Nonlinear systems identification using dynamic multi-time scale neural networks

Xuan Han, Wen-Fang Xie*, Zhijun Fu, Weidong Luo

Concordia University, Mechanical & Industrial Engineering, 1455 De Maisonneuve W., Montreal, QC, Canada H3G 1M8

A R T I C L E   I N F O

Article history:
Received 22 November 2010
Received in revised form 5 May 2011
Accepted 5 June 2011
Communicated by J. Zhang
Available online 25 June 2011

Keywords:
Dynamic multi-time scale neural networks
Nonlinear systems
On-line identification
Neural network identifiers

A B S T R A C T

In this paper, two Neural Network (NN) identifiers are proposed for nonlinear systems identification via dynamic neural networks with different time scales including both fast and slow phenomena. The first NN identifier uses the output signals from the actual system for the system identification. The on-line update laws for dynamic neural networks have been developed using the Lyapunov function and singularly perturbed techniques. In the second NN identifier, all the output signals from nonlinear system are replaced with the state variables of the neuron networks. The on-line identification algorithm with dead-zone function is proposed to improve nonlinear system identification performance. Compared with other dynamic neural network identification methods, the proposed identification methods exhibit improved identification performance. Three examples are given to demonstrate the effectiveness of the theoretical results.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Numerous systems in the industrial fields demonstrate non-linearities and uncertainties, which can be considered as partially or totally black-box. Dynamic neural networks have been applied in system identification and control for those systems for many years and due to the fast adaptation and superb learning capability, they have transcendent advantages compared to the traditional neural network methods [2,4–6,21].

A wide class of nonlinear physical systems contains slow and fast dynamic processes that occur at different moments. Recent research results show that neural networks are very effective for modeling the complex nonlinear systems with different time scales when we have incomplete model information, or even when we consider the plant as a black-box [2,3,25]. Dynamic neural networks with different time scales can model the dynamics of the short-term memory of neural activity levels and the long-term memory on the dynamics of unsupervised synaptic modifications [1]. Different methods have been applied in this domain. The stability of equilibrium of competitive neural network with short- and long-term memory was analyzed in [10] by a quadratic-type Lyapunov function. Refs. [9,11,12] presented new methods of analyzing the dynamics of a system with different time scales based on the theory of flow invariance. The K-monotone system theory was used for analyzing the dynamics of a competitive neural system with different time scales in [13].

The past decade has witnessed great activities in stability analysis, identification and control with continuous time dynamic multi-time scale neural networks [8,14,15,17,18,20]. In [8], Sandoval et al. developed new stability conditions using the Lyapunov function and singularly perturbed technique. In [7], the passivity-based approach was used to derive stability conditions for dynamic neural networks with different time scales. The passivity approach was used to prove that a gradient descent algorithm for weight adjustment was stable and robust to any bounded uncertainties, including the optimal network approximation error [14]. Several papers proposed an adaptive nonlinear identification and trajectory [18] or velocity tracking [15] via dynamic neural networks without considering the multiple time scales. Although dynamic neural networks have been used for nonlinear system identification and control in [14,15,18,24], two issues are worth mentioning with regard to these researches. First, in the dynamic neural network model [18], a strong assumption about the linear part matrices A and B was posed as a known Hurwitz matrix, which was sometimes unrealistic for the black-box nonlinear system identification. Second, the adaptive control and identification for dynamic systems with different time scales via multiple time scale neural networks has not yet been established in the literature.

In this paper, we consider the nonlinear system with multiple time scales and model it via a continuous time dynamic neural network with different time scales including both fast and slow phenomena. Two NN identifiers are proposed for nonlinear systems identification. In the first NN identifier, the structure of the neural identifier depends on the output signals from the actual system. This may risk the stability of the neural network because it is related to that of the real system. In order to get rid of this flaw and also simplify the identification scheme, we replace all the...
output signals from nonlinear system with the state variables of the neuron networks in the construction of the second NN identifier. For two NN identifiers, the Lyapunov function and singularly perturbed techniques are used to develop on-line update laws for both dynamic and neural network weights and the linear part matrices A and B. We also determine new stability conditions for identification error by means of Lyapunov-like analysis, which are inspired from [14,15]. To the best of our knowledge, it is the first time that the identification matrix is proposed for the neural networks model with multi-time scales to provide adaptability and accuracy of nonlinear system identification. In addition, new dead-zone indicators are introduced in the on-line update laws to prevent the weights of neural network from drifting when the modeling error presents for the second NN identifier. Compared with the other dynamic neural network identification methods, the two proposed NN identifiers with on-line update laws exhibit improved identification performance. Three examples are also given to demonstrate the effectiveness of the theoretical results.

The paper is organized as follows. Section 2 describes the first structure of dynamic neural networks with different time scales and the identification algorithm. Section 3 introduces the improved NN identifier with updating laws. The simulation results and the conclusion of this paper are presented in Sections 4 and 5, respectively.

2. Identification algorithm

In this section we consider the problem of identifying a class of singular perturbation [16] nonlinear systems with two different time scales described by

\[
\begin{align*}
\dot{x} &= f(x,y,u,t), \\
\dot{y} &= f_y(x,y,u),
\end{align*}
\]

where \(x \in \mathbb{R}^n\) and \(y \in \mathbb{R}^m\) are slow and fast state variables, respectively, \(u \in \mathbb{R}^p\) is the control input vector and \(\varepsilon > 0\) is a small parameter.

In order to identify the nonlinear dynamic system (1), we employ the dynamic neural networks with two time scales:

\[
\begin{align*}
\dot{x}_{nn} &= A_{nn}x_{nn} + W_{1}\sigma_1(x,y) + W_{2}\phi_2(x,y)U \\
\dot{y} &= f_y(x_{nn},y_{nn})
\end{align*}
\]

where \(x_{nn} \in \mathbb{R}^n\) and \(y_{nn} \in \mathbb{R}^m\) are the slow and fast state variables of neural networks, respectively, \(W_{1,2} \in \mathbb{R}^{n \times 2n}\) and \(W_{3,4} \in \mathbb{R}^{m \times 2n}\) are the weights in the output layers, \(\sigma_1(x,y) = [\sigma_1(x_1), \ldots, \sigma_1(x_n), \sigma_2(y_1), \ldots, \sigma_2(y_m)]^T \in \mathbb{R}^{2n}\) \((k = 1,2)\), \(\phi_2(x,y) = [\phi_2(x_1), \ldots, \phi_2(x_n), \phi_2(y_1), \ldots, \phi_2(y_m)]^T \in \mathbb{R}^{2n}\).

![Fig. 1. Structure of the first NN identifier.](image)

From (2) and (3), we can obtain the error dynamics equations

\[
\begin{align*}
\dot{x}_e &= x_{nn} - x, \\
\dot{y}_e &= y_{nn} - y.
\end{align*}
\]

Assumption 1. For the above matrices \(P_x, P_y, Q_x, Q_y\) there exist positive definite matrices \(A_x, A_y\) such that:

\[
\begin{align*}
\lambda_{min}(Q_x) &\geq \lambda_{max}(P_x A_x P_x), \\
\lambda_{min}(Q_y) &\geq \lambda_{max}(P_y A_y P_y).
\end{align*}
\]

Differentiating (5) and using (4) yield

\[
\begin{align*}
\dot{V}_{1x} &= -\Delta x^T Q_x \Delta x + 2\Delta x^T P_x W_{1} \sigma_1(x,y) + 2\Delta y^T P_y W_{2} \phi_2(x,y)U + 2\Delta x^T P_\Delta f_x \\
&+ 2\Delta x^T P_\Delta W_{1} + 2\Delta y^T P_y f_{\Delta y} - 2\Delta x^T P_\Delta W_{2}, \\
\dot{V}_{1y} &= -2\Delta y^T Q_y \Delta y + (1/\varepsilon) 2\Delta y^T P_y f_{\Delta y} + (1/\varepsilon) 2\Delta y^T P_\Delta W_{3} \sigma_2(x,y) + 2\Delta y^T P_\Delta f_y \\
&+ 2\Delta x^T P_{\Delta y} + 2\Delta y^T P_{\Delta y} f_{\Delta y} - 2\Delta y^T P_\Delta W_{4}.
\end{align*}
\]

Theorem 1. Consider the identification model (3) for (1). If the model errors and disturbances \(\Delta f_x\) and \(\Delta f_y\) are bounded and with
Assumption 1, the updating laws
\[ \begin{align*}
  &\Delta x = -\Delta x_T \quad \text{for dynamic multi-time scales neural network.} \\
  &W_1 = -\Delta x_T (x,y) \quad W_2 = -\Delta x_T (y) \quad W_3 = -\Delta x_T (y) \\
  &W_4 = -\Delta x_T (y) \quad W_5 = -\Delta x_T (y) \\
  \end{align*} \]
(8)
can guarantee the following stability properties:
\[ \Delta x, \Delta y \in L_\infty, \quad W_{1,2,3,4,5} \in L_\infty, \quad A, B \in L_\infty. \]

**Proof.** Since the neural network's weights are adjusted as (8) and the derivatives of the neural network weights and matrices satisfy the following $W_1 = W_2, W_3 = W_3, W_4 = W_4, A = \hat{A}, B = \hat{B}$, from (7), $V_{1,x}$ and $V_{1,y}$ become
\[ \begin{align*}
  V_{1,x} &= -\Delta x^2 Q_1 \Delta x + 2\Delta x^2 P_1 \Delta x \\
  V_{1,y} &= -(1/\epsilon)\Delta y^2 Q_2 \Delta y + (1/\epsilon)\Delta y^2 P_2 \Delta y \\
  \end{align*} \]
(9)

Since $\Delta x^2 P_1 \Delta x$ is scalar, using the following matrix inequality [28]:
\[ X^T Y + \lambda X = X A^{-1} X + X^T Y \]
(10)
where $X, Y \in \mathbb{R}^{n \times k}$ are any matrices, $A = A^T \in \mathbb{R}^{n \times n}$ is any positive definite matrix, we obtain:
\[ 2\Delta x^2 P_1 \Delta x \leq (1/\epsilon)2\Delta x^2 P_1 \Delta x \leq (1/\epsilon)\Delta x^2 P_1 \Delta x + (1/\epsilon)\Delta x^2 P_1 \Delta x \\
\]
Using Assumption 1, Eq. (9) can be represented as
\[ \begin{align*}
  V_{1,x} &= -\Delta x^2 Q_1 \Delta x + 2\Delta x^2 P_1 \Delta x \\
  &\leq -(1/\epsilon)\Delta x^2 Q_1 \Delta x + (1/\epsilon)2\Delta x^2 P_1 \Delta x \\
  V_{1,y} &= -(1/\epsilon)\Delta y^2 Q_2 \Delta y + (1/\epsilon)\Delta y^2 P_2 \Delta y \\
  &\leq -(1/\epsilon)\Delta y^2 Q_2 \Delta y + (1/\epsilon)\Delta y^2 P_2 \Delta y \\
\end{align*} \]
where $\Delta x = (\Delta x_1, \ldots, \Delta x_k)^T, \Delta y = (\Delta y_1, \ldots, \Delta y_k)^T, \beta_1(f_j) = \max(0, f_j), \beta_2(f_j) = \max(0, f_j)$. 

Since $\Delta x, \Delta y$ are ISS-lyapunov functions, $V_{1,x}$ and $V_{1,y}$ are ISS-lyapunov functions. Using Theorem 1 in [20], the dynamics of the identification error (4) is input to state stable, which implies $\Delta x, \Delta y \in L_\infty, \Delta x, \Delta y \in L_\infty$. Such input to state stability means that the behavior of neural network identification should remain bounded when input bounded [14].

**Remark 1.** When $\epsilon$ is very close to zero, both $W_2$ and $W_4$ exhibit a high-gain behavior, causing the instability of identification algorithm. The lyapunov function (6) can be multiplied by any positive constant $\alpha$, i.e., $\beta(\alpha P_2) + \alpha y^2 + \alpha - \lambda Q_2$, the learning gains of $W_2$ and $W_4$ become $1/\epsilon \alpha P_2$. We can also choose $\alpha$ as a very small number, which is close to zero so that the learning gain $(1/\epsilon) \alpha P_2$ does not become too large.

The structure of the identification scheme is illustrated in Fig. 2.

**3. Improved system identification**

In the proposed identification scheme, we use the signals from the actual system in the neuron networks to identify the nonlinear system (1). This may simplify the identification and the control procedure, but the updating laws depend on the actual signals of the nonlinear system. Also, this may risk the stability of the neural network because it is related to the output of the real system. In order to conquer this flaw and also simplify the identification scheme, we replace all the output signals from nonlinear system with the state variables of the neural networks in the construction of NN identifier. In the on-line update laws, we introduce new dead-zone indicators to prevent the weights of neural network from drifting when the modeling error presents for dynamic multi-time scales neural network.

Consider the nonlinear system (1). In order to identify the system, we employ the dynamic neural networks with two time scales:
\[ x_m = Ax_m + W_1 \sigma_1 (V_1[km,ym]^T) + W_2 \phi_1 (V_2[km,ym]^T) (U) \]
\[ \hat{x}_m = By_m + W_3 \sigma_2 (V_3[km,ym]^T) + W_4 \phi_2 (V_4[km,ym]^T) (U) \]
\[ x_m \in \mathbb{R}^{n_1} \text{ and } y_m \in \mathbb{R}^{n_2} \text{ are the slow and fast state variables of neural networks.} \]
\[ V_{1,2,3,4} \in \mathbb{R}^{n_1 \times n_2} \text{ and } W_{1,2,3,4} \in \mathbb{R}^{n_2 \times n_2} \text{ are the weights in the hidden layer and} \]
\[ \sigma_1 (km,ym) = [\sigma_1 (km,ym), \ldots, \sigma_1 (km,yn)] \text{ and } \sigma_2 (km,ym) = [\sigma_2 (km,yn), \ldots, \sigma_2 (km,yn)] \]
\[ (k = 1, 2, \ldots), U = [u_1, u_2, u_3, \ldots] \text{ is the control input vector, } \gamma : \mathbb{R}^{n_2} \rightarrow \mathbb{R}^{n_2} \text{ is a differentiable input–output mapping function.} \]
A relation $\sigma_2 (\cdot)$ and $\phi_2 (\cdot)$ are unknown matrices for the linear part of neural networks and the parameter $\epsilon$ is an unknown small positive number. The activation functions $\sigma_2 (\cdot)$ and $\phi_2 (\cdot)$ are still kept as sigmoid function.

In order to simplify the analysis process, we consider the simplest structure, which means
\[ p = q = n, \quad V_1 = V_2 = I, \quad \phi_1 (\cdot) = \phi_2 (\cdot) = I \]
\[ \hat{x}_m = Ax_m + W_1 \sigma_1 (x_m,ym)^T + W_2 \gamma (U) \]
\[ \hat{y}_m = By_m + W_2 \sigma_2 (x_m,ym)^T + W_4 \gamma (U) \]
\[ x_m \in \mathbb{R}^{n_1} \text{ and } y_m \in \mathbb{R}^{n_2} \text{ are the slow and fast state variables of neural networks.} \]
\[ W_{1,2,3,4} \in \mathbb{R}^{n_2 \times n_2} \text{ are the weights in the output layers,} \]
\[ \sigma_1 (km,ym) = [\sigma_1 (km,yn), \ldots, \sigma_1 (km,yn), \ldots, \sigma_1 (km,yn)] \in \mathbb{R}^{n_2 \times n_2} \]
\[ (k = 1, 2, \ldots), U = [u_1, u_2, u_3, \ldots] \text{ is the control input vector, } A \in \mathbb{R}^{n_1 \times n_1}, B \in \mathbb{R}^{n_1 \times n_2} \text{ are unknown matrices for the linear part of neural networks.} \]

Generally speaking, when the dynamic neural network (12) does not match the given nonlinear system (1) exactly, the
nonlinear system can be represented as
\[ \dot{x} = A'x + W_1 \sigma_1(x,y) + W_2 \gamma(U) - D \delta, \]
where \( W_1, W_2, W_3, W_4 \) are unknown nominal constant matrices, the vector functions \( \Delta \delta \) and \( \Delta D \) can be regarded as modeling error and disturbances, which are assumed to be bounded as \( || \Delta \delta || \leq \Delta \delta, || \Delta D || \leq \Delta D \) and \( A' \) and \( B' \) are unknown nominal constant Hurwitz matrices.

**Assumption 2.** The nominal values \( W_1, W_2, W_3, W_4 \) are bounded as
\[ W_1 A_1 W_1^T \leq W_2, \quad W_1 A_2 W_1^T \leq W_3, \quad W_1 A_3 W_1^T \leq W_4, \]
where \( A_1, A_2, A_3, A_4 \) are any positive definite symmetric matrices, \( W_1, W_2, W_3, W_4 \) are prior known matrices bounds.

As we assumed in Section 2, the state variables in system (1) are completely measurable and the state variables of the plant is equal to that of the neural networks (12). The identification errors are defined as
\[ \Delta x = x - x_{\text{nom}}, \quad \Delta y = y - y_{\text{nom}}. \]

From (12) and (13), the error dynamic equations become
\[ \dot{\Delta x} = A' \Delta x + \dot{\Delta} x_{\text{nom}} + W_1 \dot{\Delta} \sigma_1 + W_1 \dot{x}_{\text{nom}}(x_{\text{nom}},y_{\text{nom}}) + W_2 \dot{\gamma}(U) + \dot{\Delta} D, \]
\[ \dot{\Delta y} = B' \Delta y, \quad \dot{\Delta} y_{\text{nom}} + W_3 \dot{\Delta} \sigma_2 + W_3 \dot{x}_{\text{nom}}(x_{\text{nom}},y_{\text{nom}}) + W_4 \dot{\gamma}(U) + \dot{\Delta} D \]
where \( \dot{\Delta} x_{\text{nom}} = W_1 \dot{x}_{\text{nom}} - \dot{W}_1 \dot{x}_{\text{nom}}, \quad \dot{\Delta} y_{\text{nom}} = W_3 \dot{x}_{\text{nom}} - \dot{W}_3 \dot{x}_{\text{nom}} \)
\( \dot{\Delta} D \) is assumed to be bounded as \( \| \dot{\Delta} D \| \leq \Delta D \).

**Assumption 3.** The difference of the activation function \( \dot{\sigma}(\cdot) \), which is \( \dot{\sigma}(x,y) = \sigma_k(x,y) - \sigma_k(x_{\text{nom}},y_{\text{nom}}) \), satisfies the generalized Lipshitz condition
\[ \dot{\sigma}_x A_1 \dot{\sigma} < \begin{bmatrix} \dot{Ax}^T & 0 \\ D_1 & 0 \\ D_1 & D_1 \end{bmatrix} \Delta x + \begin{bmatrix} \dot{Ax}^T & 0 \\ D_1 & D_1 \\ D_1 & D_1 \end{bmatrix} \Delta y \]
\[ \dot{\sigma}_y A_2 \dot{\sigma} < \begin{bmatrix} \dot{Ay}^T & 0 \\ D_2 & 0 \\ D_2 & D_2 \end{bmatrix} \Delta x + \begin{bmatrix} \dot{Ay}^T & 0 \\ D_2 & 0 \\ D_2 & D_2 \end{bmatrix} \Delta y \]
where \( \dot{\sigma}(x,y) = \dot{\sigma}(x_1) \ldots \dot{\sigma}(x_k), \quad \dot{\sigma}(y_1) \ldots \dot{\sigma}(y_n) \) is \( \mathbb{R}^{2m} \) \((k = 1,2), \quad D_1 = D_1^T > 0 \) and \( D_2 = D_2^T > 0 \) are known normalizing matrices.
Proof. The Lyapunov synthesis method is used to derive the stable adaptive laws. Consider the Lyapunov function candidate:

\[
V_2 = V_{2x} + V_{2y}
\]

\[
V_{2x} = [P_x^{1/2} \Delta x] - H_x^2 + \frac{1}{K_x} tr(W_x^T P_x W_x) + \frac{1}{K_y} tr(W_y^T P_y W_y)
\]

\[
V_{2y} = [P_y^{1/2} \Delta y] - H_y^2 + \frac{1}{K_x} tr(W_x^T P_x W_x) + \frac{1}{K_y} tr(W_y^T P_y W_y)
\]

(26)

Differentiating (26) and following Remark 2 yield

\[
V_{2x} = 2[||P_x^{1/2} \Delta x|| - H_x^2] + \frac{1}{K_x} tr(A_x P_x A_x) + \frac{1}{K_y} tr(W_x^T P_x W_x) + \frac{1}{K_y} tr(W_y^T P_y W_y)
\]

\[
V_{2y} = 2[||P_y^{1/2} \Delta y|| - H_y^2] + \frac{1}{K_x} tr(B_y P_y B_y) + \frac{1}{K_y} tr(W_x^T P_x W_x) + \frac{1}{K_y} tr(W_y^T P_y W_y)
\]

(27)

Since the neural network's weights are adjusted as in (24), the derivatives of the neural network weights and matrices satisfy

(28)

Using the matrix inequality (10) and Assumptions 2 and 3, one obtains

\[
2 \Delta x^T P_x W_x \Delta y \leq \Delta x^T P_x W_x \Delta x + \Delta y^T D_1 \Delta y
\]

(29)

and

\[
2 \Delta x^T P_x A_2 \Delta y \leq \Delta x^T P_x W_x \Delta y + \Delta y^T D_2 \Delta y
\]

(30)

Hence, from (28) one has

\[
V_{2x} = S_x \Delta x^T [A_x^T P_x + A_x^T P_x W_x + \Delta x^T D_1 \Delta y]
\]

\[
+ S_y \Delta y^T D_2 \Delta y + S_x A_2 \Delta y
\]

(31)

Case (a): Both the identification errors are larger than the thresholds. (i.e., $S_x > 0$, $S_y > 0$). One has

\[
V_{2x} + V_{2y} = S_x \Delta x^T [A_x^T P_x + A_x^T P_x W_x + \Delta x^T D_1 \Delta y + D_1]\n
+ (1/e) S_y \Delta y^T D_2 \Delta y + S_x A_2 \Delta y
\]

\[
+ (1/e) S_y A_2 \Delta y^T D_2 \Delta y + (1/e) S_y A_2 \Delta y
\]

(32)

One can obtain

\[
V_{2x} = V_{2x} + V_{2y} \leq -S_x (\Delta x^T Q_x \Delta x - \Delta y^T A_2 \Delta y) - (1/e) S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- (1/e) S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

(33)

Case (c): The identification error of $y$ is smaller than the threshold ($S_y = 0$, $S_x > 0$). From (31) one has $[P_x^{1/2} \Delta y] \leq H_y$ and $V_{2y} = 0$. One can obtain

\[
V_{2x} = -S_x (\Delta x^T Q_x \Delta x - \Delta y^T A_2 \Delta y - \Delta y^T D_1 \Delta y)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

(35)

Case (d): The identification error of $x$ is smaller than the threshold ($S_x = 0$, $S_y > 0$). From (31) one has $[P_y^{1/2} \Delta x] \leq H_x$ and $V_{2x} = 0$. One can obtain

\[
V_{2y} = -(1/e) S_y (\Delta x^T Q_y \Delta y - \Delta x^T A_2 \Delta y - \Delta x^T D_2 \Delta x)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

(36)

One can obtain

\[
V_{2y} = -(1/e) S_y (\Delta x^T Q_y \Delta y - \Delta x^T A_2 \Delta y - \Delta x^T D_2 \Delta x)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

(37)

Case (d): Both the identification errors are smaller than the thresholds ($S_x = 0$, $S_y = 0$). One has $[P_x^{1/2} \Delta x] \leq H_x$, $[P_y^{1/2} \Delta y] \leq H_y$ and $V_{2x} = 0$. In Case (a), one has

\[
V_{2x} = S_x \Delta x^T [A_x^T P_x + A_x^T P_x W_x + \Delta x^T D_1 \Delta y + D_1]
\]

\[
+ (1/e) S_y \Delta y^T D_2 \Delta y + S_x A_2 \Delta y
\]

\[
+ (1/e) S_y A_2 \Delta y^T D_2 \Delta y + (1/e) S_y A_2 \Delta y
\]

\[
- (1/e) S_y A_2 \Delta y^T D_2 \Delta y - (1/e) S_y A_2 \Delta y
\]

(32)

One can obtain

\[
V_{2x} = V_{2x} + V_{2y} \leq -S_x (\Delta x^T Q_x \Delta x - \Delta y^T A_2 \Delta y - \Delta y^T D_1 \Delta y)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

(33)

Case (b): The identification error of $y$ is smaller than the threshold ($S_x > 0$, $S_y = 0$). From (31) one has $[P_y^{1/2} \Delta x] \leq H_x$ and $V_{2y} = 0$. One can obtain

\[
V_{2x} = -S_x (\Delta x^T Q_x \Delta x - \Delta y^T A_2 \Delta y - \Delta y^T D_1 \Delta y)
\]

\[
- S_y (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

\[
- S_x (\Delta y^T Q_y \Delta y - \Delta y^T A_2 \Delta y)
\]

(35)
\[ \begin{align*}
\leq -S_x \Delta^2 Q_{x0} \Delta x + S_x \lambda_{\text{max}}(A_2) \| \Delta y_k \|^2 - (1/\varepsilon) S_y \Delta y^T Q_{y0} \Delta y \\
+ (1/\varepsilon) S_y \lambda_{\text{max}}(A_4) \| \Delta y_k \|^2 \\
\leq -S_x \Delta^2 Q_{x0} \Delta x + S_x \left( \lambda_{\text{max}}(A_2) \Delta y_k^2 + \frac{\lambda_{\text{max}}(D_1) R_x}{\sigma_{\text{min}}(P_x)} \right) \\
\quad - (1/\varepsilon) S_y \Delta y^T Q_{y0} \Delta y + (1/\varepsilon) S_y \left( \lambda_{\text{max}}(A_4) \Delta y_k^2 + \frac{\lambda_{\text{max}}(D_2) R_y}{\sigma_{\text{min}}(P_y)} \right)
\end{align*} \]

In Case (b), one has
\[ V_{21} \leq -S_x \Delta^2 Q_{x0} \Delta x - \Delta^2 f_x(\Delta y - D_2 \Delta x) \]
\[ \leq -S_x \Delta^2 Q_{x0} \Delta x + S_x \left( \lambda_{\text{max}}(A_2) \| \Delta y_k \|^2 + \lambda_{\text{max}}(D_1) \| \Delta x \|^2 \right) \]
\[ \leq -S_x \Delta^2 Q_{x0} \Delta x + S_x \left( \lambda_{\text{max}}(A_2) \Delta y_k^2 + \frac{\lambda_{\text{max}}(D_1) R_x}{\sigma_{\text{min}}(P_x)} \right) \]
\[ \quad - (1/\varepsilon) S_y \Delta y^T Q_{y0} \Delta y + (1/\varepsilon) S_y \left( \lambda_{\text{max}}(A_4) \Delta y_k^2 + \frac{\lambda_{\text{max}}(D_2) R_y}{\sigma_{\text{min}}(P_y)} \right) \]
\[ \quad \left( \lambda_{\text{max}}(A_2) \Delta y_k^2 + \frac{\lambda_{\text{max}}(D_1) R_x}{\sigma_{\text{min}}(P_x)} \right) \]
\[ \left( \lambda_{\text{max}}(A_4) \Delta y_k^2 + \frac{\lambda_{\text{max}}(D_2) R_y}{\sigma_{\text{min}}(P_y)} \right) \]
\[ \left( \lambda_{\text{max}}(A_2) \Delta y_k^2 + \frac{\lambda_{\text{min}}(D_1) R_x}{\sigma_{\text{min}}(P_x)} \right) \]
\[ \left( \lambda_{\text{max}}(A_4) \Delta y_k^2 + \frac{\lambda_{\text{min}}(D_2) R_y}{\sigma_{\text{min}}(P_y)} \right) \]
\[ \left( \lambda_{\text{max}}(A_2) \Delta y_k^2 + \frac{\lambda_{\text{min}}(D_1) R_x}{\sigma_{\text{min}}(P_x)} \right) \]
\[ \left( \lambda_{\text{max}}(A_4) \Delta y_k^2 + \frac{\lambda_{\text{min}}(D_2) R_y}{\sigma_{\text{min}}(P_y)} \right) \]
\[ \left( \lambda_{\text{max}}(A_2) \Delta y_k^2 + \frac{\lambda_{\text{min}}(D_1) R_x}{\sigma_{\text{min}}(P_x)} \right) \]
\[ \left( \lambda_{\text{max}}(A_4) \Delta y_k^2 + \frac{\lambda_{\text{min}}(D_2) R_y}{\sigma_{\text{min}}(P_y)} \right) \]

Remark 3. \( S_x \) and \( S_y \) are the dead-zone functions, which prevent the weights drifting into infinity when the modeling error presents \[ [22]. \] This is known as "parameters drift" \[ [23]. \] phenomenon.

It is noticed that \( H_x \) and \( H_y \) are thresholds for the identification error. For case \( (a) \), where \( S_x > 0, S_y > 0 \), i.e. \( \| P_{nu}^2 \| \Delta x \| > H_x \), \( \| P_{nu}^2 \| \Delta y \| > H_y \), smaller thresholds as in \[ (33) \] could be used, but we extend those to \( H_x \) and \( H_y \) to unify the thresholds for all the possible Cases \( (a), (b), (c), (d) \) during the entire identification process.

In the all above 4 cases, since \( V_{21} = V_{2x} + V_{2y} \) are positive definite, \( V_{21} = V_{2x} + V_{2y} \leq 0 \) can be achieved using the update laws \[ (24) \]. This implies \( \Delta x, \Delta y, W_{1,2,3,4}, A, B \in L_{\infty} \). Furthermore, \( x_{\text{nom}} = \Delta x + x_{\text{nom}} = \Delta y + y \) are also bounded. From the error Eqs. \[ (16) \], with the assumption that error and disturbances are bounded, we can draw the conclusion that \( \Delta x, \Delta y \in L_{\infty} \). Since the control input \( \gamma(U) \) and \( \sigma_{1,2}(\cdot) \) are bounded, it is concluded that \( \lim_{t \to \infty} \tilde{W}_{1,2} = 0 \), \( \lim_{t \to \infty} W_{3,4} = 0 \).

The structure of the improved identification scheme is illustrated in Fig. 4.

4. Applications

To illustrate the theoretical results, we give the following three examples.

Example 1. Let us consider the nonlinear system in \[ [18] \]
\[ x_1 = x_2 x_1 + \beta_1 \text{sign}(x_2) + u_1 \]
\[ x_2 = x_2 x_2 + \beta_2 \text{sign}(x_2) + u_2 \]
\[ \text{where we use the same parameter } x_1 = -5, x_2 = 10, \beta_1 = 3, \beta_2 = 2 \], \( \gamma = x_1(0) = -5, x_2(0) = -5, \tau = 0.2 \) and identification algorithm for NN weights as those in \[ [18] \] except the additional linear matrices identification algorithm and different activated functions, which is a sigmoid function \( \sigma_{2}(\cdot) = \phi_{2}(\cdot) = 1/(1 + e^{-x}) \). \( k = 1.2 \). The same input signals are adopted where \( u_1 \) is a sinusoidal wave \( (u_1 = 8 \sin(0.05t)) \) and \( u_2 \) is a saw-tooth function with the amplitude 8 and frequency 0.02 Hz.

To show the identification performance of the proposed algorithm, the performance index—Root Mean Square (RMS)—for the states error has been adopted for the purpose of comparison:

\[ \text{RMS} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} e^2(i)} \]

where \( n \) is number of the simulation steps, \( e(i) \) is the difference between the state variables in model and system at \( i \)th step.
In the first NN identifier, for state variable \( x_1 \), the RMS value is 0.139102 and the RMS for state variable \( x_2 \) is 0.116635. In the second NN identifier, for state variable \( x_1 \), the RMS value is 0.047168 and the RMS for state variable \( x_2 \) is 0.020158. The RMS values of both state variables demonstrate that the second NN identifier has better performance than the first one.

The results in Figs. 5–9 demonstrate that the identification performance has been improved compared to those in [18]. It can be seen that the state variables of dynamic multi-time scale NN follow those of the nonlinear system more accurately and quickly.

The eigenvalues of the linear parameter matrix are shown in Fig. 9. The eigenvalues for both \( A \) and \( B \) are universally smaller than zero, which means they are kept as stable during the identification.

**Example 2.** In 1952, Hodgkin and Huxley proposed a system of differential equations describing the flow of electric current through the surface membrane of a giant nerve fiber. Later this Hodgkin–Huxley (HH) model of the squid giant axon became one of the most important models in computational neuroscience and a prototype of a large family of mathematical models quantitatively describing

![Fig. 5. Identification result for \( x_1 \): (a) first NN identifier and (b) second NN identifier.](image1)

![Fig. 6. Identification error for \( x_1 \): (a) first NN identifier and (b) second NN identifier.](image2)

![Fig. 7. Identification result for \( x_2 \): (a) first NN identifier and (b) second NN identifier.](image3)
Fig. 8. Identification error for $x_2$: (a) first NN identifier and (b) second NN identifier.

Fig. 9. Eigenvalues of the linear parameter matrices: (a) first NN identifier and (b) second NN identifier.

Fig. 10. Identification results: (a) first NN identifier and (b) second NN identifier. Solid line: state $V, n, m, h$; dotted lines: identified states of the NN model.
electrophysiology of various living cells and tissues [16,19]:

\[
\begin{align*}
\frac{dV}{dt} &= \frac{1}{C_m} (I_{\text{ext}} - g_K n^4 (V + E_K) - g_Na m^3 h (V + E_Na)) \\
\frac{dn}{dt} &= \frac{1}{\tau_n} (n_m - n) \\
\frac{dh}{dt} &= \frac{1}{\tau_h} (h_m - h) \\
\frac{dm}{dt} &= \frac{1}{\tau_m} (m_m - m)
\end{align*}
\]

(45)

where time \( t \) is measured in ms, variable \( V \) is the membrane potential in mV, and \( n, m, \) and \( h \) are dimensionless gating variables corresponding to \( K^+ \), \( Na^+ \) and leakage current channels, respectively, which can vary in the range of \( [0,1] \):

\[
\begin{align*}
n_\infty &= \frac{1}{1 + e^{(V + 40)/12}}, & m_\infty &= \frac{1}{1 + e^{(V + 55)/11}}, & h_\infty &= \frac{1}{1 + e^{(V + 65)/18}} \\
\tau_n &= \frac{1}{e^{(V + 40)/12}} - 1, & \tau_m &= \frac{1}{e^{(V + 55)/11}} - 1, & \tau_h &= \frac{1}{e^{(V + 65)/18}} - 1 \\
\alpha_n &= 0.125e^{-(V/80)}, & \alpha_m &= 4e^{-(V/18)}, & \alpha_h &= \frac{1}{e^{30-V/10} + 1}
\end{align*}
\]

\( g_K = 36 \text{mS/cm}^2, \ g_Na = 120 \text{mS/cm}^2, \ g_l = 0.3 \text{mS/cm}^2 \)

\( E_K = -12 \text{mV}, \ E_Na = 115 \text{mV}, \ E_l = 10.599 \text{mV}, \ C_m = 1 \mu \text{F/cm}^2 \)

**Remark 4.** From the electrophysiology point of view, the most important state of the HH system is the membrane potential \( V \), which exhibits multifarious electro-physic phenomena and is also the core of the numerous former researches. Instead of using the original HH model, we used the (2, 2) asymptotic embedded system [16].

**Remark 5.** We took the modified HH model with the effect of extremely low frequency (ELF) external electric field \( E_w \), which served as the other control input besides the external applied stimulation current \( I_{\text{ext}} \).

---

Fig. 11. Eigenvalues of the linear matrices \( A \) and \( B \): (a) first NN identifier and (b) second NN identifier.

Fig. 12. Identification results: (a) first NN identifier and (b) second NN identifier. Solid line: state \( V, n, m, h \); dotted lines: identified states of the NN model.
Numerous researchers have carried out research on applying various stimulations to HH model. In this research, we use the input signals with which the states of NN can still follow those of the HH system:

\[ I_{\text{ext}} = \frac{1}{2} A_I \cos \omega_I t + 1, \quad E_w = \frac{1}{2} A_E \cos \omega_E t \]

where \( \omega_I = 2\pi f_I \), and all the initial conditions for the HH system are the equilibrium (quiescent), \( V_0 = 0.00002, m_0 = 0.05293, h_0 = 0.59612, n_0 = 0.31768 \).

We pick two typical stimulations, which can result in significant and classical neuron excitation:

1. \( E_w = 0, A_I = 30 \mu \text{A/cm}^2, f_I = 10 \text{ Hz}, \epsilon = 0.2 \)
2. \( I_{\text{ext}} = 0, A_E = 10 \text{ mV}, f_E = 115 \text{ Hz}, \epsilon = 0.2 \)

The identification results are presented in Figs. 10–13 for (2, 2) asymptotic embedded HH model. For the first NN identifier, in Case I, system is in 8/1 phase locked oscillation periodic bursting. RMS values of the state variables are \( \text{RMS}_V = 0.333511, \text{RMS}_h = 0.335092, \text{RMS}_m = 0.322476 \).

In Case II, system is in the same frequency periodic spiking. RMS values of the state variables are \( \text{RMS}_V = 0.140449, \text{RMS}_h = 0.147785, \text{RMS}_m = 0.07245 \). The time scale is considered by putting \( \epsilon = 0.2 \).
flexibility of linear part matrix $A$ and $B$ enhance the identification ability of the NN identifier. Even the single layer structure is powerful enough to successfully follow the complicated electro-physic phenomena from HH model.

For the second NN identifier, in simulation Case I, System is in $8/1$ phase locked oscillation periodic bursting. RMS values of the state variables are $\text{RMS}_n=0.074642$, $\text{RMS}_h=0.083497$, $\text{RMS}_V=0.438275$, $\text{RMS}_m=0.035473$. In Case II, the system is in the same frequency periodic spiking. The RMS values of the state variables are $\text{RMS}_n=0.05695$, $\text{RMS}_h=0.061458$, $\text{RMS}_V=0.86327$, $\text{RMS}_m=0.060288$. The time scale is considered by putting $\varepsilon=0.2$. From Figs. 10–13, we can see that the states of the NN model can follow those of the HH model very closely. The identification performance of the proposed algorithm is better than that of the first NN identifier, especially for the membrane potential. The eigenvalues of $A$ and $B$ for (a) and (b) converge to the same steady values since the nominal linear matrices $A^*$ and $B^*$ do not change with different inputs.

Example 3. To further illustrate the theoretical results, we test two NN identifiers on a DC servomotor, which is a typical example of two time-scales system. DC motor modeling can be separated into two electrical and mechanical subsystems. It is well known that the time constant of the electrical system is much smaller than that of the mechanical system. Hence, the electrical sub-system is the fast subsystem and the mechanical system is the slow subsystem. The model of DC motor [8] is shown as follows:

\begin{align}
\omega_t &= I_e \\
\dot{\omega}_t &= -\omega_t - I_e + i_t
\end{align}

where $\omega_t$ is related to the speed of motor, $i_e$ is related to circuit current, $\varepsilon$ is the time scale and $i_t$ is related to the input voltage to the circuit.

The parameter of the time scale is $\varepsilon=0.5$ and the input signal is $i_t = 3 \sin 0.5 t$. The sigmoid functions $\sigma_{1,2}(\cdot)$ and $\phi_{1,2}(\cdot)$ are chosen as $1/(1+e^{-\cdot})$. In the first NN identifier, for state variable $\omega_t$, the RMS value is 0.02742 and RMS for state variable $i_e$ is 0.2325. In the second NN identifier, for state variable $\omega_t$, the RMS value is 0.02324 and RMS for state variable $i_e$ is 0.04108. The RMS values of both state variables demonstrate that the second NN identifier has better performance than the first one. The results in Figs. 14–17 demonstrate that the states of dynamic multi-scale NN follow those of the nonlinear system more accurately and quickly.

The simulation results of three nonlinear systems demonstrate that the states of dynamic multi-scale neural networks can track the nonlinear system state variables on-line. The identification errors approach to the thresholds. The eigenvalues of $A$ and $B$ converge to the steady values in the system identifications.

5. Conclusions

In this paper we propose two new NN identifiers for nonlinear systems with multi-time scales. The on-line identification algorithms for dynamic neural networks have been developed. The Lyapunov synthesis method is used to prove the stability of identification algorithms. The proposed algorithms are applied
three multi-time scale systems. The identification results show
the effectiveness of the proposed identification algorithms.

References

[1] A. Meyer-Base, F. Ohl, H. Scheich, Singular perturbation analysis of compe-


[4] E. W. Sey, Y. Cao, Nonlinear system identification for predictive control
using continuous time recurrent neural networks and automatic differentia-


based identification of SMB chromatographic processes, Control Engineering


[8] Alejandro Cruz Sandoval, Wen Yu, Xiaou Li, Some stability properties of
dynamic neural networks with different time-scales, in: Proceedings of the
2006 International Joint Conference on Neural Networks, Vancouver, BC,

different timescales, in: Proceedings of the 2002 International Conference

[10] A. Meyer-Base, F. Ohl, H. Scheich, Quadratic-Type Lyapunov functions for
competitive neural networks with different Time-Scales, IEEE Transactions on

titive neural networks with different time scales, IEEE Transactions on

organizing neural network with feedforward and feedback dynamics, in:
Proceedings IEEE International Joint Conference on Neural Networks, vol. 2,

[13] A. Meyer-Base, F. Ohl, H. Scheich, Stability analysis techniques for competi-
tive neural networks with different time-scales, in: Proceedings of the IEEE
International Conference on Neural Networks, vol. 6, 27 November–1

[14] Wen Yu, Xiaou Li, Some new results on system identification with dynamic
neural networks, IEEE Transactions on Neural Networks 12 (2) (March 2001)
412–417.

driving dynamical neural networks, IEEE Transactions on Systems, Man, and

[16] Rebecca Suckley, Vadim N. Biktashev, The asymptotic structure of the


[18] Alexander S. Poznyak, Wen Yu, Edgar N. Sanchez, Jose P. Perez, Nonlinear
adaptive trajectory tracking using dynamic neural networks, IEEE Transac-
tions on Neural Networks 10 (6) (1999) 1402–1411.

[19] Eugene M. Izhikevich, Dynamical Systems in Neuroscience: The Geometry of


Proceedings of the IEEE Communications, Computers and Power in the

[22] W. Yu, A.S. Poznyak, Indirect adaptive control via parallel dynamic neural
networks, control theory and applications, IEE Proceedings 146 (1) (1999)
30–33.

River, NJ, 1996.

[24] X.M. Ren, A.B. Rad, P.T. Chan, W.L. Lo, Identification and control of contin-
uous-time nonlinear systems via dynamic neural networks, IEEE Transac-


[26] Boris T. Polyak, Introduction to Optimization, Optimization Software, Inc,


[28] A.S. Poznyak, E.N. Sanchez, W. Yu, Differential Neural Networks for Robust
Nonlinear Control, identification, State estimation and Trajectory Tracking,

Xuan Han received his B.S. Degree from Department of Electrical & Automation Engineering, Tianjin Univer-
sity, P.R. China, in 2006 and his master degree from Concordia University in 2010. His research interest
includes neural network, nonlinear system identifica-
tion and control, mechatronics.

Wen-Fang Xie is an associate professor with the
Department of Mechanical and Industrial Engineering
at Concordia University, Montreal, Canada. She was an
Industrial Research Fellowship holder from Natural
Sciences and Engineering Research Council of Canada
(NSERC) and served as a senior research engineer in In-
CoreTec, Inc., Canada before she joined Concordia
University in 2003. She received her Ph.D. from the
Hong Kong Polytechnic University in 1999 and her
master’s degree from Beijing University of Aeronautics
and Astronautics in 1991. Her research interests
include nonlinear control and identification in mecha-
tronics, artificial intelligent control and advanced pro-
cess control and system identification.

Zhijun Fu received his B.S. degree in electrical infor-
mation engineering from Henan Normal University,
China, in 2005. He obtained his master degree from the
Science and Technology University of Beijing, China.
He is currently a Ph.D. student in the Science and
Technology University of Beijing. He is carrying out his
Ph.D. research as a visiting scholar in the Concordia
University (Canada). His research interests include
control theory applications, electrical machine drives,
power electronics, induction motor control and elec-
tric vehicle control.

Weidong Luo received his B.S. degree from the Science
and Technology University, Beijing, China, in 1982 and
the master degree in mechanical engineering from Chine
University of Petroleum in 1985. He is a profes-
sor with the Department of Mechanical and Industrial
Engineering at the Science and Technology University,
Beijing, China. His research interests include
mechanics of complex systems, fault diagnosis and
detection, vehicle intelligent control, modern design
theory and technology.