Computación y Sistemas Número Especial pp. 083 - 091
© 2002, CIC - IPN. *ISSN 1405-5546* Impreso en México

Passivity Analysis for Dynamic Neuro Identifier

Neuro Identificación Mediante Técnicas de Pasividad

Wen Yu
Centro de Investigación y Estudios Avanzados-IPN
Departamento de Control Automático Av. IPN 2508
Apartado Postal 14-740, CP. 07000, México D.F.,
E-mail: yuw@ctrl.cinvestav.mx

Article received on June 21, 2001; accepted on February 02, 2002

Abstract

In this paper, dynamic neural networks (single layer and multilayer) are used for nonlinear system on-line identification. Passivity approach is applied to access several new stability properties of the neuro identifier. The conditions for passivity, stability, asymptotic stable and input-to-state stable are established. We conclude that the backpropagation algorithm with a modification term which is determined by off-line learning may make the neuro identifier robust with respect to any bounded uncertainty.

Keywords: System Identification, Neural Networks, Passivity.

Resumen

En este artículo, por la técnica pasiva, damos varios algoritmos estables para multicapas y las solas capas de las redes de neuronales, estos algoritmos nuevos han sido exitosamente aplicados para muchos sistemas reales por simulaciones. Las condiciones para la pasividad, estabilidad, entrada para estado estable se establecen. Concluímos que por medio de la técnica de pasividad se obtiene que el algoritmo BP, en el proceso de identificación neuronal, es robusto con respecto a todos los tipos de incertidumbres acotadas.

Palabras Clave: Identificación, Redes Neuronales y Pasividad

Introduction

In manyapplication, «black-box» identification using neural networks has emerged as a viable tool for unknown nonlinear systems. This model-free approach uses the nice features of neural networks, but the lack of model makes hard to obtain theoretical results on the stability and performance of neuro identifier. For the engineers it is very important to assure the stability in theory before they apply in to a real system.

There are not many results on stability analysis of neural networks in spite of their successful applications. The global asymptotic stable (GAS) of dynamic neural networks hes been developed during the last decade. Diagonal stability (Kaszkurewics and Bhaya, 1994) and negative semi-definiteness (Forti, et al., 1994) of the interconnection matrix may make Hopfield-Thank neuro circuit GAS. Multilayer perceptrons (MPL) nad recurrent neural networks can be related to the Lur`e systems, the absolute stabilities were developed by (Suykens, et al., 1999) and (Matsouka, 1992). Input-to-state stable (ISS) analysis method (Sontag and Wang, 1995) is an effective tool for dynamic neural networks (Yu, 2001) stated that if the weights are small enough, neural networks are ISS and GAS with zero input. Many publishes investigate the stability of identification error and tracking error of neural networks. (Jagannathan and Lewis, 1996) studied the stability conditions when multilayer perceptrons are used to identify and control a nonlinear system. Lyapunov-like analysis is suitable for dynamic neural network, the signal-layer case were discussed in (Rovithakis and Christodoulou, 1994) and (Yu and Li, 2001), the high-order networks and multilayer networks may be found in (Kosmatopoulos, *et al.*, 1995) and (Poznyak, *et al.*, 1999). Since neural networks cannot match the unknown nonlinear systems exactly, some robust modifications (Ioannou and Sun, 1996) should be applied on normal gradient or backpropagation algorithm (Jagannathan and Lewis, 1996)(Rovithakis and Christodoulou, 1994)(Suykens, *et al.*, 1999)(Yu and Li, 2001). One of advantages of "black-box" identification is that the identification error can be regarded inside the "black-box", so the gradient algorithm without any modification maybe have robust properties.

In this paper, we will use passivity theory to analyze the stability of the multilayer dynamic neuro identifier. Passivity approach may deal with the poor define nonlinear systems, usually by means of sector bounds, and offers elegant solutions for the proof of absolute stable. It can lead to general conclusions on the stability using only input-output characteristics. The passivity properties of MLP were examined in (Commuri and Lewis, 1996). By means of analyzing the interconnected of error models, they derived the relationship between passivity and closed-loop stable. To the best of our knowledge, open loop analysis based on the passivity method for multilayer dynamic neural networks has not yet been established in the literature. We will show that a backpropagation-like learning law can make the identification error stable, asymptotic stable and input-to-state stable. A Simulations of vehicle idle speed identification gives the effective of the algorithm of this paper.

2. SYSTEM IDENTIFICATION WITH SINGLE LAYER NEURAL NETWORKS

The nonlinear system to be identified is given as:

$$x_t = f(x_t, u_t), \quad y_t = x_t, \quad x_t \in \Re^n,\ u_t \in \Re^m \tag{1}$$

We construct the following single layer dynamic neural network:

$$\dot{\widehat{x}}_t = A\widehat{x}_t + W_{1,t}\sigma(\widehat{x}_t) + W_{2,t}\phi(\widehat{x}_t)\gamma(u_t) \tag{2}$$

where $\widehat{x}_t \in \Re^n$ is the state of the neural network, $A \in \Re^{n\times n}$ is a stable matrix. $W_{1,t} \in \Re^{n\times n}$, $W_{2,t} \in \Re^{n\times n}$ are weight matrices of neural networks. The vector functions $\sigma(x_t) \in \Re^n$ is assumed to be $n-$dimensional with the elements increasing monotonically. The matrix function $\phi(\cdot)$ is assumed to be $\Re^{n\times m}$ diagonal: $\phi(\widehat{x}_t) = diag(\phi_1(\widehat{x}_1)\cdots\phi_n(\widehat{x}_n))$. $\gamma(u_t) \in \Re^m$, u_t is the control input of the plant (1). Function $\gamma(\cdot)$ is selected as $\|\gamma(u_t)\|^2 \le \overline{u}$. The typical presentation of the elements

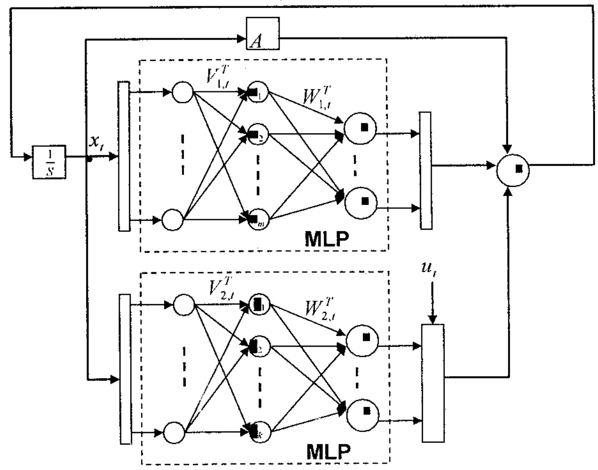

Fig. 1. The structure of the dynamic neural network.

$\sigma_i(\cdot)$ and $\phi_i(.)$ are as sigmoid functions, i.e. $\sigma_i(x_i) = \frac{a_i}{1+e^{-b_i x_i}} - c_i$.

The structure of the neuro identifier is shown in Fig.1.

Remark 1. The neural networks have been discussed by many authors, for example (Rovithakis and Christodoulou, 1994), (Kosmatopoulos, *et al.*, 1995), (Poznyak, *et al.*, 1999) and (Yu and Li, 2001). It can be seen that Hopfield model is the special case of this networks with $A = diag\{a_i\}$, $a_i := -1/R_iC_i$, $R_i > 0$ and $C_i > 0$. R_i and C_i are the resistance and capacitance at the ith node of the network respectively.

Let us define identification error as $\Delta_t := \widehat{x}_t - x_t$. Because $\sigma(\cdot)$ and $\phi(\cdot)$ are chosen as sigmoid functions, clearly they satisfy the following assumption.

A1: *The function $\sigma(\cdot)$ and $\phi(\cdot)$ fulfill generalized Lipshitz condition*

$$\widetilde{\sigma}^T\Lambda_1\widetilde{\sigma} \le \Delta_t^T D_\sigma \Delta_t$$
$$\left(\widetilde{\phi}_t\gamma(u_t)\right)^T \Lambda_2 \left(\widetilde{\phi}_t\gamma(u_t)\right) \le \overline{u}\Delta_t^T D_\phi \Delta_t$$

where $\widetilde{\sigma} := \sigma(\widehat{x}_t) - \sigma(x_t)$, $\widetilde{\phi} := \phi(\widehat{x}_t) - \phi(x_t)$, Λ_1, Λ_2, *D_σ and D_ϕ are known positive constants, nonlinear.*

Since dynamic neural networks are like black-boxes, they can follow any nonlinear systems in any accuracy if the neural networks are big enough (Matsouka, 1992), we may represent the nonlinear system (1) with the single layer neural network (2) plus a modeling error as

$$\dot{x}_t = Ax_t + W_1^*\sigma(x_t) + W_2^*\phi(x_t)\gamma(u_t) - \widetilde{f}_t \tag{3}$$

where W_1^* and W_2^* are bounded unknown matrices

$$W_1^*\Lambda_1^{-1}W_1^{*T} \le \overline{W}_1, \quad W_2^*\Lambda_2^{-1}W_2^{*T} \le \overline{W}_2 \tag{4}$$

$\overline{W}_1$ and $\overline{W}_2$ are priory known matrices, vector function $\widetilde{f}_t$ can be regarded as modelling error and disturbances.

We may define matrix errors as $\widetilde{W}_{1,t} := W_{1,t} - W_1^*$, $\widetilde{W}_{2,t} := W_{2,t} - W_2^*$. The error dynamic is obtained from (2) and (3)

$$\dot{\Delta}_t = A\Delta_t + \widetilde{W}_{1,t}\sigma(\widehat{x}_t) + \widetilde{W}_{2,t}\phi(\widehat{x}_t)\gamma(u_t) + W_1^*\widetilde{\sigma} + W_1^*\widetilde{\phi}\gamma(u_t) + \widetilde{f}_t \quad (5)$$

If we define

$$R := \overline{W}_1 + \overline{W}_2, \quad Q := D_\sigma + \overline{u}D_\phi + Q_0 \quad (6)$$

and the matrices A and Q_0 are selected to fulfill the following conditions:

(1) the pair $(A, R^{1/2})$ is controllable, the pair $(Q^{1/2}, A)$ is observable,

(2) local frequency condition satisfies

$$A^T R^{-1} A - Q \geq \frac{1}{4}\left[A^T R^{-1} - R^{-1}A\right] R\left[A^T R^{-1} - R^{-1}A\right]^T \quad (7)$$

then the following assumption can be established:

A2: There exist a stable matrix A and a strictly positive defined matrix Q_0 such that the matrix Riccati equation

$$A^T P + PA + PRP + Q = 0 \quad (8)$$

has a positive solution $P = P^T > 0$.

This conditions is easily fulfilled if we select A as stable diagonal matrix. Next theorem states the learning procedure of neuro identifier.

Theorem 1. If the weights $W_{1,t}$ and $W_{2,t}$ are updated as

$$\begin{aligned} \dot{W}_{1,t} &= -K_1 P\Delta_t \sigma^T(\widehat{x}_t) \\ \dot{W}_{2,t} &= -K_2 P\phi(\widehat{x}_t)\gamma(u_t)\Delta_t^T \end{aligned} \quad (9)$$

where P is the solution of Riccati equation (8), then the dynamic of identification error (5) is strictly passive from $\widetilde{f}_t$ to the identification error $2P\Delta_t$

Proof. Select a Lyapunov function (storage function) as

$$S_t = \Delta_t^T P\Delta_t + tr\left\{\widetilde{W}_{1,t}^T K_1^{-1}\widetilde{W}_{1,t}\right\} + tr\left\{\widetilde{W}_{2,t}^T K_2^{-1}\widetilde{W}_{2,t}\right\} \quad (10)$$

where $P \in \Re^{n\times n}$ is positive definite matrix. According to (5), the derivative is

$$\begin{aligned} \dot{S}_t &= \Delta_t^T\left(PA + A^T P\right)\Delta_t + 2\Delta_t^T P\widetilde{W}_{1,t}\sigma(\widehat{x}_t) \\ &+2\Delta_t^T P\widetilde{W}_{2,t}\phi(\widehat{x}_t)\gamma(u_t) + 2\Delta_t^T P\widetilde{f}_t \\ &+2\Delta_t^T P\left[W_1^*\widetilde{\sigma} + W_1^*\widetilde{\phi}\gamma(u_t)\right] \\ &+2tr\left\{\dot{\widetilde{W}}_{1,t}^T K_1^{-1}\widetilde{W}_{1,t}\right\} + 2tr\left\{\dot{\widetilde{W}}_{2,t}^T K_2^{-1}\widetilde{W}_{2,t}\right\} \end{aligned}$$

Since $\Delta_t^T P W_1^*\widetilde{\sigma}_t$ is scalar, using (20) and matrix inequality

$$X^T Y + \left(X^T Y\right)^T \leq X^T\Lambda^{-1}X + Y^T\Lambda Y \quad (11)$$

where $X, Y, \Lambda \in \Re^{n\times k}$ are any matrices, Λ is any positive definite matrix, we obtain

$$\begin{aligned} 2\Delta_t^T P W_1^*\widetilde{\sigma}_t &\leq \Delta_t^T P W_1^*\Lambda_1^{-1}W_1^{*T}P\Delta_t \\ &+\widetilde{\sigma}_t^T\Lambda_1\widetilde{\sigma}_t \leq \Delta_t^T\left(P\overline{W}_1 P + D_\sigma\right)\Delta_t \\ 2\Delta_t^T P W_2^*\widetilde{\phi}_t\gamma(u_t) &\leq \Delta_t^T\left(P\overline{W}_2 P + \overline{u}D_\phi\right)\Delta_t \end{aligned} \quad (12)$$

So we have

$$\begin{aligned} \dot{S}_t &\leq \Delta_t^T\begin{bmatrix} PA + A^T P + P\left(\overline{W}_1 + \overline{W}_2\right)P \\ +\left(D_\sigma + \overline{u}D_\phi + Q_0\right)\end{bmatrix}\Delta_t \\ &+2tr\left\{\dot{\widetilde{W}}_{1,t}^T K_1^{-1}\widetilde{W}_{1,t}\right\} + 2\Delta_t^T P\widetilde{W}_{1,t}\sigma(\widehat{x}_t) + 2\Delta_t^T P\widetilde{f}_t \\ &+2tr\left\{\dot{\widetilde{W}}_{2,t}^T K_2^{-1}\widetilde{W}_{2,t}\right\} + 2\Delta_t^T P\widetilde{W}_{2,t}\phi(\widehat{x}_t)\gamma(u_t) \\ &-\Delta_t^T Q_0\Delta_t \end{aligned}$$

Since $\dot{\widetilde{W}}_{1,t} = \dot{W}_{1,t}$, if we use the updating law as in (25) and **A1**, we have

$$\dot{S}_t \leq -\Delta_t^T Q_0\Delta_t + 2\Delta_t^T P\widetilde{f}_t \quad (13)$$

From Definition 1, if we define the input as $\widetilde{f}_t$ and the output as $2P\Delta_t$, then the system is strictly passive with $V_t = \Delta_t^T Q_0\Delta_t \geq 0$.

Remark 2. Since the updating rate is $K_i P$ $(i = 1, 2)$, and K_i can be any positive matrix, the learning process of dynamic neural network (9) is free of P, the solution of Riccati equation (8).

Corollary 2. If only parameters uncertainty present $(\widetilde{f}_t = 0)$, then the updating law as (9) can make the identification error asymptotic stable,

$$\lim_{t\to\infty}\Delta_t = 0, \quad W_{1,t} \in L_\infty, \quad W_{2,t} \in L_\infty \quad (14)$$

Proof. Since the dynamic of identification error (5) is passive, the storage function $S(x_t)$ satisfies

$$\dot{S}(x_t) \leq \widetilde{f}_t^T 2P\Delta_t = 0$$

The positive defined $S(x_t)$ implies Δ_t, $W_{1,t}$ and $W_{2,t}$ are bounded.

From the error equation (5) $\Delta_t \in L_\infty$

$$\dot{S} \leq -\Delta_t^T Q_0 \Delta_t \leq 0 \tag{15}$$

Integrate (15) both sides

$$\int_0^\infty \|\Delta_t\|_{Q_0} \leq S_0 - S_\infty < \infty$$

So $\Delta_t \in L_2 \cap L_\infty$,, using Barlalat's Lemma we have (14). As u_t, $\sigma(\widehat{x}_t)$, $\phi(\widehat{x}_t)$ and P are bounded, $\lim_{t\to\infty} W_{1,t} = 0$ and $\lim_{t\to\infty} W_{1,t} = 0$.

Remark 3. For model matching case, Lyapunov-like analysis can reach the same result as Corollary 1 and Corollary 2 (Yu and Li, 2001). But in the case of modeling error, we will give a new conclusion on neuro identification: the gradient algorithm (9) is also robust respect to unmodeled dynamic, bounded disturbance and stochastic noise.

Theorem 3. Using the updating law as (9), the dynamic of neuro identifier (5) is input-to-state stable (ISS).

Proof. In view of the matrix inequality

$$X^TY + \left(X^TY\right)^T \leq X^T\Lambda^{-1}X + Y^T\Lambda Y$$
$$2\Delta_t^T P\widetilde{f}_t \leq \Delta_t^T P\Lambda_f P\Delta_t + \widetilde{f}_t^T \Lambda_f^{-1}\widetilde{f}_t$$

(15) can be represented as

$$\begin{aligned}\dot{S}_t &= -\Delta_t^T Q_0 \Delta_t + 2\Delta_t^T P\widetilde{f}_t \\ &\leq -\lambda_{\min}(Q_0)\|\Delta_t\|^2 + \Delta_t^T P\Lambda_f P\Delta_t + \widetilde{f}_t^T \Lambda_f^{-1}\widetilde{f}_t \\ &\leq -\alpha_{\|\Delta_t\|}\|\Delta_t\| + \beta_{\|\widetilde{f}_t\|}\left\|\widetilde{f}_t\right\|\end{aligned}$$

where

$$\alpha_{\|\Delta_t\|} := \left[\lambda_{\min}(Q_0) - \lambda_{\max}(P\Lambda_f P)\right]\|\Delta_t\|$$
$$\beta_{\|\widetilde{f}_t\|} := \lambda_{\max}\left(\Lambda_f^{-1}\right)\left\|\widetilde{f}_t\right\|$$

We can select a positive defined matrix Λ_f such that

$$\lambda_{\max}(P\Lambda_f P) \leq \lambda_{\min}(Q_0) \tag{16}$$

So α and β are $\mathcal{K}_\infty$ functions, S_t is an ISS-Lyapunov function. Using Theorem 1 of (Sontag and Wang, 1995), the dynamic of identification error (5) is input to state stable.

Corollary 4. If the modelling error $\widetilde{f}_t$ is bounded, then the updating law as (9) can make the identification procedure stable

$$\Delta_t \in L_\infty, \quad W_{1,t} \in L_\infty, \quad W_{2,t} \in L_\infty$$

Proof. From Property 2 we know input-to-state stable means that the behavior of the dynamic neural networks should remain bounded when its input is bounded.

Remark 4. Since the state and output variables are physically bounded, the modelling error $\widetilde{f}_t$ can be assumed to be bounded too (see, for example (Jagannathan and Lewis, 1996)(Poznyak, *et al.*, 1999)(Rovithakis and Christodoulou, 1994)). The condition (16) can be established if Λ_f is a small enough constant matrix. Unlike robust adaptive laws, such as dead-zone (Poznyak, *et al.*, 1999) and $\sigma-$modification (Kosmatopoulos, *et al.*, 1995), we do not need to know the upper bound of uncertainties.

Remark 5. It is well known that structure uncertainties will cause parameters drift for adaptive control, so one has to use robust modification to make identification stable (Ioannou and Sun, 1996). Robust adaptive methods may be extended to neuro identification directly (Jagannathan and Lewis, 1996)(Poznyak, *et al.*, 1999) (Rovithakis and Christodoulou, 1994). But neuro identification is a kind of "black-box" method, nobody needs structure information and all of uncertainties are inside the black box. Although robust adaptive algorithms are suitable for neuro identification, they are not the simplest. By means of passivity technique, we success to prove our conclusion: pure gradient algorithm is robust with respect to all kinds of bounded uncertainties for neuro identification.

3. SYSTEM IDENTIFICATION WITH MULTILAYER NEURAL NETWORKS

We construct the following multilayer dynamic neural network:

$$\dot{\widehat{x}}_t = A\widehat{x}_t + W_{1,t}\sigma(V_{1,t}\widehat{x}_t) + W_{2,t}\phi(V_{2,t}\widehat{x}_t)\pi(u_t) \tag{17}$$

where $\widehat{x}_t \in \Re^n$ is the state of the neural network, $A \in \Re^{n\times n}$ is a stable matrix. $W_{1,t} \in \Re^{n\times n}$, $W_{2,t} \in \Re^{n\times n}$ are weight matrices of output layers, $V_{1,t} \in \Re^{m\times n}$, $V_2 \in \Re^{m\times n}$ are weight matrices of hidden layers. The vector field $\sigma(x_t) : \Re^n \to \Re^m$ is assumed to have the elements increasing monotonically. The function $\phi(\cdot)$ is the transformation from $\Re^n$ to $\Re^{m\times m}$. $\pi(u_t) \in \Re^m$, is selected as saturation function: $\|\pi(u_t)\|^2 \leq \overline{u}$. The typical presentation of the elements $\sigma_i(\cdot)$ and $\phi_{ij}(.)$ are as sigmoid functions

$$\sigma_i(x_{i,t}) = a_i / \left(1 + e^{-b_i x_{i,t}}\right) - c_i$$

Generally, dynamic neural network (17) cannot follow the nonlinear system (1) exactly, it may be written as

$$\begin{array}{c} \dot{x}_t = Ax_t + W_1^*\sigma(V_1^0 x_t) \\ +W_2^*\phi(V_2^0 x_t)\pi(u_t) - \widetilde{f}_t\left(V_1^0, V_2^0\right) \end{array} \tag{18}$$

W_1^* and W_2^* are optimal matrix which may minimize modelling error $\widetilde{f}_t$, they are bounded as

$$W_1^*\Lambda_1^{-1}W_1^{*T} \leq \overline{W}_1, \quad W_2^*\Lambda_2^{-1}W_2^{*T} \leq \overline{W}_2 \tag{19}$$

V_1^0 and V_2^0 are prior given matrices which are obtained from off-line learning. Let us define identification error as $\Delta_t := \widehat{x}_t - x_t$, $\widetilde{\sigma}_t := \sigma(V_1^0\widehat{x}_t) - \sigma(V_1^0 x_t)$, $\widetilde{\phi}_t := \phi(V_2^0\widehat{x}_t)\pi(u_t) - \phi(V_2^0 x_t)\pi(u_t)$, $\widetilde{\sigma}_t' := \sigma(V_{1,t}\widehat{x}_t) - \sigma(V_1^0\widehat{x}_t)$, $\widetilde{\phi}_t' := \phi(V_{2,t}\widehat{x}_t)\pi(u_t) - \phi(V_2^0\widehat{x}_t)\pi(u_t)$, $\widetilde{V}_{1,t} := V_{1,t} - V_1^0$, $\widetilde{V}_{2,t} := V_{2,t} - V_2^0$, $\widetilde{W}_{1,t} := W_{1,t} - W_1^*$, $\widetilde{W}_{2,t} := W_{2,t} - W_2^*$. Because $\sigma(\cdot)$ and $\phi(\cdot)$ are chosen as sigmoid functions, clearly they satisfy Lipschitz condition

$$\begin{array}{cc} \widetilde{\sigma}_t^T\Lambda_1\widetilde{\sigma}_t \leq \Delta_t^T\Lambda_\sigma\Delta_t, & \widetilde{\phi}_t^T\Lambda_2\widetilde{\phi}_t \leq \overline{u}\Delta_t^T\Lambda_\phi\Delta_t \\ \widetilde{\sigma}_t' = D_\sigma\widetilde{V}_{1,t}\widehat{x}_t + \nu_\sigma, & \widetilde{\phi}_t' = D_\phi\widetilde{V}_{2,t}\widehat{x}_t + \nu_\phi \end{array} \tag{20}$$

where

$$D_\sigma = \frac{\partial\sigma^T(Z)}{\partial Z}\mid_{Z=V_{1,t}\widehat{x}_t}, \|\nu_\sigma\|_{\Lambda_1}^2 \leq l_1\left\|\widetilde{V}_{1,t}\widehat{x}_t\right\|_{\Lambda_1}^2$$
$$D_\phi = \frac{\partial\left[\phi(Z)\pi(u_t)\right]^T}{\partial Z}\mid_{Z=V_{2,t}\widehat{x}_t}, \|\nu_\phi\|_{\Lambda_2}^2 \leq l_2\left\|\widetilde{V}_{2,t}\widehat{x}_t\right\|_{\Lambda_2}^2$$

$l_1 > 0$, $l_2 > 0$, Λ_1, Λ_2, Λ_σ and Λ_ϕ are positive define matrices. The identification error dynamic is obtained from (17) and (18)

$$\begin{array}{c} \dot{\Delta}_t = A\Delta_t + \widetilde{W}_{1,t}\sigma(V_{1,t}\widehat{x}_t) + \widetilde{W}_{2,t}\phi(V_{2,t}\widehat{x}_t)\pi(u_t) \\ +W_1^*\widetilde{\sigma}_t + W_1^*\widetilde{\sigma}_t' + W_2^*\widetilde{\phi}_t + W_2^*\widetilde{\phi}_t' + \widetilde{f}_t\left(V_1^0, V_2^0\right) \end{array} \tag{21}$$

If we define

$$R := 2\overline{W}_1 + 2\overline{W}_2, \quad Q := D_\sigma + \overline{u}D_\phi + Q_0 \tag{22}$$

and the matrices A and Q_0 are selected to fulfill the following conditions:

(1) the pair $(A, R^{1/2})$ is controllable, the pair $(Q^{1/2}, A)$ is observable,

(2) local frequency condition satisfies

$$\begin{array}{c} A^TR^{-1}A - Q \geq \frac{1}{4}\left[A^TR^{-1} - R^{-1}A\right] \\ R\left[A^TR^{-1} - R^{-1}A\right]^T \end{array} \tag{23}$$

then the following assumption can be established:

A3: There exist a stable matrix A and a strictly positive defined matrix Q_0 such that the matrix Riccati equation

$$A^TP + PA + PRP + Q = 0 \tag{24}$$

has a positive solution $P = P^T > 0$.

This conditions is easily fulfilled if we select A as stable diagonal matrix. Next Theorem states the learning procedure of neuro identifier.

Theorem 5. If the weights $W_{1,t}$, $W_{2,t}$, $V_{1,t}$ and $V_{2,t}$ are updated as

$$\begin{array}{l} \dot{W}_{1,t} = -K_1P\sigma(V_{1,t}\widehat{x}_t)\Delta_t^T + K_1PD_\sigma\widetilde{V}_{1,t}\widehat{x}_t\Delta_t^T \\ \quad\dot{W}_{2,t} = -K_2P\phi(V_{2,t}\widehat{x}_t)\pi(u_t)\Delta_t^T \\ \quad\quad +K_2PD_\phi\widetilde{V}_{2,t}\widehat{x}_t\pi(u_t)\Delta_t^T \\ \dot{V}_{1,t} = -K_3PW_{1,t}D_\sigma\Delta_t\widehat{x}_t^T - \frac{l_1}{2}K_3\Lambda_1\widetilde{V}_{1,t}\widehat{x}_t\widehat{x}_t^T \\ \dot{V}_{2,t} = -K_4PW_{2,t}D_\phi\widehat{x}_t\Delta_t^T - \frac{l_2}{2}K_4\Lambda_2\widetilde{V}_{2,t}\widehat{x}_t\widehat{x}_t^T \end{array} \tag{25}$$

where P is the solution of Riccati equation (24) $\widetilde{V}_{i,t} = V_{i,t} - V_i^0$, $i = 1, 2$, then the dynamic of identification error (21) is strictly passive from $\widetilde{f}_t\left(V_1^0, V_2^0\right)$ to the identification error $2P\Delta_t$

Proof. see (Yu, 2002)

Remark 6. Since the updating gain is K_iP $(i = 1\cdots4)$ and K_i can be any positive matrix, the learning process of dynamic neural network (25) dose not depend on the solution of Riccati equation (24). $W_{1,t}D_\sigma\Delta_t$ is the error backpropagation for the hidden layer, $\widehat{x}_t^T$ is the input to the hidden layer; $\sigma(V_{1,t}\widehat{x}_t)$ is the input for the output layer, so the first parts $K_1P\sigma(V_{1,t}\widehat{x}_t)\Delta_t^T$ and $-K_3PW_{1,t}D_\sigma\Delta_t\widehat{x}_t^T$ are the same as the backpropagation scheme of multilayer perceptrons. The second parts are used to assure the passivate properties of identification error.

Corollary 6. If V_1^0 and V_2^0 are optimal values, and only parameters uncertainty present $(\widetilde{f}_t = 0)$, then the updating law as (25) can make the identification error asymptotic stable,

$$\lim_{t\to\infty}\Delta_t = 0 \tag{26}$$

Theorem 7. Using the updating law as (25), the dynamic of neuro identifier (21) is input-to-state stable (ISS).

Corollary 8. If the modelling error $\widetilde{f}_t$ is bounded, then the updating law as (25) can make the identification procedure stable

$$\Delta_t \in L_\infty, \quad W_{1,t} \in L_\infty, \quad W_{2,t} \in L_\infty$$

Proof. From Property 2 we know input-to-state stable means that the behavior of the dynamic neural networks should remain bounded when its input is bounded.

For model matching case, Lyapunov-like analysis (Yu and Li, 2001) can reach the same result as Corollary 1. But in the case of modeling error ($\widetilde{f}_t \neq 0$), robust modification terms have to be added in the updating law in order to assure stability (Jagannathan and Lewis, 1996)(Poznyak, *et al.*, 1999) (Rovithakis and Christodoulou, 1994). Since the state and output variables are physically bounded, the modelling error $\widetilde{f}_t$ can be assumed to be bounded too (see, for example (Jagannathan and Lewis, 1996)(Poznyak, *et al.*, 1999)(Rovithakis and Christodoulou, 1994)). The robust modification usually depends on the upper bound of modeling error $\widetilde{f}_t$. Unlike robust adaptive laws, such as dead-zone (Poznyak, *et al.*, 1999) and $\sigma-$modification (Kosmatopoulos, *et al.*, 1995), the updating law does not need the upper bound of uncertainties.

Theorem 9. If the modelling error $\widetilde{f}_t\left(V_1^0, V_2^0\right)$ is bounded as $\widetilde{f}_t^T \Lambda_f \widetilde{f}_t \leq \overline{\eta}\left(V_1^0, V_2^0\right)$, P in the updating law (25) is the solution of following Riccati equation

$$\begin{gathered} A^T P + PA + P\left(2\overline{W}_1 + 2\overline{W}_2 + \Lambda_f\right)P \\ + \left(D_\sigma + \overline{u}D_\phi + Q_0\right) = 0 \end{gathered} \tag{27}$$

then *the average of the identification error satisfies*

$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T \|\Delta_t\|^2_{Q_0}\, dt \leq \overline{\eta}\left(V_1^0, V_2^0\right) \tag{28}$$

Q_0 is any positive defined matrix.

Proof. Let define a Lyapunov function as (10), in view of the matrix inequality

$$\begin{aligned} 2\Delta_t^T P \widetilde{f}_t &\leq \Delta_t^T P \Lambda_f P \Delta_t + \widetilde{f}_t^T \Lambda_f^{-1} \widetilde{f}_t \\ &\leq \Delta_t^T P \Lambda_f P \Delta_t + \overline{\eta}\left(V_1^0, V_2^0\right) \end{aligned}$$

and the updating law (25), the derivative of the Lyapunov function (10) is

$$\begin{aligned} \dot{S}_t \leq \Delta_t^T [&PA + A^T P + P\left(2\overline{W}_1 + 2\overline{W}_2 + \Lambda_f\right)P \\ &+ \left(D_\sigma + \overline{u}D_\phi + Q_0\right)]\Delta_t - \Delta_t^T Q_0 \Delta_t + \overline{\eta} \end{aligned}$$

From (27) we have

$$\dot{S}_t \leq -\Delta_t^T Q_0 \Delta_t + \overline{\eta}\left(V_1^0, V_2^0\right) \tag{29}$$

Integrating (29) from 0 up to T yields

$$V_T - V_0 \leq -\int_0^T \Delta_t^T Q_0 \Delta_t dt + \overline{\eta} T$$

So

$$\int_0^T \Delta_t^T Q_0 \Delta_t dt \leq V_0 - V_T + \overline{\eta} T \leq V_0 + \overline{\eta} T$$

(28) is established

Remark 7. The learning law (25) is depends on prior known matrices V_1^0 and V_2^0, these two matrices will influence the modeling error $\widetilde{f}_t$ in (18). The *identification error will convergence to the ball radium the upper bounded of* $\widetilde{f}_t$. So an off-line learning is required to determine suitable V_1^0 and V_2^0, after that the neuro identifier may be used for nonlinear system on-line identification.

4. SIMULATION

The engine operation at idle is a nonlinear process that is far from its optimal range. Because it does not require any large degree of instrumentation or external sensing capabilities, the idle speed control is also accessible and can be formatted as a benchmark problem for control society. The process of engine at idle has time delays that vary inversely with engine speed and is time-varying due to aging of components and environmental changes such as engine warm-up after a cold start. The measurement of system outputs occurs asynchronously with the calculation of control signals. We assume that the occurrence of plant disturbances, such as engagement of air conditioner compressor, shift from neutral to drive in automatic transmissions, application and release of electric loads, and power steering lock-up, are not directly measured. The dynamic engine model a two inputs and two outputs system (Puskorius and Feldkamp, 1994):

$$\begin{gathered} \dot{P} = k_P\left(\dot{m}_{ai} - \dot{m}_{ao}\right), \quad \dot{N} = k_N\left(T_i - T_L\right) \\ \dot{m}_{ai} = \left(1 + k_{m1}\theta + k_{m2}\theta^2\right) g(P), \\ \dot{m}_{ao} = -k_{m3}N - k_{m4}P + k_{m5}NP + k_{m6}NP^2 \end{gathered}$$

The engine model parameters are for a 1.6 liter, 4-cylinder fuel injected engine

$$\begin{gathered} g(P) = \begin{cases} 1 & P < 50.6625 \\ 0.0197\sqrt{101.325P - P^2} & P \geq 50.6625 \end{cases} \\ T_i = -39.22 + 325024 m_{ao} - 0.0112\delta^2 + 0.635\delta \\ + \frac{2\pi}{60}\left(0.0216 + 0.000675\delta\right)N - \left(\frac{2\pi}{60}\right)^2 0.000102N^2 \\ T_L = \left(\frac{N}{263.17}\right)^2 + T_d, \quad m_{ao} = \dot{m}_{ao}(t-\tau)/\left(120N\right) \\ k_P = 42.40, \quad k_N = 54.26 \\ k_{m1} = 0.907, \quad k_{m2} = 0.0998 \\ k_{m3} = 0.0005968, \quad k_{m4} = 0.0005341 \\ k_{m5} = 0.000001757, \quad \tau = 45/N \end{gathered}$$

The system outputs are manifold press P (kPa) and engine speed N (rpm). The control inputs are thottle angle θ (degree) and the spark advance δ (degree). Disturbances act to the engine in the form of unmeasured

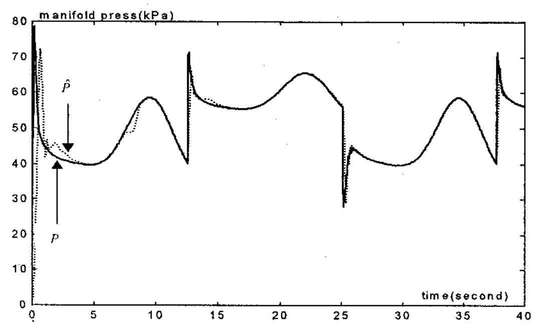

Fig. 2. Manifold press

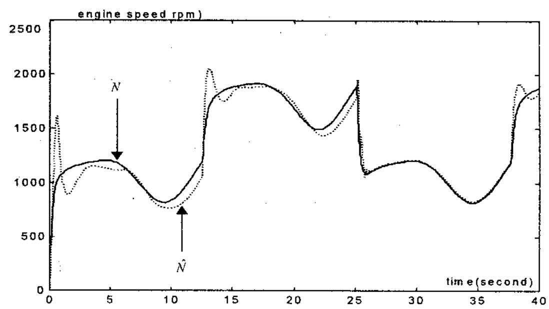

Fig. 3. Engine speed

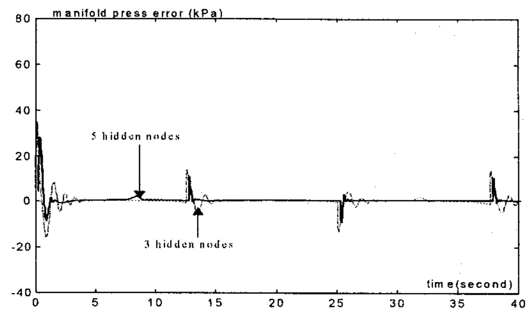

Fig. 4. Manifold press error with different V_1^0 and V_2^0

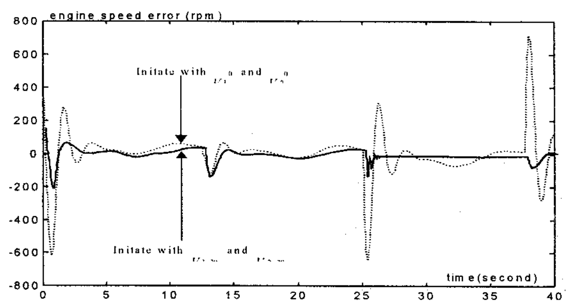

Fig. 5. Engine speed error with different V_1^0 and V_2^0

accessory torque T_d (N-m). The variable $\dot{m}_{ai}$ and $\dot{m}_{ao}$ refer to the mass air flow into and out of the manifold. m_{ao} is the air mass in the cylinder. The parameter τ is a dynamic transport time delay. The function $g(P)$ is a manifold pressure influence function. T_i is the engine's internally developed torque, T_L is the load torque. If we define $x = (P,\ N)^T$, $\quad u = (\theta,\ \delta)^T$, then the dynamic of vehicle idle speed are $\dot{x}_t = \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} f_1(x,u) \\ f_2(x,u) \end{pmatrix}$. f_1 and f_2 are assumed to be unknown and only x and u are measurable. In order to do the simulation, we select input as $\delta = 30\sin\frac{t}{2}$, θ is sawtooth wave with amplitude 10, frequency $\frac{1}{2}$, T_d is square wave with amplitude 20, frequency $\frac{1}{4}$, $x_0 = [10, 500]^T$.

Let us select dynamic neural network as

$$\dot{\widehat{x}}_t = A\widehat{x}_t + W_{1,t}\sigma(V_{1,t}\widehat{x}_t) + W_{2,t}\phi(V_{2,t}\widehat{x}_t)\pi(u_t)$$

where $A = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}$, $\widehat{x}_0 = [0, 0]^T$, $W_{1,t}$ and $W_{2,t} \in \Re^{2\times 3}$, $V_{1,t}$ and $V_{2,t} \in \Re^{3\times 2}$. The sigmoid functions are $\sigma(x_i) = \frac{2}{1+e^{-2x_i}} - 0.5$, $\phi(x_i) = \frac{0.2}{1+e^{-0.2x_i}} - 0.05$. $\pi(u_t) = u_t$. $D_\sigma = diag[D_{\sigma_1}, D_{\sigma_2} D_{\sigma_3}]$, $D_\phi = diag[D_{\phi_1}, D_{\phi_2}, D_{\phi_3}]$, $u_3 = 0$,

$$D_{\sigma_i} = \frac{4e^{-2Z_{1,i}}}{(1+e^{-2Z_{1,i}})^2}, \quad Z_{1,i} = (V_{1,t}\widehat{x})_i$$

$$D_{\sigma_i} = \frac{0.04e^{-0.2Z_{2,i}}}{(1+e^{-0.2Z_{2,i}})^2}u_i, \quad Z_{2,i} = (V_{2,t}\widehat{x})_i$$

We select $K_1P = K_2P = K_3P = K_4P = 2I$. The learning law as (25). If we choose $V_1^0 = V_2^0 = \begin{bmatrix} 0.2 & 0.3 \\ 0.2 & 0.3 \\ 0.2 & 0.3 \end{bmatrix}$,. the identification results are shown in Fig.2 and Fig.3

If we choose another $V_1^0 = V_2^0 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$, the dash lines in Fig.4 and Fig.5 correspond to the identification error with these new matrices V_1^0 and V_2^0. So we select $\begin{bmatrix} 0.2 & 0.3 \\ 0.2 & 0.3 \\ 0.2 & 0.3 \end{bmatrix}$ as V_1^0 and V_2^0.If we use signal layer neural networks as

$$\dot{\widehat{x}}_t = A\widehat{x}_t + W_{1,t}\sigma(\widehat{x}_t) + W_{2,t}\phi(\widehat{x}_t)\pi(u_t)$$

all of conditions are the same as multilayer neural networks, the dash lines in Fig.6 and Fig.7 show the compensation of the identification error with single layer and multilayer neural networks. One can see that the multilayer dynamic neural networks are more powerful than single-layer dynamic neural networks, any they are robust with respect to bounded uncertainties.

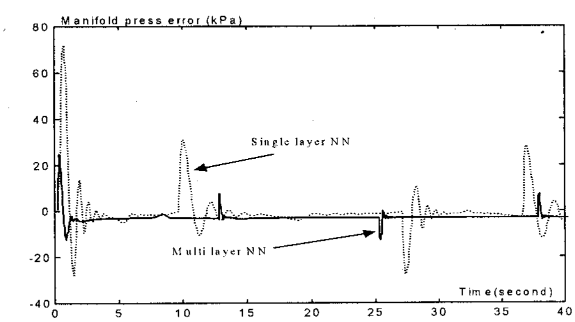

Fig. 6. Manifold press error with single layer and multilayer networks

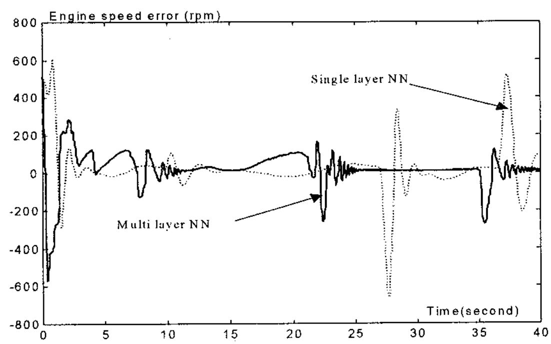

Fig. 7. Engine speed error with single layer and multilayer networks

5. CONCLUSION

By means of passivity technique, we give some new results on neuro identification with multilayer dynamic neural networks . Compared with other stability analysis of neuro identifications, our algorithm is more simple because robust modifications are not applied, so the algorithm proposed in this paper is more suitable for engineering application. We success to prove that even the simple gradient learning algorithm may guarantee the identification error robust stable, and the backpropagation learning algorithm may guarantee the identification error robust stable.

6. REFERENCES

Byrnes C.I., Isidori A., Willems J.C. (1991), Passivity, Feedback Equivalence, and the Global Stabilization of Minimum Phase Nonlinear Systems, *IEEE Trans. Automat. Contr.*, Vol.36, 1228-1240.

S.Commuri and F.L.Lewis (1996), CMAC Neural Networks for Control of Nonlinear Dynamical Systems: Structure, Stability and Passivity, *Automatica*, vol.33, no.4, 635-641.

P.A.Ioannou and J.Sun (1996), *Robust Adaptive Control*, Prentice-Hall, Inc, Upper Saddle River: NJ.

S.Jagannathan and F.L. Lewis (1996), Identification of Nonlinear Dynamical Systems Using Multilayered Neural Networks, *Automatica*, vol.32, no.12, 1707-1712.

K.Kaszkurewics and A.Bhaya (1994), On a Class of Globally Stable Neural Circuits, *IEEE Trans. on Circuit and Systems-I*, Vol. 41, 171-174.

M.Forti, S.Manetti and M.Marini (1994), Necessary and Sufficient Condition for Absolute Stability of Neural Networks, *IEEE Trans. on Circuit and Systems-I*, Vol. 41, 491-494.

E.B.Kosmatopoulos, M.M.Polycarpou, M.A.Christodoulou and P.A.Ioannpu (1995), High-Order Neural Network Structures for Identification of Dynamical Systems, *IEEE Trans. on Neural Networks*, Vol.6, No.2, 442-431.

K.Matsouka (1992), Stability Conditions for Nonlinear Continuous Neural Networks with Asymmetric Connection Weights, *Neural Networks*, Vol.5, 495-500

A.S.Poznyak, W.Yu, E.N.Sanchez and J.P.Perez (1999), Nonlinear Adaptive Trajectory Tracking Using Dynamic Neural Networks, *IEEE Trans. on Neural Networks*, Vol.10, No.5, 1402-1411.

G.V.Puskorius and L.A.Feldkamp (1994), Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks, *IEEE Trans. on Neural Networks*, Vol.5, No.2, 279-297

G.A.Rovithakis and M.A.Christodoulou (1994), Adaptive Control of Unknown Plants Using Dynamical Neural Networks, *IEEE Trans. on Syst., Man and Cybern.*, Vol. 24, 400-412.

E.D.Sontag and Y.Wang (1995), On Characterization of the Input-to-State Stability Property, System & Control Letters, Vol.24, 351-359.

J.A.K.Suykens, J.Vandewalle and B.De Moor (1999), Lur's Systems with Multilayer Perceptron and Recurrent Neural Networks; Absolute Stability and Dissipativity, *IEEE Trans. on Automatic Control*, Vol.44, 770-774.

Wen Yu and Xiaoou Li (2001), Some New Results on System Identification with Dynamic Neural Networks, *IEEE Trans. Neural Networks*, Vol.12, No.2, 412-417

Wen Yu (2001), Some Stability Properties of Dynamic Neural Networks, *IEEE Trans. Circuits and Systems, Part I*, Vol.48, No.1, 256-259

Wen Yu (2002), Passivity Analysis for Dynamic Multilayer Neuro Identifier, *Automatica*, accepted for publication.

Wen Yu, *was born in Shenyang, China, in 1966. He received the B.S. degree from Tsinghua University, China in 1990 and the M.S. and Ph.D. degrees, both from Northeastern University, China, in 1992 and 1995, respectively. From 1995 to 1996, he served as a Lecture in Northeastern University, China. In 1996, he joined CINVESTAV-IPN, where he is a professor in the Departamento de Control Automatico. His research interests include neural networks, visual servoing and hybrid systems.*