_1 f(x,y) = \frac{\partial f(x,y)}{\partial x}$; $D_2 f(x,y) = \frac{\partial f(x,y)}{\partial y}$. The notation $h_y(x)$ or $h_x(y)$ means $h(x,y)$ for a fixed y or x, respectively. Also, $h_y \circ f(x) = h(f(x), y)$ for a fixed y. For $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $\|x\|$ denotes the norm of x; for $y \in \mathbb{R}^m$ $(x,y) = (x_1, \ldots, x_n, y_1, \ldots, y_m)$; $B^n_\varepsilon = \{x \in \mathbb{R}^n : \|x\| < \varepsilon\}$. For matrices M_1, M_2, define $[M_1, M_2] := \begin{bmatrix} M_1 & M_2 \end{bmatrix}$ and $(M_1, M_2) := \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}$. The notation x_k for a scalar x denotes the value of vector or scalar x at discrete–time k. The i–th component of a vector x_k is denoted by $x_{i,k}$. The expression $y = O(x)$ means that $y \to 0$ as $x \to 0$ while the expression $x = o(y)$ means that $x/y \to 0$ as $y \to 0$ or equivalently $x = O(y)y$. In some cases, the short hand notation $x = O(y_1, \ldots, y_n)$ is used to denote $\|x\| \to 0$ as $\|y\| \to 0$.

II. Problem Statement

Consider an n–th order discrete–time system given by the state-space representation

$$x_{k+1} = F(x_k, u_k) \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^n$ is the input, and $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is the state transition map assumed to be at least continuously differentiable. Throughout the paper we shall assume that the Jacobian of F with respect to x, $D_1F(x, u)$, is nonsingular everywhere. Note that this is not too restrictive an assumption as it is automatically satisfied for discrete–time systems resulting from discretizing continuous–time systems that are both forward and backward integrable.

Remark 1: When the plant is not in the form of (1), the Block Input-State approach [24], [25] can be used to *lift* the dimension of the input vector to that of the state resulting in a *square* system. For instance, a single–input system $x_{k+1} = f(x_k, u_k)$ can be lifted to the required form by defining the state and input vectors to be $\mathbf{x}_k = x_{nk}$ and $\mathbf{u}_k = (u_{nk}, \ldots, u_{nk+n-1})$, respectively. In this case, the resulting square system is given by $\mathbf{x}_{k+1} = F(\mathbf{x}_k, \mathbf{u}_k)$ where $F(\mathbf{x}, \mathbf{u}) := f_{u_i} \circ f_{u_{i-1}} \circ \cdots \circ f_{u_2} \circ f_{u_1}(\mathbf{x})$. As can be seen in Section (V), even multi–input underactuated systems can be discretized to this form.

To state the control objective precisely, we need to specify the set of all *admissible* desired states and outputs. We let the set of admissible desired states, denoted by $\mathcal{X}$, be a bounded and convex open subset of $\mathbb{R}^n$ (e.g., $\mathcal{X} \subset \{x \in \mathbb{R}^n : \|x\| < r\}$ for some finite $r > 0$).

The control objective is to formulate a controller that forces the state of the plant x_k to asymptotically converge to x_k^d, a prescribed desired state trajectory. This is to be accomplished without requiring exact information about the state–space map F of the plant. In fact, it is desired to learn the inverse input-state map of the plant at the same time that the state tracking control is enforced.

The partial derivatives of F, which play a key role in our subsequent development, are denoted by D_1F and D_2F. In particular, D_2F, which shall be referred to as the controllability matrix, is the generalizations of its counterpart for linear systems.

To ensure controllability of the system (1), we make the following assumptions:

A_1 System (1) is strongly controllable (with respect to $\mathcal{X}$) [19], [14]: $\forall x, z \in \mathcal{X}$, there exists a unique input vector $u \in \mathbb{R}^n$ such that $F(x, u) = z$. We shall denote the corresponding set of *admissible* inputs by

$$\mathcal{U} = \{u \in \mathbb{R}^n : z = F(x, u), \text{ for some } x, z \in \mathcal{X}\}$$

which is nonempty by the assumption.

A_2 The controllability matrix, $D_2F(x, u)$, has full rank for all $u \in \overline{\mathcal{U}}$ and $x \in \overline{\mathcal{X}}$.

Remark 2: The assumptions made here about the system are rather standard controllability assumptions for nonlinear systems also adopted by others [26], [21], [14], [27] in the literature. They are automatically satisfied if the linearized system is controllable about an equilibrium point with $\mathcal{X}$ confined to an open neighborhood of the equilibrium state (see for example [14], [27]). Assumptions A_1 and A_2 guarantee that any initial state can be transferred to any final state by means of a control sequence of length n. In the case of a linear system, this definition coincides with the usual controllability definition.

The following proposition shows that fulfillment of Assumptions A_1 and A_2 imply existence of a continuously differentiable control function for the system.

Proposition 1: Consider the discrete–time system (1) satisfying Assumptions A_1 and A_2. There exists $\Psi \in C^1(\mathcal{X} \times \mathcal{X}, \mathcal{U})$ such that:
i) The admissible input set, $\mathcal{U}$, is bounded and open.
ii) For all $x, z \in \mathcal{X}$ and $u \in \mathcal{U}$ we have $u = \Psi(z, x) \iff z = F(x, u)$.
Moreover, the Jacobians of $\Psi(z, x)$ are given by $D_1\Psi = (D_2F)^{-1}$and $D_2\Psi = -(D_2F)^{-1}D_1F$.

Proof: Assumption A_1 guarantees the existence of $\Psi : \mathcal{X} \times \mathcal{X} \to \mathcal{U}$, which uniquely determines $\mathbf{u}$ in terms of $\mathbf{x}$ and $\mathbf{z}$: $\mathbf{u} = \Psi(\mathbf{x}, \mathbf{z})$. Consider the map

$$\bar{F}(x, u) = (x, F(x, u))$$

By the hypothesis, the Jacobian of this map,

$$D\bar{F}(\mathbf{x}, \mathbf{u}) = \begin{bmatrix} I_{n\times n} & D_1F \\ \mathbf{0} & D_2F \end{bmatrix}$$

is invertible at any $x \in \mathcal{X}$ and $u \in \mathbb{R}^n$. The inverse function theorem [28] implies that $\bar{F}(x, u)$ has a unique *local* (not necessarily covering the entire $\mathcal{X}$) inverse $\bar{F}^{-1}(x, z)$ such that $\bar{F}(\bar{F}^{-1}(x, z)) = (x, z)$. Moreover, $D\bar{F}^{-1}(x, z) = \left[D\bar{F}(x, u)\right]^{-1}$. Noting that $\bar{F}^{-1}(x, z) = (x, \Psi(x, z))$ for all (x, z) in the domain of $\bar{F}^{-1}$, it follows that $\Psi(x, z) \in C^1(\mathcal{X} \times \mathcal{X}, \mathcal{U})$ with the Jacobians specified in the statement of the proposition. The set $\mathcal{U} = \Psi(\mathcal{X}, \mathcal{X})$, which is open by the continuity of $\bar{F}(x, u)$ (i.e., $(\mathcal{X},\mathcal{U}) = \bar{F}^{-1}(\mathcal{X}, \mathcal{X})$), and bounded since its closure $\overline{\mathcal{U}} = \Psi(\overline{\mathcal{X}}, \overline{\mathcal{X}})$ is compact. ■

III. Learning Control Formulation

In this section we formulate a nonlinear learning control law that forces the state of the system $\mathbf{x}_k$ to follow a sequence of desired states, $\mathbf{x}_k^d \in \mathcal{X}$ without the exact knowledge of the system model. Given that $\mathbf{x}_k \in \mathcal{X}$ and $\mathbf{x}_{k+1}^d \in \mathcal{X}$, by proposition 1, setting the control input to

$$u_k = \Psi(x_{k+1}^d, x_k) \tag{2}$$

gives $x_{k+1} = F(x_k, u_k) = x_{k+1}^d$. Of course, as can be seen, this control law requires the knowledge of the nonlinear control function $\Psi(.,.)$, which may be thought of as the algebraic inverse of (1) relative to the input vector. The main goal of the remainder of this paper is to formulate a learning controller that can generate the required input sequence without requiring the system model. To make this formulation possible we shall make 2 additional assumptions regarding the control function Ψ and the desired state trajectory:

A$_3$ There exists a sequence of *known* functions $\phi_i \in C^1(\mathcal{X} \times \mathcal{X}, \mathbb{R})$, $i = 1, 2, \ldots$, called the *basis* functions, such that $\overline{\text{span}}\{\phi_i\}_{i=1}^{\infty} \supset C^1(\mathcal{X} \times \mathcal{X}, \mathbb{R})$. Equivalently, given $\Psi \in C^1(\mathcal{X} \times \mathcal{X}, \mathbb{R}^n)$ and $\varepsilon > 0$, there exist a positive integer N and $\mathbf{w}_i \in \mathbb{R}^n$, $i = 1, \ldots, N$, such that $\sup_{x,z} \left\| \sum_{i=1}^{N} \mathbf{w}_i \phi_i(z, x) - \Psi(z, x) \right\| < \varepsilon$. In the sequel, we shall use the compact notations $\Phi = (\phi_1, \cdots, \phi_N)$, $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_n]$, and $\mathbf{W}^s = (\mathbf{w}_1, \ldots, \mathbf{w}_n)$ where the superscript s stands for the *stack* operator.

A$_4$ The desired state sequence $x_k^d \in \mathcal{X}$ is Persistently Exciting (PE) relative to Φ: There exist $\underline{\lambda} > 0$ and integer $s > 0$ (depending on N) such that

$$\sum_{j=k}^{k+s-1} \Phi(x_{j+1}^d, x_j^d)\Phi(x_{j+1}^d, x_j^d)^T \geq \underline{\lambda} I_{N\times N}$$

holds, $\forall\ k \geq 0$.

Remark 3: The type of basis functions specified in Assumption A_3 for approximating the control function Ψ is quiet general. For instance, it is well known that polynomials may be used to approximate continuous functions over a compact set to within a prescribed accuracy. Other examples of basis functions are Radial Basis Functions [16] and Nodal Link Perceptron Network (NLPN) [2]. The PE assumption (A$_4$) is similar to those made in [2], [17] and guarantees the convergence of the parameter estimation error to zero in the absence of any functional approximation error (i.e., $\varepsilon = 0$).

In view of Assumption A$_3$, the control function $\Psi(.,.) = \mathbf{W}^T\Phi(.,.)$ in (2) may be estimated via

$$u_k = \hat{\Psi}\left(x_{k+1}^d, x_k\right) := \hat{\mathbf{W}}_k^T \Phi\left(x_{k+1}^d, x_k\right) \tag{3}$$

where $\hat{\mathbf{W}}_k$ is the k–th *estimate* of $\mathbf{W}$ to be specified shortly. Even if $\hat{\mathbf{W}}_k = \mathbf{W}$, the control law (3) is a deadbeat control action often leading to input saturation and other undesirable behavior. An alternative approach may be considered using the affine Taylor series approximation

$$\hat{\Psi}\left(x_{k+1}^d, x_k\right) = \hat{\Psi}\left(x_{k+1}^d, x_k^d\right) + \left[D_2\hat{\Psi}\left(x_{k+1}^d, x_k^d\right)\right] e_k + o\left(|e_k|\right)$$

where $e_k = x_k - x_k^d$ is the state tracking error. This result motivates the use of the following *non-deadbeat* feedforward control law

$$u_k = \hat{\mathbf{W}}_k^T \Phi\left(x_{k+1}^d, x_k^d\right) + \hat{K}_k e_k \tag{4}$$

where $\hat{K}_k = \alpha_k \hat{\mathbf{W}}_k^T D_2\Phi\left(\mathbf{x}_{k+1}^d, \mathbf{x}_k^d\right)$ for some $0 < \alpha_k \leq 1$. Denoting the control function approximation error by

$$d_k = \Psi\left(x_{k+1}^d, x_k^d\right) - \mathbf{W}^T\Phi\left(\mathbf{x}_{k+1}^d, \mathbf{x}_k^d\right)$$

and applying the control law (4) to system (1) results in the closed-loop *error* dynamics $e_{k+1} = G_k(e_k, \tilde{\mathbf{W}}_k^s, d_k)$ where

$$\begin{aligned} G_k(e_k, \tilde{\mathbf{W}}_k^s, d_k) \quad = \quad & F(x_k^d + e_k, u_k^d + \tilde{\mathbf{W}}_k^T\Phi_k + K_k e_k \\ & + \alpha_k \tilde{\mathbf{W}}_k^T \Phi'_k e_k + d_k) - \mathbf{x}_{k+1}^d \end{aligned}$$

$\Phi_k = \Phi\left(x_{k+1}^d, x_k^d\right)$, $\Phi'_k = D_2\Phi\left(x_{k+1}^d, x_k^d\right)$, $u_k^d = \mathbf{W}^T\Phi_k$, $\tilde{\mathbf{W}}_k = \hat{\mathbf{W}}_k - \mathbf{W}$, and $K_k = \alpha_k \mathbf{W}^T\Phi'_k$. Using Proposition 1.*ii*, it can be seen that the linearized error dynamics $e_{k+1} = G_k(e_k, \tilde{\mathbf{W}}_k, 0)$ about $(0, 0, 0)$ is given by

$$e_{k+1} = (D_1\Psi)^{-1}\left[(\alpha_k - 1)(D_2\Psi)e_k + \tilde{\mathbf{W}}_k^T\Phi_k\right] \tag{5}$$

Based on this error equation, we propose the following Steepest Decent (SD) estimation law to minimize $\|e_{k+1}\|^2$ with respect to $\hat{\mathbf{W}}_k$:

$$\hat{\mathbf{W}}_{k+1} = \hat{\mathbf{W}}_k - c_k \Phi_k \bar{e}_{k+1}^T \tag{6}$$

where $\sup_{k\geq 1} c_k \|\Phi_k\|^2 < 2$, $0 < \underline{c} \leq c_k \leq \bar{c} < \infty$ for some constant scalars $\underline{c}$ and $\bar{c}$, and

$$\bar{e}_k = (D_1\hat{\Psi})e_k + (1 - \alpha_{k-1})(D_2\hat{\Psi})e_{k-1} \tag{7}$$

Unfortunately, this update law may be unrealizable as it requires knowledge of the future value of the measured error. However, this problem can be easily fixed by using an *augmented* error, similar to those in [2], [29]. Define the augmented error to be

$$e_k^a = \bar{e}_k + (\hat{\mathbf{W}}_k^T - \hat{\mathbf{W}}_{k-1}^T)\Phi_{k-1} \approx \tilde{\mathbf{W}}_k^T\Phi_{k-1} \tag{8}$$

The SD update law corresponding to (8) becomes

$$\hat{\mathbf{W}}_{k+1} = \hat{\mathbf{W}}_k - c_k\Phi_{k-1}e_k^{aT} \tag{9}$$

in matrix form, or equivalently $\hat{\mathbf{w}}_{i,k+1} = \hat{\mathbf{w}}_{i,k} - c_k\phi_{i,k-1}e_k^a$, both of which are realizable. In view of (9) the augmented error in (8) can be updated recursively via

$$e_k^a = \bar{e}_k - c_{k-1}\Phi_{k-2}^T\Phi_{k-1}e_{k-1}^a \tag{10}$$

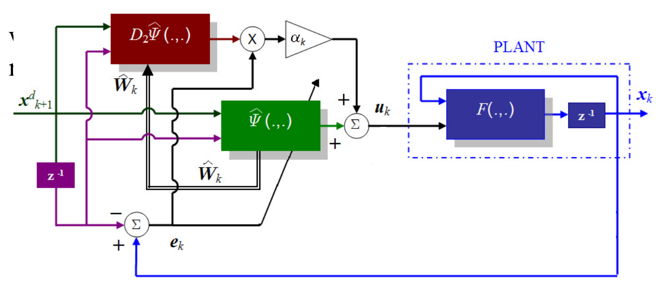

Fig. 1. Learning Control Block Diagram

Remark 4: The combined feedback control (4) and parameter estimation (9) laws, referred to as the *learning* control law (see Figure 1), is model–free and only requires the plant to be controllable.

IV. Stability Analysis

The stability analysis of the overall closed loop system is carried out in several steps. We first show that the linearized closed–loop error system given by (5) and (9) is *exponentially* stable. Subsequently, we will show that the original nonlinear error dynamics coupled with the estimation law (9) is *locally* asymptotically stable in the absence of any control function approximation error and bounded–input bounded–output (BIBO) stable otherwise.

The following Lemma (see Appendix A for a proof) shows that the difference equations given by (10) and (9) for the disturbance free linearized error system is *uniformly exponentially* stable if the PE Assumption (A_4) is satisfied. To state the lemma, we first need to express the evolution of the parameter error matrix ($\tilde{\mathbf{W}}_k$) for the linearized error system. This can achieved by subtracting $\mathbf{W}$ from both sides of (9) and replacing the augmented error by $\mathbf{e}^a_k = \tilde{\mathbf{W}}^T_k \Phi_{k-1}$:

$$\tilde{\mathbf{W}}_{k+1} = \mathbf{A}_k \tilde{\mathbf{W}}_k,\ \mathbf{A}_k = I - c_k \Phi_{k-1}\Phi^T_{k-1} \qquad (11)$$

Lemma 1: The linear time–varying system (11), $\tilde{\mathbf{W}}_{k+1} = \mathbf{A}_k \tilde{\mathbf{W}}_k$ is *uniformly* exponential stable if the desired state sequence $\mathbf{x}^d_k \in \mathcal{X}$ is PE (Assumption A_4). More precisely, there exists $0 \leq \sigma < 1$ such that $\left\|\tilde{\mathbf{W}}_{k+s}\right\| \leq \sigma \left\|\tilde{\mathbf{W}}_k\right\|$, $\forall k \geq 0$, or equivalently $\left\|\mathbf{\Gamma}_{(k+s,k)}\right\| \leq \sigma$, $\forall k \geq 0$, where $\mathbf{\Gamma}_{(k+s,k)} := \mathbf{A}_{k+s-1} \cdots \mathbf{A}_k$ is the system state–transition matrix and s is as specified in Assumption A_4.

The next Theorem (see Appendix B for a proof), which is important in its own right, shows that uniform exponential stability of the linearized error system implies its local asymptotic and BIBO stability of the original system.

Theorem 1: Consider the nonlinear time varying system $\xi_{k+1} = F_k(\xi_k, \mu_k)$, $F_k \in C^1(B^m_{\bar{\xi}} \times \mathbb{R}^p, \mathbb{R}^m)$, for some $\bar{\xi} > 0$ and integers $m, p \geq 1$. Let $A_k := D_1 F_k(0,0)$, $B_k := D_2 F_k(0,0)$ and suppose that $F_k(0,0) = 0$, $\forall k \geq 0$, $a := \sup_{k\geq 1} \|A_k\| < \infty$, $b := \sup_{k \geq 1}\|B_k\| < \infty$, and that the linearized system, $\tilde{\xi}_{k+1} = A_k \tilde{\xi}_k$, is uniformly exponentially stable, i.e., there exist $s \geq 1$ and $0 < \sigma < 1$ such that $\left\|\tilde{\xi}_{k+s}\right\| \leq \sigma \left\|\tilde{\xi}_k\right\|$, $\forall k \geq 0$. There exist $\bar{\mu}$ and $\bar{\xi}_0$ such that if $\|\xi_0\| < \bar{\xi}_0$ and $\sup_{k\geq 1} \|\mu_k\| < \bar{\mu}$, then $\|\xi_k\| < \bar{\xi}$, $\forall k \geq 0$. Moreover the upper limit $\overline{\lim}_{k\to\infty} \|\xi_k\| = O(\overline{\lim}_{k\to\infty} \|\mu_k\|)$.

We are finally in position to state and prove our main stability Theorem:

Theorem 2: Consider the discrete-time system (1) subjected to the learning control law given by (4) and (9) satisfying Assumptions A_1–A_4 and further assume that $x^d_k + B^n_{\bar{e}} \subset \mathcal{X}$, $\forall k \geq 0$, for some $\bar{e} > 0$. There exist positive scalars $\bar{\alpha}$, $\bar{e}_0$, $\bar{w}_0$, and $\bar{\varepsilon}$ such that if $\sup_{k\geq 0} |1 - \alpha_k| < \bar{\alpha}$, $|e_0| < \bar{e}_0$, $\left\|\tilde{\mathbf{W}}\right\| < \bar{w}_0$, and $\varepsilon \leq \bar{\varepsilon}$, then e_k and u_k are bounded for all k. Moreover, $\overline{\lim}_{k\to\infty} |e_k| = O(\varepsilon)$.

Proof: Defining the overall error state vector by $\xi_k = (e_k, e_{k-1}, e^a_k, \tilde{\mathbf{W}}^s_k)$ then the closed-loop error system is given by $\xi_{k+1} = \bar{G}_k(\xi_k, d_k)$ where $\bar{G}_k$ is the state map specified by

$$\begin{aligned} e_{k+1} &= G_k(e_k, \tilde{\mathbf{W}}^s_k, d_k) \\ \tilde{\mathbf{W}}_{k+1} &= \tilde{\mathbf{W}}_k - c_k \Phi_{k-1} e^{aT}_k \end{aligned}$$

together with (7), (10), and $\hat{\Psi} = (\mathbf{W} + \tilde{\mathbf{W}}_k)^T \Phi_k$. Linearizing $\xi_{k+1} = \bar{G}_k(\xi_k, 0)$ about $\xi_k = 0$ yields $\bar{\mathbf{e}}_k = \tilde{\mathbf{W}}_{k-1}\Phi_{k-1}$, $\mathbf{e}^a_k = \tilde{\mathbf{W}}_k \Phi_{k-1}$,

$$\begin{aligned} \mathbf{e}_{k+1} &= (1-\alpha_k)(D_1 F)\mathbf{e}_k + (D_1\Psi)^{-1}\tilde{\mathbf{W}}^T_k \Phi_k \\ \tilde{\mathbf{W}}_{k+1} &= \mathbf{A}_k \tilde{\mathbf{W}}_k,\ \mathbf{A}_k = I - c_k \Phi_{k-1}\Phi^T_{k-1} \end{aligned}$$

where we used that $D_1 F = -(D_1\Psi)^{-1}(D_2\Psi)$. By Lemma (1), the 2nd subsystem $\tilde{\mathbf{W}}_{k+1} = \mathbf{A}_k \tilde{\mathbf{W}}_k$ is uniformly exponentially stable. The first subsystem, hence the overall system, is also exponentially stable provided that

$$\bar{\alpha} \sup_{\mathbf{z},\mathbf{x} \in \mathcal{X}} \|D_1 F(\mathbf{z}, \mathbf{x})\| < 1$$

where $\bar{\alpha} = \sup_{k\geq 1} |1 - a_k|$. Thus the hypothesis of Theorem (1) is satisfied with $\bar{\xi} = \bar{e}$ and the conclusions follow. ■

V. Application to a Differential Drive Robot

In this section we apply the learning controller developed in the preceding section to a nonholonomic nonlinear system arising from a differential wheel mobile robot [23] (Figure 2). The equation of motion for this system is given by $\dot{\mathbf{x}} = g_1(\mathbf{x})v_R + g_2(\mathbf{x})v_L$ where $\mathbf{x} = (x, y, \theta)$ specifies the robot configuration (position and orientation), v_R and v_L are the linear speeds of the right and left wheels, respectively, $g_1(\mathbf{x}) = \frac{1}{2}(\cos\theta, \sin\theta, 1/L)$, and $g_2(\mathbf{x}) = \frac{1}{2}(\cos\theta, \sin\theta, -1/L)$.

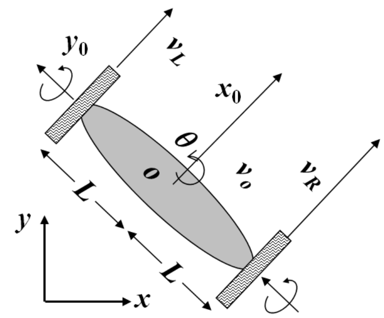

Fig. 2. Differential Drive Robot

To discretize this system and transform it into the state–space model (1) the inputs $u_1(t)$ and $u_2(t)$ are chosen to be piecewise constant and are collectively identified with the discrete–time input vector $u_k = [u_{1k}\ u_{2k}\ u_{3k}]^T \in \mathbb{R}^3$. Letting T denote the sampling period, the relationship between the inputs $v_R(t)$ and $v_L(t)$ on $[kT, (k+1)T]$ and u_k is chosen to be

$$\begin{aligned} v_R(t) &= u_{1k} + u_{2k}\sigma^1_k + u_{3k}\sigma^2_k \\ v_L(t) &= u_{1k} - u_{2k}\sigma^1_k - u_{3k}\sigma^2_k \end{aligned}$$

where $\delta = T/2$ and $\sigma^j_k(t) = 1$ if $kT + (j-1)\delta \leq t < kT + j\delta$ and 0 otherwise. Note that the first component of u_k is the robot linear speed and the last two components are proportional to robot angular speed over the k–th sampling period. Integrating $\dot{\mathbf{x}} = g_1(\mathbf{x})v_R + g_2(\mathbf{x})v_L$ from $t = kT$ to $t = (k+1)T$ gives $\mathbf{x}(kT + T)$ as a function of $\mathbf{x}(kT)$ and u_k:

$$\begin{aligned} x_{k+1} &= x_k + T u_{1k} f_1(\theta_k, u_{2k}, u_{3k}) \\ y_{k+1} &= y_k + T u_{1k} f_2(\theta_k, u_{2k}, u_{3k}) \\ \theta_{k+1} &= \theta_k + T(u_{2k} + u_{3k})/2 \end{aligned}$$

where f_1 and f_2 are smooth functions. It can be shown that the Jacobian ($D_1 F$) and the Controllability ($D_2 F$) matrices associated with this system are nonsingular everywhere thus satisfying Assumptions A_1 and A_2. The control function Ψ is modeled as a cubic polynomial function of $(\delta x', \delta y')$ where $\delta x'$ and $\delta y'$ are the changes in the x and y coordinates of the robot position relative to (x_0, y_0) frame attached to the robot (see Figure 2). A randomly selected desired trajectory consisting of 5000 points were used during the learning process ($T = 0.1$ sec). The feedback and learning control gain values of $\alpha_k = 0.9$ and $c_k = 0.2$, $\forall k \geq 0$, were used in the simulations. No other information regarding the system model was assumed or used. The simulation results show that the norm of the tracking error is reduced by 2 orders of magnitude within 5000 trials (Figure 3). Even though not shown, the learned linear (u_{1k}) and angular velocity (u_{2k} and u_{3k}) inputs are well–behaved and stay within their intended lower and upper bounds.

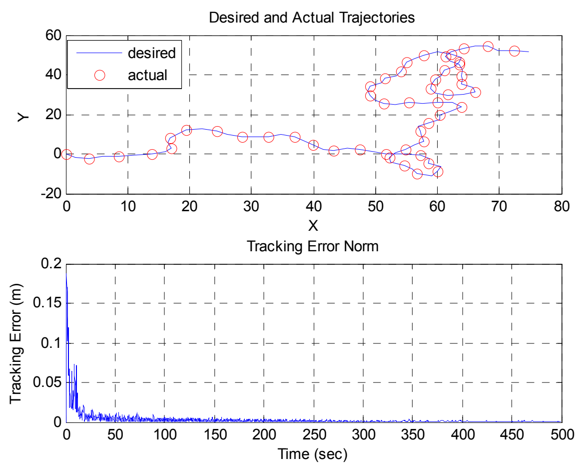

Fig. 3. Tracking Error Time History and Desired State Trajectory

VI. Conclusions

A tracking learning controller for nonlinear discrete–time systems employing a generalized function approximator such as an NLPN was introduced. The developed controller was shown to require minimal modeling information and can be used to learn the inverse of the input-state mapping of the plant while in operation. The stability of the overall system under ideal and realistic conditions was rigorously analyzed. Finally, the application of the learning controller to a nonholonomic underactuated mobile robot was investigated. The simulation results of this application confirmed the theoretical assertions of the paper.

References

[1] N. Sadegh, "A perceptron network for functional identification and control of nonlinear systems," *IEEE Transactions on Neural Networks*, vol. 4, no. 6, pp. 982–989, 1993.

[2] ——, "A nodal link perceptron network with applications to control of a nonholonomic system," *IEEE Transactions on Neural Networks*, vol. 6, no. 6, pp. 1516–1523, 1995.

[3] J. Ghosh and B. Paden, "Iterative learning control for nonlinear nonminimum phase plants with input disturbances," in *Proceedings of the 1999 American Control Conference*, San Diego, CA, 1999, pp. 2584–2589.

[4] N. Sadegh, "Trajectory learning and output feedback control of nonlinear discrete-time systems," in *40th IEEE Conference on Decision and Control*, vol. 4, 2001, pp. 4032–4037.

[5] A. Alleyne and M. Pomykalski, "Control of a class of nonlinear systems subject to periodic exogenous signals," *IEEE Transactions on Control System Technology*, vol. 8, no. 2, pp. 279–287, 2000.

[6] F. Dong and M. Tomizuka, "An iterative learning control design for self-servowriting in hard disk drives using l_1 optimal control," in *American Control Conference*, St. Louis, Missouri, 2009, pp. 240–245.

[7] S. Arimoto, "Mathematical theory of learning with applications to robot control," in *Adaptive and Learning Systems*, K. S. Narendra, Ed. Plenum Publishing, 1986.

[8] N. Sadegh and K. Guglielmo, "Design and implementation of adaptive and repetitive controllers for mechanical manipulators," *IEEE Transactions on Robotics and Automation*, vol. 8, no. 3, pp. 395–400, 1992.

[9] P. Opdenbosch, "Auto-calibration and control applied to electro-hydraulic valves," Ph.D. Dissertation, Georgia Institute of Technology, 2007.

[10] P. Opdenbosch, N. Sadegh, and W. J. Book, "Learning control applied to electro-hydraulic poppet valves," in *American Control Conference*, 2008, pp. 1525–1532.

[11] S. Liu and B. Yao, "On-board system identification of systems with unknown input nonlinearity and system parameters," *ASME Dynamic Systems and Control Division*, vol. 74, pp. 1079–1085, 2005.

[12] A. Isidori, *Nonlinear Control Systems*. Springer-Verlag, 1995.

[13] J. Hauser, S. Sastry, and P. Kokotovic, "Nonlinear control via approximate input-output linearization: The ball and beam example," *IEEE Transactions on Automatic Control*, vol. 37, no. 3, 1992.

[14] A. Levin and K. Narendra, "Control of nonlinear dynamical systems using neural networks–controllability and stabilization," *IEEE Transactions on Neural Networks*, vol. 4, no. 2, pp. 192–206, 1993.

[15] S. Devasi, D. Chen, and B. Paden, "Nonlinear inversion–based output tracking," *IEEE Transactions on Automatic Control*, vol. 41, no. 7, 1996.

[16] J. Park and J. Sandberg, "Universal approximation using radial basis functions network," *Neural Computation*, vol. 3, pp. 246–257, 1991.

[17] N. Sadegh, "Realization of nonlinear systems described by input–output difference equations," in *Proceedings of the 1998 American Control Conference*, Tampa, FL, 1998, pp. 4614–4619.

[18] P. Albertos, "Block multirate input-ouput model for sampled-data control systems," *IEEE Transactions on Automatic Control*, vol. 35, pp. 1085–1088, 1990.

[19] J. Grizzle and P. Kokotovic, "Feedback linearization of sampled-data systems," *IEEE Transactions on Automatic Control*, vol. 33, no. 9, pp. 857–859, 1988.

[20] N. Sadegh, "Minimal realization of nonlinear systems described by input-output difference equations," *IEEE Transactions on Automatic Control*, vol. 46, no. 5, pp. 698–710, 2001.

[21] P. Moraal and J. Grizzle, "Observer design for nonlinear systems with discrete–time measurements," *IEEE Transactions on Automatic Control*, vol. 40, no. 3, pp. 395–404, 1995.

[22] A. Khayatian, *Multirate and Block Methods for Modeling and Control of Pulse Modulated Systems*. Georgia Institute of Technology: Ph. D. Dissertation, 1993.

[23] E. Papadopoulos and M. Misailidis, "On differential drive robot odometry with application to path planning," in *European Control Conference*, Kos, Greece, 2007, pp. 5492–5499.

[24] P. Albertos, "Block multirate input-output model for sampled-data control systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 9, pp. 1085–1088, 1990.

[25] N. Sadegh and B. Maqueira, "Output control of discrete-time nonlinear systems," in *Proceedings of the American Control Conference*, vol. 2, 1994, pp. 2175–2179.

[26] J. Grizzle and P. Moraal, "Newton observers and nonlinear discrete–time control," in *Proceedings of the 29th Conference on Decision and Control*, Honolulu, HI, 1990, pp. 760–767.

[27] A. Levin and K. Narendra, "Control of nonlinear dynamical systems using neural networks–part ii: Observability, identification, and control," *IEEE Transactions on Neural Networks*, vol. 7, no. 1, pp. 30–42, 1996.

[28] J. Marsden, *Elementary Classical Analysis*. San Francisco, CA.: Freeman Publishing, 1974.

[29] K. Narendra and Y. Lin, "Stable discrete adaptive control," *IEEE Transactions on Automatic Control*, vol. 25, pp. 456–461, 1980.

APPENDIX

A. Proof of Lemma 1

We first show that $\left\|\mathbf{\Gamma}_{(k+j,k)}\right\| \leq 1$, $\forall k, j \geq 0$. Each A_k is symmetric, has $N-1$ repeated eigenvalues at 1, and a single eigenvalue λ corresponding to the eigenvector $v = \Phi_k / \|\Phi_k\|$: $\lambda = v^T A_k v = 1 - c_k \|\Phi_k\|^2$. Clearly $|\lambda| < 1$ if $c_k \|\Phi_k\|^2 < 2$, which shows that $\|A_k\| \leq 1$. Thus $\left\|\mathbf{\Gamma}_{(k_1,k_0)}\right\| \leq \|\mathbf{A}_{k+j-1}\| \cdots \|\mathbf{A}_k\| \leq 1$. To prove the main assertion of the Lemma, suppose on the contrary that $\forall \varepsilon > 0$, $\exists k \geq 1$ such that $\left\|\mathbf{\Gamma}_{(k+s,k)}\right\|^2 > 1 - \varepsilon^2$. Consequently, $\exists v_k \in \mathbb{R}^N$, $\|v_k\| = 1$, such that $\left\|\mathbf{\Gamma}_{(k+s,k)} v_k\right\|^2 > 1 - \varepsilon^2$. The inequality

$$\left\|\mathbf{\Gamma}_{(k+s,k)} v_k\right\| \leq \|\mathbf{A}_{k+s-1}\| \cdots \|\mathbf{A}_{k+j}\| \left\|\mathbf{\Gamma}_{(k+j,k)} v_k\right\|$$

combined with $\|A_k\| \leq 1$, $\forall k \geq 0$, implies that $\left\|\mathbf{\Gamma}_{(k+j,k)} v_k\right\|^2 > 1 - \varepsilon^2$, $1 \leq j \leq s$. Expanding $\left\|\mathbf{\Gamma}_{(k+1,k)} v_k\right\|^2 = \left\|v_k - c_k \Phi_k \Phi_k^T v_k\right\|^2 > 1 - \varepsilon^2$ implies that

$$(2 - c_k |\Phi_k|^2) c_k \left|\Phi_k^T v_k\right|^2 < \varepsilon$$

or $\left\|\Phi_k^T v_k\right\| < \gamma\varepsilon$, where $\gamma^2 = \sup_{k \geq 1} (2 - c_k \|\Phi_k\|)^{-1} c_k^{-1} < \infty$. Furthermore, $\left\|\mathbf{\Gamma}_{(k+1,k)} v_k - v_k\right\| < \gamma \bar{c} \bar{\Phi} \varepsilon$ where $\bar{\Phi} := \sup_{\mathbf{x},\mathbf{z} \in \mathcal{X}} \|\Phi(\mathbf{z}, \mathbf{x})\|$. Letting $w_0 = v_k$ and $w_j = \mathbf{\Gamma}_{(k+j,k)} v_k$ we shall use an induction argument to show that $\left\|\Phi_{k+j}^T v_k\right\| < (1 + j\bar{c}\bar{\Phi})\gamma\varepsilon$ and $\|\tilde{w}_{j+1}\| \leq (j+1)\gamma\bar{c}\bar{\Phi}\varepsilon$, $j = 0, \ldots, s-1$, where $\tilde{w}_j := w_j - v_k$. We have already established this for $j = 0$. To prove it in general, suppose that the induction hypothesis is satisfied for $j = 0, \ldots, q-1$, $q < s$. Expanding $\left\|\mathbf{\Gamma}_{(k+q+1,k)} v_k\right\|^2 = \|\mathbf{A}_{k+q} w_q\|^2$, we have

$$\|w_q\|^2 - \gamma^{-2} \left\|\Phi_{k+q}^T w_q\right\|^2 > 1 - \varepsilon^2$$

This inequality combined with $\|w_q\| \leq 1$ and $\left\|\Phi_{k+q}^T w_q\right\| \geq \left\|\Phi_{k+q}^T v_k\right\| - \bar{\Phi} \|\tilde{w}_q\|$ imply that $\left\|\Phi_{k+q}^T w_q\right\| < \gamma\varepsilon$ and that $\left\|\Phi_{k+q}^T v_k\right\| < (1 + q\bar{c}\bar{\Phi}^2)\gamma\varepsilon$. Furthermore,

$$\begin{aligned}
\|\tilde{w}_{q+1}\| &= \|\mathbf{A}_{k+q} w_q - v_k\| = \left\|\tilde{w}_q - c_{k+q}\Phi_{k+q}\Phi_{k+q}^T w_q\right\| \\
&< \|\tilde{w}_q\| + \gamma\bar{c}\bar{\Phi}\varepsilon = (q+1)\gamma\bar{c}\bar{\Phi}\varepsilon
\end{aligned}$$

which completes the induction proof. Thus

$$\begin{aligned}
v_k^T \left(\sum_{\nu=k}^{k+s-1} \Phi_\nu \Phi_\nu^T \right) v_k &= \sum_{\nu=k}^{k+s-1} |\Phi_\nu^T v_k|^2 \\
&< \sum_{j=0}^{s-1} (1 + j\bar{c}\bar{\Phi}^2)\gamma\varepsilon \\
&< s(1 + (s-1)\bar{c}\bar{\Phi}^2/2)\gamma\varepsilon
\end{aligned}$$

Since $\varepsilon > 0$ can be made arbitrarily small, this proves that no $\underline{\lambda} > 0$ exits such that $\sum_{\nu=k}^{k+s-1} \Phi_\nu \Phi_\nu^T \geq \underline{\lambda} I_{N \times N}$, $\forall k \geq 0$, thus contradicting Assumption A_4 and completing the proof.

B. Proof of Theorem 1

First by the continuity of F_k and $\xi_k = F_{k-1,\mu_{k-1}} \circ \cdots \circ F_{0,\mu_0}(\xi_0)$ it follows that there exists $\varepsilon_0 > 0$ and $\varepsilon_1 > 0$ such that if $\|\xi_0\| < \varepsilon_0$ and $\sup_{k \geq 0} \|\mu_k\| < \varepsilon_1$ then $\|\xi_k\| < \bar{\xi}$, $0 \leq k < s$. By the Taylor series expansion

$$\begin{aligned}
\xi_{k+1} &= A_k \xi_k + d_k && (12) \\
d_k &= B_k \mu_k + \tilde{d}_k && (13)
\end{aligned}$$

where $\left\|\tilde{d}_k\right\| = \eta_k \|\mu_k\| + \gamma_k \|\xi_k\|$ and $\eta_k, \gamma_k = O(\|\xi_k\|, \|\mu_k\|)$. Evaluating $\xi_1, \ldots, \xi_s$ recursively yields

$$\xi_s = \mathbf{\Gamma}_{(s,0)} \xi_0 + d_{s-1} + \sum_{k=0}^{s-2} \mathbf{\Gamma}_{(s,k+1)} d_k$$

where $\mathbf{\Gamma}_{(k_1,k_0)}$ is as defined in Lemma 1. Using the hypothesis that $\left\|\mathbf{\Gamma}_{(s,0)}\right\| \leq \sigma$ it follows that

$$\|\xi_s\| \leq \sigma \|\xi_0\| + a_s \sup_{0 \leq k < s} \|d_k\| \qquad (14)$$

where $a_s = (a^s - 1)/(a - 1)$. To establish an upper bound on $\sup_{0 \leq k < s} \|d_k\|$, we may also recursively evaluate $\|\xi_1\|, \ldots,$ and$\|\xi_s\|$ based on

$$\left\|\xi_{k+1}\right\| \leq (a + \gamma) \|\xi_k\| + (b + \eta) \|\mu_k\|$$

implied by (12) with $\gamma = \sup_{0 \leq k < s} \|\gamma_k\|$ and $\eta = \sup_{0 \leq k < s} \|\eta_k\|$ to get

$$\|\xi_k\| \leq (a + \gamma)^s \|\xi_0\| + b_s \sup_{0 \leq k < s} \left\|\mu_j\right\| \qquad (15)$$

where $b_s = \frac{((a+\gamma)^s - 1)(b+\eta)}{a+\gamma-1}$. Taking the norm of both sides of (13) using (15) establishes

$$\sup_{0 \leq k < s} \|d_k\| \leq \gamma(a + \gamma)^s \|\xi_0\| + \bar{b} \sup_{0 \leq k < s} \|\mu_k\|$$

where $\bar{b} = \gamma b_s + b + \eta$. The preceding inequality combined with (14) yields

$$\|\xi_s\| \leq (\sigma + a_s \gamma (a + \gamma)^s) \|\xi_0\| + a_s \bar{b} \sup_{0 \leq k < s} \|\mu_k\|$$

Once more, by the continuity of F_k there exists $\bar{\xi}_0 \leq \varepsilon_0$ and $\bar{\mu} \leq \varepsilon_1$ such that if $\|\xi_0\| < \bar{\xi}_0$ and $\sup_{k \geq 1} \|\mu_k\| < \bar{\mu}$ then $\bar{\sigma} := \sigma + a_s \gamma (a + \gamma)^{s-1} < 1$ and $a_s \bar{b} \bar{\mu} < (1 - \bar{\sigma})\bar{\xi}_0$ implying that $\|\xi_s\| < \bar{\xi}_0$. A simple induction argument also establishes that

$$\begin{aligned}
\left\|\xi_{k+js}\right\| &\leq (a + \gamma)^s \left\|\xi_{js}\right\| + b_s \sup_{js \leq k < js+s} \|\mu_k\| \\
\left\|\xi_{(j+1)s}\right\| &< \bar{\sigma} \left\|\xi_{js}\right\| + a_s \bar{b} \sup_{js \leq k < js+s} \|\mu_k\| < \bar{\xi}_0
\end{aligned}$$

for $j = 1, 2, \ldots$ and $\|\xi_k\| < \bar{\xi}$, $\forall k \geq 0$. Finally, taking the upper limits of the preceding equations yields

$$\overline{\lim_{k \to \infty}} \|\xi_k\| \leq \left[\frac{a_s \bar{b} (a + \gamma)^s}{(1 - \bar{\sigma})} + b_s \right] \overline{\lim_{k \to \infty}} \|\mu_k\|$$

proving that $\overline{\lim}_{k \to \infty} \|\xi_k\| = O(\overline{\lim}_{k \to \infty} \|\mu_k\|)$.