Direct Adaptive NN Control of MIMO Nonlinear Discrete-Time Systems using Discrete Nussbaum Gain

Lianfei Zhai*  Chenguang Yang**  Shuzhi Sam Ge***1  Tianyou Chai* Tong Heng Lee**

* Key Laboratory of Process Industry Automation, Ministry of Education, and Research Center of Automation, Northeastern University, Shenyang, P. R. China 110014  ** Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576

Abstract: In this paper, direct adaptive neural network (NN) control is developed for a class of multi-input and multi-output (MIMO) nonlinear systems in discrete-time. To solve the difficulty of nonaffine appearance of control, implicit function theorem is exploited to assert the existence of an ideal desired feedback control (IDFC). Then, high-order-neural-network (HONN) is employed to approximate the IDFC. Under the assumption that the inverse control gain matrix has an either positive definite or negative definite symmetric part, the obstacle in NN weights tuning for the MIMO systems is transformed to as similar as unknown control direction problem for SISO system. Then, the difficulty in NN weights tuning is overcome by exploiting the discrete Nussbaum gain, which is combined with deadzone method to treat with external disturbance with unknown upper bound. All signals in the closed-loop system are guaranteed to be semi-globally-uniformly-ultimately-bounded (SGUUB). The effectiveness of the proposed control is demonstrated in the simulation.

1. INTRODUCTION

Adaptive neural network (NN) control has drawn ever increasing research in the past two decades due to its universal approximation capability of nonlinear functions. In early works, backpropagation (BP) algorithm was used as major NN learning method and the resultant control requires sufficient off-line training to guarantee that the stability and convergence of the closed-loop system. With the progress of NN control, several excellent adaptive NN control approaches have been proposed for discrete-time nonlinear systems based on Lyapunov’s method and they guarantee the stability of closed-loop systems without the requirement of off-line training (Ge et al. [2001], Jagannathan [2006]).

To overcome the difficulty caused by nonaffine appearance of input in nonaffine nonlinear systems, control design using implicit function theory was first introduced in (Goh [1994], Goh and Lee [1994]) to identify an “inverse” control. In (Goh [1994]), the control design was based on NN trained model, while in (Goh and Lee [1994]), adaptive NN control was proposed. The issues in the application of NN “inverse control” was discussed in (Cabrera and Narendra [1999]). The implicit function based adaptive NN control has been further developed with high order neural network (HONN) for nonaffine pure-feedback system in (Ge et al. [2007a]).

The aforementioned results are restricted in single-input and single-output (SISO) nonlinear systems. For MIMO nonlinear systems, the control problem becomes very difficult to deal with when there exist uncertain parameters and unknown nonlinear functions in the input coupling matrix. As indicated in (Yao and Tomizuka [2001]), sometimes even the presentation of MIMO systems in a meaningful manner is a difficult task. Due to these difficulties, there are relatively fewer results available for the broader class of MIMO nonlinear systems, in comparison with the vast amount of results on SISO nonlinear systems.

In (Jagannathan and Lewis [1996]), the NN control is studied for a class of discrete-time MIMO nonlinear systems with relative degree one and without any interconnections between subsystems. For non-affine MIMO nonlinear systems, NN “inverse control” approach based on off-line training has been proposed in (Adetona et al. [2000]). In (Ge et al. [2004b]), NN control for MIMO system with n triangular subsystems was investigated with backstepping design. However, as mentioned in (Ge et al. [2004a]), how to tune the NN weights for general MIMO systems is still an open problem, especially when there exists unknown strong interconnections between subsystems. To deal with the tuning problem, it is proposed in (Ge et al. [2004a]) to seek an orthogonal matrix to tune the NN weight matrix for adaptive NN control of a general class of unknown discrete-time NARMAX MIMO systems in affine form. However, the existence of the orthogonal matrix required for tuning is not theoretically guaranteed.

In this paper, we investigate adaptive NN control of a class of uncertain discrete-time MIMO NARMAX non-affine
systems. Assuming the inverse control gain matrix has an either positive definite or negative definite symmetric part, the adaptive tuning of NN weights in control of MIMO nonlinear system can be simplified to as similar as that for SISO system with unknown control direction, which is defined as the signs of “control variable” gains in affine systems or signs of partial derivatives over “control variables” in non-affine systems. Based on this observation, we only restrict on the inverse control gain matrix of the system instead of assuming the existence of an orthogonal matrix (Ge et al. [2004a]) for tuning.

Throughout this paper, the following notations are used.

- \( \| \cdot \| \) denotes the Euclidean norm of vectors and induced norm of matrices.
- \( (\cdot)^T \) represents the transpose of a vector or a matrix.
- \( (\cdot)^{-1} \) denotes the inverse of a vector or a matrix.
- \( 0_{[p]} \) stands for \( p \)-dimension zero vector.
- \( (\cdot) \) and \( (\cdot) \) denote the estimate of parameters and estimation error, respectively.

2. SYSTEM DESCRIPTION AND PRELIMINARIES

2.1 System Dynamics

Consider \( p \)-input and \( p \)-output nonlinear discrete-time systems described in the NARMAX model as follows

\[
g(k+\tau) = F(Y(k), U_{k-1}(k), u(k), D_{k-1}(k), d(k)) + d(k+\tau -1) \tag{1}
\]

where \( \tau \) is the system delay, \( F(\cdot) \in \mathbb{R}^p \) is unknown smooth vector valued system function, \( u(k) = [u_1(k), \ldots, u_p(k)]^T \in \mathbb{R}^p \) and \( d(k) = [d_1(k), \ldots, d_p(k)]^T \in \mathbb{R}^p \) denotes the external disturbance which is bounded by an unknown constant \( d_b > 0 \), i.e., \( \|d(k)\| \leq d_b \), and the vectors \( Y(k), U_{k-1}(k), D_{k-1}(k), \) and \( d(k) \) are defined as

\[
Y(k) = [y_1(k), \ldots, y_1(k-n_1+1), y_2(k), \ldots, y_2(k-n_2+1), \ldots, y_p(k), \ldots, y_p(k-n_p+1)]^T
\]

\[
U_{k-1}(k) = [u_1(k-1), \ldots, u_1(k-m_1), u_2(k-1), \ldots, u_2(k-m_2), \ldots, u_p(k-1), \ldots, u_p(k-m_p)]^T
\]

\[
D_{k-1}(k) = [d_1(k-1), \ldots, d_1(k-t_1+1), d_2(k-1), \ldots, d_2(k-t_2+1), \ldots, d_p(k-1), \ldots, d_p(k-t_p+1)]^T
\]

\[
d(k) = [d(k+\tau -2), \ldots, d(k)]^T, \quad \text{if} \quad \tau \geq 2
\]

with \( n_i \) denotes the length of the \( i \)-th inputs, \( m_i \) the length of the \( i \)-th disturbance, \( t_i \) the length of the \( i \)-th disturbance, \( i = 1, \ldots, p \).

**Assumption 1.** The vector valued system function \( F(Y(k), U_{k-1}(k), D_{k-1}(k), d(k)) \) satisfies Lipschitz condition w.r.t. \( D_{k-1}(k) \) and \( d(k) \), i.e., there exists Lipschitz constants \( L_1 \) and \( L_2 \) such that

\[
\|F(Y(k), U_{k-1}(k), u(k), D_{k-1}(k), d(k))
-F(Y(k), U_{k-1}(k), u(k), 0, 0)\| \\
\leq L_1\|D_{k-1}(k)\| + L_2\|d(k)\|
\]

**Assumption 2.** The control gain matrix \( G(k) = \frac{\partial F(\cdot)}{\partial u(k)} \) \( \forall k \geq 0 \), is a full rank matrix, and its inverse, \( G^{-1}(k) \), has an either positive definite or negative definite symmetric part.

\[
G_{12}(k) = \frac{G^{-1}(k)+G^{-T}(k)}{2}
\]

In addition, the eigenvalues of \( G_{12}(k) \) are assumed to be bounded.

**Remark 1.** It should be pointed that matrices \( G(k) \) and \( G^{-1}(k) \) are general real matrices and they are not required to be symmetric.

**Remark 2.** Assumption 2 is quite looser than Assumption 4 in (Ge et al. [2004a]), which requires existence of an orthogonal matrix \( Q(k) \) multiplying \( G^{-1}(k) \) to guarantee the eigenvalues of the product matrix are all positive.

2.2 Discrete Nussbaum Gain

The discrete-time Nussbaum gain \( N(\cdot) \) was firstly proposed in (Lee and Narendra [1986]). It is defined on a discrete sequence \( x(k) \) with \( \Delta x(k) = x(k+1) - x(k) \).

**Lemma 1.** Consider the discrete-time Nussbaum gain \( N(x(k)) \) defined in (Lee and Narendra [1986]). For the summation

\[
S_N'(x(k)) = \sum_{\sigma=0}^{k} g(\sigma)N(x(\sigma))\Delta x(\sigma) \tag{2}
\]

where \( 0 \leq \Delta x(\sigma) \leq \delta_0 \) with \( \Delta x(0) = 0 \), \( x(k) = \sum_{\sigma=0}^{k-1}\Delta x(\sigma) \) and \( g(k) \) is a bounded coefficient satisfying \( g_1 \leq |g(\sigma)| \leq g_2 \) with positive constants \( g_1 \) and \( g_2 \), we have the following conclusions:

(i) If \( x(k) \) increases without bound, then

\[
\sup_{x(k) \geq \delta_0} \frac{1}{x(k)} S_N'(x(k)) = +\infty
\]

\[
\inf_{x(k) \geq \delta_1} \frac{1}{x(k)} S_N'(x(k)) = -\infty \tag{3}
\]

(ii) If \( x(k) \leq \delta_1 \), then \( |S_N'(x(k))| \leq \delta_2 \), where \( \delta_1 \) and \( \delta_2 \) are some positive constants.

**Proof.** See (Ge et al. [2007b]).

2.3 HONN Approximation

The structure of HONN is expressed as the following (Ge et al. [2001]):

\[
\phi(W, z) = W^T S(z) \quad W \quad \text{and} \quad S(z) \in \mathbb{R}^l
\]

\[
S(z) = [s_1(z), s_2(z), \ldots, s_l(z)]^T
\]

\[
s_i(z) = \prod_{j \in I_i} [a(s_j)]^j \tau_i, \quad i = 1, 2, \ldots, l
\]

where \( z \in \mathbb{R}^m \) is the input to HONN, \( l \) is a positive integer and denotes the NN node number, \( \{I_1, I_2, \ldots, I_l\} \) is a collection of \( l \) unordered subsets of \( \{1, 2, \ldots, m\} \), e.g., \( I_1 = \{1, 3, m\} \), \( I_2 = \{2, 4, m\} \), \( d_1(j) \) is non-negative integers (\( d_1(j) \)'s are larger when the order of the function to be approximated is higher), \( W \) is an adjustable synaptic weight vector, and \( s(z) \) is a monotone increasing and differentiable sigmoidal function. In this paper, it is chosen as a hyperbolic tangent function, i.e., \( s(z) = \frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \).

For a smooth function \( \varphi(z) \) over a compact set \( \Omega \subset \mathbb{R}^m \), given a constant real number \( \mu^* > 0 \), if \( \int \) is sufficiently large, according to (Girosi and Poggio [1989]), there exist a set of ideal bounded weights \( W^* \) such that
max |ϕ(z) − φ(W∗, z)| < µ(z) |µ(z)| < µ∗ (6)
From the universal approximation results for neural networks (Gupta and Rao [1994]), it is known that the constant µ∗ can be made small enough by increasing the NN node number.

3. CONTROL DESIGN AND STABILITY ANALYSIS

The control objective is to design a control input u(k), such that the system output tracks the bounded desired trajectory yd(k) = [yd1(k), ..., ydN(k)]T ∈ Rp, while all the closed loop signals remain bounded. Define error vector e(k) = y(k) − yd(k) = [e1(k), e2(k), ..., eN(k)]T. From (1), the error dynamics is

e(k + τ) = F(Y(k), Uk−1(k), u(k), Dk−1(k), ˙d(k)) − yd(k + τ) + d(k + τ − 1) = F(Y(k), Uk−1(k), u(k), 0, 0) − yd(k + τ) + ∆F(k) + d(k + τ − 1) (7)

where

\[ \Delta F(k) = F(Y(k), U_{k-1}(k), u(k), D_{k-1}(k), \dot{d}(k)) - F(Y(k), U_{k-1}(k), u(k), 0, 0) \]

According to the boundedness of disturbance D_{k-1}(k) and \dot{d}(k), and Assumption 1, ∆F(k) is also bounded. From Assumption 2, the control gain matrix G(k) is non-singular, ∀k ≥ 0. According to implicit function theorem, there exists a unique and smooth IDFC \( u^*(k) = \alpha^*(Y(k), U_{k-1}(k), yd(k + τ)) \), where \( \alpha^*(\cdot) \) is an implicit function such that

\[ F(Y(k), U_{k-1}(k), u^*(k), 0, 0) − yd(k + τ) = 0 \] (8)

Assumption 3. Given a bounded output y(k) ∈ Ωy ⊂ Rp, ∀k > 0, where Ωy can be any bounded compact set, there is a corresponding bounded compact set Ωy, such that the desired control u*(k) is within the compact set Ωy.

Remark 3. The desired trajectory is assumed to be achievable, because it is meaningless to drive the system to track an unrealistic trajectory.

Consider employing HONN in Section 2.3 to approximate the IDFC u*(k) as follows

\[ u^*(k) = W^*T S(\bar{z}(k)) + \mu(k) \] (9)

where \( \bar{z}(k) = [Y^T(k), U_{k-1}^T(k), yd^T(k + τ)]^T \) ∈ Ωz ⊂ Rp with \( q = \sum_{i=1}^{p} n_i + m_i + 1 \) and \( \mu(k) \) is the bounded NN approximation error vector satisfying \( \|\mu(k)\| \leq \mu^* \), which can be reduced by increasing the number of NN nodes. Then the adaptive NN control u(k) is constructed as

\[ u(k) = \hat{W}^T(k) S(\bar{z}(k)) \] (10)

where \( \hat{W}(k) \) ∈ Rl×q and S(\( \bar{z}(k) \)) ∈ Rl. The NN weight adaptation law is given as

\[ \dot{W}(k) = \hat{W}(k − τ) − γ N(x(k)) S(\bar{z}(k − τ)) a(k) e^T(k)/D(k) \] (11)

\[ \Delta \dot{x}(k) = \alpha(k) \gamma e^T(k)/D(k), \dot{x}(0) = 0 \] (12)

\[ D(k) = (1 + |N(x(k))|^2)(1 + |S(\bar{z}(k − τ))|)|e(k)|^2 \] (13)

\[ a(k) = \begin{cases} 1, & \text{if } |e(k)|/(1 + |N(x(k))|) > \lambda \\ 0, & \text{otherwise} \end{cases} \] (14)

where γ > 0 and λ > 0 can be arbitrary positive constants, and \( N(\cdot) \) is discrete Nussbaurn gain defined in Secon 2.2.

Remark 4. To deal with the disturbance and the neural network approximation error, dead zone (14) is introduced into the neural network weight adaptation law (11) to realize robust adaptive control. Although the disturbance is assumed to be bounded, we do not need to know the exact boundary of disturbance.

Theorem 1. Consider the closed-loop system consisting of system (1), controller (10), and neural network weights adaptation law (11)-(14). All signals in the closed-loop system are SGUUB, the discrete Nussbaurn gain N(x(k)) will converge to a constant ultimately, and the tracking error satisfies \( \lim_{k \to \infty} \|e(k)\| < CL \), where \( C = \lim_{k \to \infty} (1 + |N(x(k))|) \).

Proof. The proof is proceed in two parts: Firstly, we assume inputs and outputs are within Ωz such that NN approximation holds; Secondly, given any initial condition, we show that there exists a determined compact set such that if initially the NN approximation range covers this set then the inputs and outputs are guaranteed to be within Ωz without priori assumption in the first step.

Using mean value theorem, (7) can be written as

\[ e(k + τ) = F(Y(k), U_{k-1}(k), y^*(k), 0, 0) − yd(k + τ) + ∆F(k) + G_ξ(ξ) [u(k) − \bar{u}(k)] + d(k + τ − 1) \] (15)

where \( G_ξ(ξ) = \frac{∂F(ξ)}{∂u(k)} \), \( u_ξ(k) \) is a point of line \( L(u(k), \bar{u}(k)) = \{ξ | ξ = \bar{u}(k) + (1 − θ)u^*(k), 0 ≤ θ ≤ 1\} \). Considering (8)-(10) and (15), we obtain

\[ e(k + τ) = G_ξ(ξ) [\hat{W}^T(k) S(\bar{z}(k)) − \mu(k)] + ∆F(k) + d(k + τ − 1) \] (16)

where \( \hat{W}(k) = \hat{W}(k) − W^* \) is the NN weights estimation error.

According to Assumption 2, there exist two positive constants \( \bar{g} \) and \( g \) such that \( gI \leq \frac{1}{2} (G_ξ^T(k) + G_ξ^T(k)) \leq \bar{g}I \) or \( −\bar{g}I \leq \frac{1}{2} (G_ξ^T(k) + G_ξ^T(k)) \leq −\bar{g}I \), where I is the identity matrix. It implies there exists a sequence \( g(k) \) satisfying \( \bar{g} \leq |g(k)| \leq \bar{g} \) such that

\[ e^T(k)G_ξ^T(k) e(k) = e^T(k)G_ξ^T(k) [G_ξ^T(k) G_ξ(k) + G_ξ^T(k) G_ξ(k)] e(k) \]

\[ = g(k) e^T(k) e(k) \] (17)

From (16), we have

\[ \hat{W}^T(k − τ) S(\bar{z}(k − τ)) = G_ξ^T(k − τ) e(k) + d^*(k + 1) \] (18)

where \( d^*(k + 1) = −G_ξ^T(k − τ) [\Delta F(k − τ) + d(k − 1)] + \mu(k) \). According to the boundedness of \( d(k) \), \( \Delta F(k − τ) \) and \( \mu(k) \), and Assumption 2, \( d^*(k) \) is bounded, i.e., \( \|d^*(k − 1)\| \leq d^*_1 \), where \( d^*_1 \) an unknown constant.

Choose a positive definite \( V(k) = \sum_{j=1}^r |\hat{W}^T(k − τ + j) (k − τ + j)\hat{W}(k − τ + j) \} \). Considering (18) and (17), we have

6510
\[
\begin{align*}
&\text{tr}\{2a(k)\gamma N(x(k)) \tilde{W}^T(k - \tau)S(\bar{z}(k - \tau))e^T(k)D(k)\} \\
&= 2a(k)\gamma N(x(k)) \left[ g(k)e^T(k)e(k)D(k) + e^T(k)d^*(k - 1) \right]
\end{align*}
\]

Then, the difference of \(V(k)\) along (18) is
\[
\Delta V(k) = V(k) - V(k - 1) = \text{tr}\{\tilde{W}^T(k)\tilde{W}(k) - \tilde{W}^T(k - \tau)\tilde{W}(k - \tau)\} = \text{tr}\{[\tilde{W}^T(k) - \tilde{W}(k - \tau)][\tilde{W}(k) - \tilde{W}(k - \tau)]\} + 2\tilde{W}^T(k - \tau)[\tilde{W}(k - \tau) - \tilde{W}(k - \tau)]
\]
\[
= a(k)\gamma^2 N^2(x(k)) S^T(\bar{z}(k - \tau))S(\bar{z}(k - \tau))e^T(k)e(k)D^2(k)
\]

Then the boundedness of \(y(k)\) is obvious. The boundedness of \(\bar{W}(k)\) leads to the boundedness of \(u(k)\).

For discrete-time system, the boundedness of \(y(k)\) and \(u(k)\) implies there is a largest bounding set depending on initial condition such that it includes \(y(k)\) and \(u(k)\), \(\forall k > 0\). If initially the NN approximate range \(\Omega_z\) is constructed to cover this set, then NN approximation will hold \(\forall k > 0\), such that the priori assumption the NN approximation range is large enough can be replaced by that NN approximation range covers a specified set depending on initial condition. According to the definition of SGUUB (given any initial condition, there is a corresponding control that can guarantee the closed-loop stability), the proof is completed.

4. SIMULATION

The following NARMAX model is used for simulation.
\[
y_1(k + 2) = \pm 0.2\sin(u_1(k)) \pm u_1(k) + d_1(k) + 0.6\cos(y_2(k - 1)) + y_1(k - 1) + 1.2u_2(k) + y_2(k - 2) \\
y_2(k + 2) = \pm \cos(y_2(k)) \pm 0.5u_2(k) + d_2(k) + 1y_2(k - 1) + 1.6\sin(y_1(k))u_1(k - 1)
\]

where \(d_1(k) = 0.01\cos(0.1k)\) and \(d_2(k) = 0.01\sin(0.1k)\) are disturbance.

The control objective is to make the outputs \(y_1(k)\) and \(y_2(k)\) track the desired reference trajectories \(y_1(k) = 0.5 + 0.25\sin(0.25\pi T k)\) and \(y_2(k) = 0.5 + 0.25\sin(0.25\pi T k)\) and \(y_2(k) = 0.5 + 0.25\sin(0.25\pi T k)\), with \(T = 0.01\), respectively. The initial system states are \(Y(i) = [0, 0]^T\), \(i = -1, 0\), \(U(0) = [0, 0]^T\) and the initial NN weights estimates \(\bar{W}(0)\) are zero matrix and the initial regression function \(S(0)\) is randomly chosen.

Firstly, we choose \(\pm\) to be \(\pm\) in (22), such that inverse control gain matrix \(G^{-1}(k)\) has a positive definite symmetric part. The simulation results are shown in Figures 1-3. The tuning factor and the threshold value are chosen as \(\gamma = 0.95\) and \(\lambda = 0.001\).

Secondly, let us choose \(\pm\) to be \(\pm\) in (22) such that \(G^{-1}(k)\) has a negative symmetric part. Using the same control law with same parameters, the simulation results are shown in Figures 4-6. The simulation results demonstrate the proposed NN control works properly in both cases.

It is noted in Figure 1 that \(N(x(k))\) always keeps positive while in Figure 4 it turns to negative and remains so. This is because \(g(k)N(x(k))\) must be positive to make \(\Delta V(k)\) negative in (21).

5. CONCLUSION

Under the assumption on the inverse control gain matrix, direct adaptive neural network control has then been proposed for a class of MIMO discrete-time nonlinear systems with non-affine appearance of controls. Implicit function theory has been used to assert the existence of IDFC and discrete Nussbaum gain was introduced in the
NN adaptation law. The proposed NN control guarantees SGUUB stability of the closed-loop system.

Fig. 1. Discrete Nussbaum gain $N(k)$ and $x(k)$

Fig. 2. Control $u_1(k)$, $u_2(k)$ and norm of NN weights

Fig. 3. Output tracking

Fig. 4. Discrete Nussbaum gain $N(k)$ and $x(k)$

Fig. 5. Control $u_1(k)$, $u_2(k)$ and norm of NN weights

Fig. 6. Output tracking

REFERENCES


