

# CONVERGÊNCIA DO ERRO DE PREDIÇÃO EM IDENTIFICAÇÃO NEURAL SUJEITA A DISTÚRBIOS E PARÂMETROS VARIANTES NO TEMPO

José A. Ruiz Vargas

Elder M. Hemerly

Departamento de Sistemas e Controle, Instituto Tecnológico de Aeronáutica  
12228-900 São José dos Campos, São Paulo, Brasil  
vargas@ieee.org hemerly@ita.br

**Resumo** – Neste artigo é proposto um algoritmo para identificação neural de sistemas contínuos não-lineares incertos com parâmetros variantes no tempo e sujeitos a distúrbios. Prova-se, via argumentos usuais de Lyapunov e uma técnica de limitação adaptativa (*adaptive bounding technique*), que o erro de predição converge para zero, inclusive na presença de erros de aproximação e distúrbios, enquanto os outros erros associados permanecem limitados. Simulações são apresentadas para ilustrar a aplicação e desempenho do algoritmo proposto.

**Abstract** – In this paper an algorithm to identify uncertain continuous nonlinear systems in the presence of time-varying parameters and non-vanishing disturbances is proposed. Using standard Lyapunov procedures and an adaptive bounding technique the prediction error convergence to zero is proved, even when approximation error and disturbances are present, while guaranteeing uniform ultimate boundedness of all others associate estimation errors. A simulation study to illustrate the application and performance of the proposed algorithm is provided.

**Palavras Chaves** – Identificação, sistemas não-lineares, redes neurais artificiais, métodos de Lyapunov.

## 1 Introdução

A identificação e modelagem de sistemas como metodologia de análise científica data de Galileo (1564-1642), que também é considerado fundador da Dinâmica. A partir daí a identificação e modelagem de sistemas têm seguido diferentes caminhos com igual força na matemática, estatística, ciência da computação, controle de sistemas dinâmicos e processamento de sinais entre outras [1]. Entretanto, a identificação de sistemas não-lineares está longe, em geral, de ser um problema bem resolvido: representa um problema conjunto de determinação de estrutura e estimação de parâmetros [2], onde a presença de incertezas é inevitável. Isto porque é impossível se obter modelos para identificação conhecidos perfeitamente, devido às simplificações necessárias para a modelagem matemática, falhas intempestivas, mudanças nas condições de operação, envelhecimento dos equipamentos e assim por diante.

Uma abordagem relativamente recente e promissora para tratar o problema de identificação de sistemas dinâmicos não-lineares baseia-se no emprego de redes neurais artificiais (RNAs). As principais características que fundamentam o emprego de RNAs em identificação são: 1) RNAs têm estrutura conhecida, 2) RNAs podem aproximar mapeamentos entrada-saída complexos via aprendizagem, sempre que satisfeitas as hipóteses de aproximação universal e regularidade [3], 3) RNAs são adequadas ao processamento paralelo e exibem tolerância a falhas. Portanto, mediante o emprego de RNAs o problema de determinação de estrutura é simplificado, sendo o projeto do preditor resumido principalmente ao projeto de leis de adaptação, o que pode ser feito com base na análise de estabilidade. Como exemplos podem ser citados [3]-[11], onde foi considerada a identificação de varias classes de sistemas incertos com parâmetros constantes e desconhecidos, sendo os pesos dos modelos neurais empregados ajustados com leis de adaptação projetadas com base na análise de estabilidade [3]-[8] e passividade [9]-[11]. Em todos estes trabalhos, contudo, unicamente foi provado que os erros de predição residuais são limitados e, em geral, de magnitude proporcional a limitantes superiores dos erros de aproximação e distúrbios.

Em particular, em [3], [4] e [8] foram utilizadas RNAs parametrizáveis linearmente para modelar o sistema incerto e, com base nesta parametrização e argumentos usuais de Lyapunov, projetar algoritmos de adaptação para os pesos que asseguraram a limitação dos sinais de erro. Por exemplo em [3], RNAs de base radial e o método direto de Lyapunov foram empregados para garantir a convergência do erro de estado para uma vizinhança da origem. Já em [4] e [8] foram utilizadas RNAs de alta ordem para melhorar o desempenho da identificação, pois as redes de alta ordem apresentam uma capacidade de aproximação superior àquela de base radial, como reportado em [5]. A principal contribuição de [8] foi a proposição de algoritmos de aprendizado genéricos para RNAs de alta ordem, que incluíram como casos particulares alguns outros algoritmos da literatura, como por exemplo aqueles em [3] e [4]. Mais recentemente em [9] e [11] foram propostos algoritmos com base em argumentos de passividade e baseados em RNAs parametrizáveis linearmente. Nesses trabalhos foi mostrado que modificações robustas do algoritmo do gradiente não são necessárias para assegurar a convergência do erro de estimação do estado para uma vizinhança da origem na presença de erros limitados de modelagem. Contudo, a principal carência destes trabalhos, como também em [3], [4] e [8], reside na impossibilidade de assegurar que o erro de predição converge para zero na presença de erros de aproximação e distúrbios limitados. Adicionalmente os algoritmos nesses trabalhos são conservadores,

pois não é considerado no projeto dos mesmos a possível presença de parâmetros variantes no tempo. Algoritmos baseados em RNAs parametrizáveis não-linearmente, como [5]-[7] e [10]-[11] também padecem das limitações supracitadas.

A não garantia de convergência do erro de predição na presença de distúrbios não é um problema peculiar da identificação neural, já que inclusive na estimação paramétrica *on-line* de sistemas lineares não existem leis de adaptação robustas que permitam assegurar a convergência do erro residual de predição para zero na presença de erros de modelagem e distúrbios. Vide [12] para maiores detalhes.

Como já mencionado, a principal deficiência de [3]-[11] consiste na não garantia da convergência do erro de predição na presença de erros de aproximação e distúrbios. Neste sentido, em [13]-[14] foram propostos, recentemente, algoritmos baseados em RNAs parametrizáveis linearmente que asseguraram, ao contrário de [3]-[11], a convergência assintótica do erro de predição para zero, inclusive na presença de erros de aproximação e distúrbios. Nestes trabalhos a convergência assintótica foi alcançada introduzindo modelos para identificação neural convenientes e leis de adaptação robustas baseadas em *leakage gains* [12] dinâmicos. Contudo, em [13]-[14] foram impostas condições que restringem a aplicação e não foi considerada a presença de parâmetros variantes no tempo, o qual constitui a principal motivação do presente trabalho.

Mais precisamente, neste artigo é proposto um algoritmo baseado em RNAs para identificação de uma classe genérica de sistemas não-lineares nos quais os parâmetros e distúrbios são limitados, desconhecidos e variantes no tempo. Para tanto, foram projetados, com base na análise de estabilidade e usando uma técnica de limitação adaptativa [15], um modelo para identificação e leis de adaptação robustas para seus pesos para garantir a convergência do erro de predição para zero, e limitação dos outros erros de estimação associados. O algoritmo proposto estende os resultados prévios em [3]-[11], uma vez que é assegurada a convergência do erro de predição na presença de distúrbios, e os em [13]-[14], já que possibilita a aplicação, ao contrário de [13]-[14], a sistemas com parâmetros variantes no tempo, sempre que algumas condições sobre os parâmetros de projeto sejam satisfeitas. Convém ressaltar que, mesmo no caso de parâmetros constantes, o algoritmo proposto assume hipóteses de aplicação menos restritivas que aquelas em [13]-[14].

A contribuição deste trabalho é dupla: por um lado propõe-se uma metodologia para identificação de sistemas incertos genéricos que assegura a convergência assintótica do erro de predição, apesar da presença de parâmetros variantes no tempo, erros de aproximação e distúrbios. Por outro lado, propõe-se também uma metodologia para identificação que assegura a convergência exponencial do erro de predição para uma bola em torno da origem, cujo raio pode ser previamente especificado pelo usuário, uma vez que o algoritmo proposto, com mínimas alterações, também pode ser utilizado para este propósito. Portanto, o presente trabalho estende o estado da arte em identificação baseada em RNAs e, em particular, em identificação de sistemas lineares, já que a metodologia proposta pode também ser usada sem RNAs em uma forma similar a [12] no contexto de estimação paramétrica *on-line*.

O restante do trabalho é organizado conforme a seguir: na seção 2 são apresentadas as classes de RNAs utilizadas na parametrização das não-linearidades desconhecidas. Na seção 3 o problema de identificação é formulado. Já na seção 4 o modelo para identificação e os erros associados são apresentados. Na seção 5 é apresentado o principal resultado do artigo, a prova de estabilidade assintótica do erro de predição. Um exemplo de simulação é discutido na seção 6. Finalmente, na seção 7, as principais contribuições do trabalho são resumidas.

## 2 Redes neurais artificiais parametrizáveis linearmente

As redes neurais parametrizáveis linearmente (RNAPLs) podem ser expressas matematicamente como

$$\rho_{nn}(W, \zeta) = W\pi(\zeta) \quad (1)$$

onde  $W \in \mathcal{R}^{n \times L_\rho}$ ,  $\zeta \in \mathcal{R}^{L_\zeta}$  e  $\pi: \mathcal{R}^{L_\zeta} \mapsto \mathcal{R}^{L_\rho}$  é um vetor de funções básicas que pode ser considerado como uma função vetorial não-linear, cujos argumentos são pré-processados por uma função escalar  $s(\cdot)$ , e  $n, L_\rho, L_\zeta$  são inteiros estritamente positivos. As funções escalares  $s(\cdot)$  usualmente utilizadas incluem sigmoide, tanh, gaussiana, *Hardy's*, *inverse Hardy's multiquadratic*, etc [16]. Entretanto, neste artigo estamos interessados somente na classe de RNAPLs nas quais  $s(\cdot)$  é limitado, uma vez que neste caso tem-se,

$$\|\pi(\zeta)\| \leq \pi_0 \quad (2)$$

sendo  $\pi_0$  uma constante estritamente positiva.

A classe de RNAPLs consideradas neste trabalho inclui HONNs [4], RBFNNs [17], *wavelet networks* [18], e também

outros aproximadores parametrizáveis linearmente como *Takagi-Sugeno fuzzy systems* [19], os quais satisfazem a propriedade de aproximação universal e regularidade [3]:

(P1) Para quaisquer constante  $\varepsilon_0 > 0$  e função  $f \in C[\Omega]$  existe uma matriz  $W^* = W$ , onde  $W^*$  é uma matriz “ótima” e  $L_\rho$  é suficientemente grande, tal que,

$$\sup_{\zeta \in \Omega} |f(\zeta) - W^* \pi(\zeta)| \leq \varepsilon_0 \quad (3)$$

onde  $C[\Omega]$  denota o espaço de funções contínuas definidas em um domínio compacto  $\Omega \subset \mathfrak{R}^n$ ,  $|\cdot|$  denota o valor absoluto se o argumento é um escalar. Se o argumento é uma função vetorial em  $\mathfrak{R}^n$  então  $|\cdot|$  denota qualquer norma em  $\mathfrak{R}^n$ .

(P2) A saída da RNA é contínua com respeito aos seus argumentos e satisfaz a condição de Lipschitz localmente para todo  $(\zeta, W)$  finito.

### 3 Formulação do problema

Considere a classe de sistemas não-lineares genéricos representados por

$$\dot{x} = F(x, u, \rho, t) \quad (4)$$

onde  $x \in X$  é o vetor de estados de dimensão  $n$ ,  $u \in U$  é um vetor de entradas admissíveis de dimensão  $m$ ,  $\rho \in V$  é um vetor de parâmetros (variante no tempo e incerto) de dimensão  $p$ ,  $F : X \times U \times V \times [0, \infty) \mapsto \mathfrak{R}^n$  é um mapeamento contínuo desconhecido. Com a finalidade de ter o problema bem colocado, assuma que  $X, U, V$  são conjuntos compactos e  $F$  é Lipschitz localmente com respeito a  $x$  em  $X \times U \times V \times [0, \infty)$ , tal que (4) tem uma única solução passando por  $x(0)$ .

Assuma-se que o seguinte pode ser estabelecido

**Hipótese 1:** Na região  $X \times U \times V \times [0, \infty)$

$$\|h(x, u, \rho, t)\| \leq h_0 \quad (5)$$

onde

$$h(x, u, \rho, t) = F(x, u, \rho, t) - f(x, u) - \sum_{i=1}^p g_i(x, u) \rho_i(t) \quad (6)$$

$f : X \times U \mapsto \mathfrak{R}^n$  e  $g_i : X \times U \mapsto \mathfrak{R}^n, \forall i = 1, 2, \dots, p$  são mapeamentos desconhecidos,  $\rho(t) = [\rho_1(t) \ \rho_2(t) \ \dots \ \rho_p(t)]^T$  é tal que  $\rho \in C^1$ , isto é,  $\rho$  possui primeira derivada temporal contínua,  $\|\cdot\|$  é a norma Euclideana e  $h_0$  é uma constante positiva.

**Hipótese 2:** Os parâmetros  $\rho_i(t)$  têm sinal constante conhecido e  $d\rho_i^2(t)/dt \leq 0 \ \forall i = 1, 2, \dots, p$ .

Portanto, excetuando-se pelas hipóteses 1 e 2, diz-se que  $F(x, u, v, t)$  é um mapeamento desconhecido.

**Comentário 1:** A hipótese 1, a qual impõe um limitante superior sobre a norma dos distúrbios, é completamente usual em identificação. A hipótese 2 impõe que a “energia” de  $\rho$  seja constante ou continuamente dissipada, o qual implica

$$\|\rho(t)\| \leq \beta_0 \quad (7)$$

para alguma constante positiva  $\beta_0$ .

Objetiva-se projetar um preditor *on-line* para (4), baseado em RNAPLs, o qual assegure a convergência assintótica do erro residual de predição, apesar da presença de erros de aproximação e distúrbios limitados.

Convém ressaltar que a identificação de sistemas dinâmicos usando RNAs é relevante não somente para prever o comportamento de sistemas incertos, mas também porque fornece uma parametrização que pode ser usada no projeto de sistemas de controle, pois em geral, não é possível se projetar um controlador de alto desempenho sem que se tenha alguma informação prévia sobre a dinâmica do sistema. Este requisito aparentemente trivial pode não estar satisfeito em uma ampla classe de problemas de interesse prático, e é nestes casos que também é importante a identificação baseada em RNAs.

#### 4 Modelo para identificação e equação do erro de predição do estado

Nesta seção é proposto o modelo para identificação e, com base no mesmo, é obtida a equação do erro de predição associado com o problema.

Com base em (3), usando-se RNAPLs, os mapeamentos  $f(x,u)$  e  $g_i(x,u)$  podem ser substituídos por  $W_f^* \pi_f(x,u)$  e  $W_{gi}^* \pi_{gi}(x,u)$  mais os erros de aproximação  $\varepsilon_f(x,u)$  e  $\varepsilon_{gi}(x,u)$ . Mais exatamente, (4) pode ser escrita como

$$\dot{x} = W_f^* \pi_f(x,u) + \sum_{i=1}^p W_{gi}^* \pi_{gi}(x,u) \rho_i + \varepsilon_f(x,u) + \sum_{i=1}^p \varepsilon_{gi}(x,u) \rho_i + h(x,u, \rho, t) \quad (8)$$

onde  $W_f^* \in \mathfrak{R}^{n \times L_f}$  e  $W_{gi}^* \in \mathfrak{R}^{n \times L_{gi}}$  são matrizes “ótimas”, requeridas somente para propósitos analíticos, que pode ser definidas como

$$W_f^* := \arg \min_{\hat{W}_f \in \Gamma_f} \left\{ \sup_{\substack{x \in X, \\ u \in U}} |f(x,u) - \hat{W}_f \pi_f(x,u)| \right\}, \quad W_{gi}^* := \arg \min_{\hat{W}_{gi} \in \Gamma_{gi}} \left\{ \sup_{\substack{x \in X, \\ u \in U}} |g_i(x,u) - \hat{W}_{gi} \pi_{gi}(x,u)| \right\} \quad (9)$$

com  $\Gamma_f = \{ \hat{W}_f \mid \|\hat{W}_f\|_F \leq \alpha_{\hat{W}_f} \}$ ,  $\Gamma_{gi} = \{ \hat{W}_{gi} \mid \|\hat{W}_{gi}\|_F \leq \alpha_{\hat{W}_{gi}} \}$ ,  $\|\cdot\|_F$  é a norma de Frobenius,  $\alpha_{\hat{W}_f}$  e  $\alpha_{\hat{W}_{gi}}$  são constantes estritamente positivas, as quais dependem das leis de ajuste dos pesos,  $\hat{W}_f$  e  $\hat{W}_{gi}$  são estimações de  $W_f^*$  e  $W_{gi}^*$ , respectivamente. Os erros de aproximação  $\varepsilon_f(x,u)$  e  $\varepsilon_{gi}(x,u)$ , correspondentes a  $W_f^*$  e  $W_{gi}^*$ , podem ser definidos como

$$\varepsilon_f(x,u) := f(x,u) - W_f^* \pi_f(x,u), \quad \varepsilon_{gi}(x,u) := g_i(x,u) - W_{gi}^* \pi_{gi}(x,u) \quad (10)$$

Uma vez que  $X, U$  são conjuntos compactos e usando-se (2), na região  $X \times U$  tem-se

$$\|\varepsilon_f(x,u)\| \leq \varepsilon_{f0}, \quad \|\varepsilon_{gi}(x,u)\| \leq \varepsilon_{gi0} \quad (11)$$

onde  $\varepsilon_{f0}$  e  $\varepsilon_{gi0}$  são constantes positivas.

**Comentário 2:** É importante ressaltar que  $W_f^*$  e  $\varepsilon_f(x,u)$  ( $W_{gi}^*$  e  $\varepsilon_{gi}(x,u)$ ) podem não ser únicos. Entretanto,  $\|\varepsilon_f(x,u)\|$  ( $\|\varepsilon_{gi}(x,u)\|$ ) é único por (9).

A estrutura (8) sugere um modelo para identificação da forma

$$\dot{\hat{x}} = \hat{W}_f \pi_f(\xi, u) + \sum_{i=1}^p \hat{W}_{gi} \pi_{gi}(\xi, u) \hat{\rho}_i - L\tilde{x} - l(\tilde{x}, \hat{\psi}, t) \quad (12)$$

onde  $\hat{x}$  é o estado estimado,  $\xi = x$  ou  $\hat{x}$ ,  $\hat{\rho}$  é o vetor de parâmetros estimado,  $\tilde{x} = \hat{x} - x$  é o erro de predição,  $L \in \mathfrak{R}^{n \times n}$  é uma matriz positiva definida de ganho de realimentação introduzida para atenuar os efeitos de incertezas e condição inicial  $x(0)$  não nulas,  $\hat{\psi}$  e  $l$  são, respectivamente, uma função escalar e vetorial, as quais serão definidas apropriadamente na próxima seção (vide (18) e (22)). Mostrar-se-á na próxima seção que o modelo para identificação (12), com mecanismos de adaptação adequadamente projetados para  $\hat{W}_f$ ,  $\hat{W}_{gi}$ ,  $\hat{\rho}$ ,  $\hat{\psi}$  e  $l$ , assegura a convergência do erro residual de predição para zero inclusive na presença de erros de aproximação e distúrbios.

**Comentário 3:** Embora o modelo (12) seja relevante do ponto de vista de identificação, pois possibilita a obtenção de modelos de erro, há uma deficiência no que se refere à predição, devido a que (12) pode unicamente trabalhar *on-line* e os pesos podem não convergir para os “ótimos”. Entretanto, esta deficiência não é peculiar, exclusivamente, ao modelo proposto, senão à maioria de modelos baseados em RNAs [5]. A disponibilidade do estado, portanto, não reduz a importância do método, uma vez que sem (12) não seria possível obter modelos de erro, que são usados na prova de estabilidade e convergência. Procedimento similar pode ser encontrado, por exemplo, em [12], no contexto do problema de estimação *on-line*. Com base em (8) e (12), obtém-se a equação do erro de predição

$$\dot{\tilde{x}} = -L\tilde{x} + \tilde{W}_f \pi_f(\xi, u) + \sum_{i=1}^p \tilde{W}_{gi} \pi_{gi}(\xi, u) \hat{\rho}_i + \sum_{i=1}^p W_{gi}^* \pi_{gi}(\xi, u) \tilde{\rho}_i + \omega(\xi, x, u, \rho, t) - l(\tilde{x}, \hat{\psi}, t) \quad (13)$$

onde  $\tilde{W}_f = \hat{W}_f - W_f^*$ ,  $\tilde{W}_{gi} = \hat{W}_{gi} - W_{gi}^*$ ,  $\tilde{\rho}_i = \hat{\rho}_i - \rho_i$  e

$$\omega(\xi, x, u, \rho, t) = W_f^* [\pi_f(\xi, u) - \pi_f(x, u)] + \sum_{i=1}^p W_{gi}^* [\pi_{gi}(\xi, u) - \pi_{gi}(x, u)] \rho_i - \varepsilon_f(x, u) - \sum_{i=1}^p \varepsilon_{gi}(x, u) \rho_i - h(x, u, \rho, t)$$

Com base em (5), (7), (11) e (P2), o termo de distúrbio  $\omega$  pode ser limitado por

$$\|\omega(\xi, x, u, \rho, t)\| \leq \omega_0 + \omega_1 \|\tilde{x}\| \quad (14)$$

onde  $\omega_0 = \varepsilon_{f0} + \beta_0 \sum_{i=1}^p \varepsilon_{gi0} + h_0$ ,  $\omega_1 = \gamma_f \|W_f^*\|_F + \beta_0 \sum_{i=1}^p \|W_{gi}^*\|_F \gamma_{gi}$ ,  $\gamma_f$  e  $\gamma_{gi}$  são constantes de Lipschitz maiores ou iguais a zero.

**Comentário 4:** Convém ressaltar que  $\omega_1 = 0$  se  $\xi = x$ , e  $\omega_0 = 0$  na ausência de erros de aproximação e distúrbios. Entretanto, em geral, as mesmas são desconhecidas *a priori*, uma vez que dependem de limitantes para os erros de aproximação, pesos nominais e distúrbios, os quais são desconhecidos. Contudo, mesmo sendo desconhecidas, sempre é possível selecionar  $\bar{\omega}_0 > \omega_0$  e  $\bar{\omega}_1 > \omega_1$ , possivelmente grandes, tal que (14) é satisfeita. Neste trabalho assume-se que tais limitantes são conhecidos *a priori*.

## 5 Leis de adaptação para os pesos e análise de estabilidade

Nesta seção, objetivando-se assegurar a convergência do erro de predição, são projetadas as leis de adaptação para os pesos e função de realimentação  $l$  em (12). Em trabalhos prévios, vide [3]-[11] para maiores detalhes, as leis de adaptação são projetadas com base em argumentos usuais de Lyapunov para assegurar limitação dos erros de estimação e estabilidade prática do erro de predição na presença de distúrbios, isto é, convergência do erro de predição para uma bola em torno da origem. Em [13] foram introduzidos uma função de realimentação e leis de adaptação baseadas em *leakage gain* dinâmicos para assegurar, adicionalmente, a convergência assintótica do erro residual de predição para zero. Neste trabalho empregam-se argumentos usuais de Lyapunov, como em [3]-[11], e uma técnica de limitação adaptativa [15], para assegurar estabilidade assintótica, similarmente a [14], porém relaxando importantes hipóteses de aplicação necessárias em [13] e [14]: parâmetros constantes e conhecimento prévio de uma região para os pesos “ótimos”. Basicamente, a técnica de limitação adaptativa consiste em projetar uma função limitadora que domine termos positivos em  $\dot{V}$ , a derivada em relação ao tempo de uma candidata a função de Lyapunov  $V$ , para fazer  $\dot{V}$  negativa semi-definida em todo o espaço dos erros.

Inicialmente, objetivando-se provar que os erros estimação são *ultimately bounded* [12], considere o vetor composto de erros

$$\tilde{\chi} = \left[ \tilde{x}^T \quad \text{vec}(\tilde{W}_f)^T \quad \text{vec}(\tilde{W}_1)^T \cdots \text{vec}(\tilde{W}_p)^T \quad \tilde{\psi} \quad \tilde{\rho}^T \right]^T \quad (15)$$

e a maior bola na região de operação

$$B_R = \{ \tilde{\chi} \mid \| \tilde{\chi} \| \leq R \} \quad (16)$$

tal que  $\forall \tilde{\chi} \in B_R$  tem-se  $(x, u) \in X \times U$ , onde  $\tilde{\psi} = \hat{\psi} - \psi^*$ ,  $\psi^*$  e  $R$  são constantes estritamente positivas. Então, a seguinte hipótese é necessária

**Hipótese 3:** Assuma que

$$R > \gamma_0 \sqrt{\frac{T_M}{T_m}} \geq \gamma_0 \quad (17)$$

onde

$$T_M = \frac{1}{2} \max [2\lambda_{\max}(P), \alpha_f^{-1}, \alpha_{gi}^{-1}, \alpha_{\psi}^{-1}, \alpha_{\rho}^{-1}, \quad i=1,2,\dots,p], \quad T_m = \frac{1}{2} \min [2\lambda_{\min}(P), \alpha_f^{-1}, \alpha_{gi}^{-1}, \alpha_{\psi}^{-1}, \alpha_{\rho}^{-1}, \quad i=1,2,\dots,p],$$

$$\gamma_0 = \max \left( \frac{\psi_1^2}{4\alpha_1\psi_0}, \frac{\psi_1}{2\sqrt{\psi_0 c_{f \min}}}, \frac{\psi_1}{2\sqrt{\psi_0 c_{gi \min}}}, \frac{\psi_1}{2\psi_0}, \frac{\psi_1}{2\sqrt{\psi_0 (c_{\rho \min} - 1/2 - \alpha_3^2/2)}} \right), \quad i=1,2,\dots,p,$$

$P \in \mathfrak{R}^{n \times n} > 0$ ,  $\lambda_{\min}(\Gamma)$  ( $\lambda_{\max}(\Gamma)$ ) é o menor (maior) autovalor de uma matriz quadrada  $\Gamma$ ,  $\alpha_f, \alpha_{gi}, \alpha_{\psi}, \alpha_{\rho}, \psi_0, \psi_1$

são constantes estritamente positivas,  $\alpha_1 = \lambda_{\min}(Q)$ ,  $Q = L^T P + PL$ ,  $\alpha_3 = \|K\|_F \left( \sum_{i=1}^p \gamma_{gi}^2 \|W_{gi}^*\|_F^2 \right)^{\frac{1}{2}}$ ,  $K = P^T + P$ ,

$c_{f \min} = \lambda_{\min}(C_f)$ ,  $C_f = \text{diag}(c_{ff}, j=1,2,\dots,n)$ ,  $c_{gi \min} = \lambda_{\min}(C_{gi})$ ,  $C_{gi} = \text{diag}(c_{gij}, j=1,2,\dots,n)$ ,  $c_{gij} > 0$ ,  $c_{\rho \min} = \lambda_{\min}(C_{\rho})$ ,  $C_{\rho} = \text{diag}(c_{\rho i}, i=1,2,\dots,p)$  e  $c_{\rho i}$  são constantes escolhidas com base em (25).

O resultado principal do trabalho é apresentado na seqüência.

**Teorema 1:** Considere a classe de sistemas não-lineares genéricos descritos por (4), satisfazendo as hipóteses 1-3, o modelo para identificação (12) com

$$l(\tilde{x}, \hat{\psi}, t) = \left[ \frac{l_0 \hat{\psi}}{\| \tilde{x} \| + z(t)} + l_1 \| \tilde{x} \| \right] \tilde{x} \quad (18)$$

$$z(t) = z_0 \exp(-z_1 t) \quad (19)$$

e as leis adaptativas,

$$\dot{W}_f = -\alpha_f [2C_f \| \tilde{x} \| (\hat{W}_f - W_{f0}) + K \tilde{x} \pi_f^T(\xi, u)] \quad (20)$$

$$\dot{W}_{gi} = -\alpha_{gi} [2C_{gi} \| \tilde{x} \| (\hat{W}_{gi} - W_{gi0}) + K \tilde{x} \pi_{gi}^T(\xi, u) \hat{\rho}_i], \quad i=1,2,\dots,p \quad (21)$$

$$\dot{\hat{\psi}} = -\alpha_{\psi} (2\psi_0 \| \tilde{x} \| \hat{\psi} - \psi_1 \| \tilde{x} \|), \quad \hat{\psi}(0) > 0 \quad (22)$$

$$\dot{\hat{\rho}}_i = -2\alpha_\rho c_{\rho i} \|\tilde{x}\| \hat{\rho}_i, \quad \hat{\rho}_i(0) \begin{cases} > 0 & \text{se } \text{sign}(\rho_i) = -1 \\ < 0 & \text{se } \text{sign}(\rho_i) = 1 \end{cases} \quad (23)$$

Se  $\tilde{\chi}(0) \in \Omega_m$  e

$$\psi_1 \geq 2\sqrt{\alpha_0 \psi_0} \quad (24)$$

$$c_{\rho \min} > 1/2 + \alpha_3^2/2 \quad (25)$$

Então,

(i)  $\tilde{\chi}(t)$  é *ultimately bounded*.

(ii)  $\lim_{t \rightarrow \infty} \tilde{\chi}(t) = 0$ .

Onde  $l_0 = 2\psi_1/\lambda_{\min}(K)$ ,  $l_1 = 1/\lambda_{\min}(K)$ ,  $z_0$  e  $z_1$  são constantes estritamente positivas,  $W_{f0} \in \mathfrak{R}^{n \times L_f}$ ,  $W_{gi0} \in \mathfrak{R}^{n \times L_{gi}}$ ,  $\Omega_m = \{\tilde{\chi} \mid V \leq V_m\}$ ,  $V_m = \min_{\|\tilde{z}\|=R} V = R^2 T_m$ ,  $V$  é definido a seguir em (26),

$$\alpha_0 = (\bar{\omega}_0 + \bar{\omega}_1^2 \|K\|_F / 2) \|K\|_F + c_{f \max} \|W_f^* - W_{f0}\|_F^2 + \sum_{i=1}^p c_{gi \max} \|W_{gi}^* - W_{gi0}\|_F^2 + \alpha_2^2 / 2 + \beta_0^2 \sum_{i=1}^p c_{gi} \quad ,$$

$$c_{f \max} = \lambda_{\max}(C_f), \quad c_{gi \max} = \lambda_{\max}(C_{gi}), \quad \alpha_2 = \|K\|_F \left( \sum_{i=1}^p \pi_{gi0}^2 \|W_{gi}^*\|_F^2 \right)^{\frac{1}{2}} \text{ e } \pi_{gi0} = \sup_{t \geq 0} \|\pi_{gi}(x(t), u(t))\|.$$

### Prova:

Considere a seguinte candidata a função de Lyapunov

$$V = \tilde{x}^T P \tilde{x} + \text{tr}(\tilde{W}_f^T \alpha_f^{-1} \tilde{W}_f) / 2 + \sum_{i=1}^p \text{tr}(\tilde{W}_{gi}^T \alpha_{gi}^{-1} \tilde{W}_{gi}) / 2 + \tilde{\psi} \alpha_\psi^{-1} \tilde{\psi} / 2 + \sum_{i=1}^p \tilde{\rho}_i \alpha_\rho^{-1} \tilde{\rho}_i / 2 \quad (26)$$

onde  $\text{tr}(\cdot)$  é o operador traço.

Derivando-se (26) em relação ao tempo, obtém-se

$$\dot{V} = \tilde{x}^T P \dot{\tilde{x}} + \dot{\tilde{x}}^T P \tilde{x} + \text{tr}(\tilde{W}_f^T \alpha_f^{-1} \dot{\tilde{W}}_f) + \sum_{i=1}^p \text{tr}(\tilde{W}_{gi}^T \alpha_{gi}^{-1} \dot{\tilde{W}}_{gi}) + \tilde{\psi} \alpha_\psi^{-1} \dot{\tilde{\psi}} + \sum_{i=1}^p \tilde{\rho}_i \alpha_\rho^{-1} \dot{\tilde{\rho}}_i \quad (27)$$

Substituindo-se (13) e (20)-(23) em (27), advém

$$\begin{aligned}
 \dot{V} = & -\tilde{x}^T Q \tilde{x} + \tilde{x}^T K \tilde{W}_f \pi_f(\xi, u) \\
 & + \tilde{x}^T K \sum_{i=1}^p \tilde{W}_{gi} \pi_{gi}(\xi, u) \hat{\rho}_i + \tilde{x}^T K \sum_{i=1}^p W_{gi}^* \pi_{gi}(\xi, u) \tilde{\rho}_i \\
 & + \tilde{x}^T K \omega - \tilde{x}^T K l \\
 & - \|\tilde{x}\| \operatorname{tr} [2\tilde{W}_f^T C_f (\hat{W}_f - W_{f0})] - \operatorname{tr} [\tilde{W}_f^T K \tilde{x} \pi_f^T(\xi, u)] \\
 & - \|\tilde{x}\| \sum_{i=1}^p \operatorname{tr} [2\tilde{W}_{gi}^T C_{gi} (\hat{W}_{gi} - W_{gi0})] - \sum_{i=1}^p [\tilde{W}_{gi}^T K \tilde{x} \pi_{gi}^T(\xi, u) \hat{\rho}_i] \\
 & - 2\psi_0 \|\tilde{x}\| \tilde{\psi} \hat{\psi} + \psi_1 \tilde{\psi} \|\tilde{x}\| \\
 & - \|\tilde{x}\| \sum_{i=1}^p 2c_{\rho_i} \tilde{\rho}_i \hat{\rho}_i - \sum_{i=1}^p \tilde{\rho}_i \alpha_{\rho_i}^{-1} \dot{\rho}_i
 \end{aligned} \tag{28}$$

onde o último termo no lado direito de (28), isto é,  $-\sum_{i=1}^p \tilde{\rho}_i \alpha_{\rho_i}^{-1} \dot{\rho}_i$  é menor ou igual a zero, já que  $-\sum_{i=1}^p \hat{\rho}_i \alpha_{\rho_i}^{-1} \dot{\rho}_i \leq 0$

devido à hipótese 2 e (23), e  $\sum_{i=1}^p \rho_i \alpha_{\rho_i}^{-1} \dot{\rho}_i \leq 0$  pela hipótese 2.

Uma vez que usando uma propriedade simples do traço resulta,

$$\operatorname{tr} [\tilde{W}_f^T K \tilde{x} \pi_f^T(\xi, u)] = \tilde{x}^T K \tilde{W}_f \pi_f(\xi, u) \tag{29}$$

$$\sum_{i=1}^p [\tilde{W}_{gi}^T K \tilde{x} \pi_{gi}^T(\xi, u) \hat{\rho}_i] = \sum_{i=1}^p \tilde{x}^T K \tilde{W}_{gi} \pi_{gi}(\xi, u) \hat{\rho}_i \tag{30}$$

substituindo-se (29) e (30) em (28), decorre

$$\begin{aligned}
 \dot{V} \leq & -\tilde{x}^T Q \tilde{x} + \tilde{x}^T K \sum_{i=1}^p W_{gi}^* \pi_{gi}(\xi, u) \tilde{\rho}_i + \tilde{x}^T K \omega - \tilde{x}^T K l \\
 & - \|\tilde{x}\| \operatorname{tr} [2\tilde{W}_f^T C_f (\hat{W}_f - W_{f0})] - \|\tilde{x}\| \sum_{i=1}^p \operatorname{tr} [2\tilde{W}_{gi}^T C_{gi} (\hat{W}_{gi} - W_{gi0})] \\
 & - 2\psi_0 \|\tilde{x}\| \tilde{\psi} \hat{\psi} + \psi_1 \tilde{\psi} \|\tilde{x}\| - \|\tilde{x}\| \sum_{i=1}^p 2c_{\rho_i} \tilde{\rho}_i \hat{\rho}_i
 \end{aligned} \tag{31}$$

Por outro lado, vide [14] para maiores detalhes, tem-se

$$\operatorname{tr} [2\tilde{W}_f^T C_f (\hat{W}_f - W_{f0})] \geq c_{f \min} \|\tilde{W}_f\|_F^2 - c_{f \max} \|W_f^* - W_{f0}\|_F^2 \tag{32}$$

$$\sum_{i=1}^p \operatorname{tr} [2\tilde{W}_{gi}^T C_{gi} (\hat{W}_{gi} - W_{gi0})] \geq \sum_{i=1}^p [c_{gi \min} \|\tilde{W}_{gi}\|_F^2 - c_{gi \max} \|W_{gi}^* - W_{gi0}\|_F^2] \tag{33}$$

$$2\tilde{\psi} \hat{\psi} = \tilde{\psi}^2 + \hat{\psi}^2 - \psi^{*2} \tag{34}$$

$$\sum_{i=1}^p 2c_{\rho_i} \tilde{\rho}_i \hat{\rho}_i = \sum_{i=1}^p c_{\rho_i} (\tilde{\rho}_i^2 + \hat{\rho}_i^2 - \rho_i^2) \tag{35}$$

Adicionalmente, com base em (P2), (14) e efetuando-se simples manipulações, resulta

$$\tilde{x}^T K \sum_{i=1}^p W_{gi}^* \pi_{gi}(\xi, u) \tilde{\rho}_i \leq \alpha_2 \|\tilde{\rho}\| \|\tilde{x}\| + \alpha_3 \|\tilde{\rho}\| \|\tilde{x}\|^2 \quad (36)$$

$$\tilde{x}^T K \omega \leq \omega_0 \|K\|_F \|\tilde{x}\| + \omega_1 \|K\|_F \|\tilde{x}\|^2 \quad (37)$$

Logo, com base em (32)-(37), (31) implica

$$\begin{aligned} \dot{V} \leq & -\|\tilde{x}\| \left[ \alpha_1 \|\tilde{x}\| + c_{f \min} \|\tilde{W}_f\|_F^2 + \sum_{i=1}^p c_{gi \min} \|\tilde{W}_{gi}\|_F^2 + \psi_0 \tilde{\psi}^2 + (c_{\rho \min} - 1/2 - \alpha_3^2/2) \tilde{\rho}^2 \right. \\ & \left. - \psi_0 \psi^{*2} - \alpha_0 - \psi_1 \tilde{\psi} + \psi_0 \hat{\psi}^2 \right] + \|\tilde{x}\|^3 - \tilde{x}^T K l \end{aligned} \quad (38)$$

Definindo-se agora

$$\psi^* = \frac{\psi_1 + \sqrt{\psi_1^2 - 4\alpha_0\psi_0}}{2\psi_0} \quad (39)$$

e empregando (18) em (38), segue que

$$\begin{aligned} \dot{V} \leq & -\|\tilde{x}\| \left[ \alpha_1 \|\tilde{x}\| + c_{f \min} \|\tilde{W}_f\|_F^2 + \sum_{i=1}^p c_{gi \min} \|\tilde{W}_{gi}\|_F^2 + \psi_0 \tilde{\psi}^2 + (c_{\rho \min} - 1/2 - \alpha_3^2/2) \tilde{\rho}^2 \right. \\ & \left. - \psi_1 \hat{\psi} + \psi_0 \hat{\psi}^2 \right] \end{aligned} \quad (40)$$

Completando-se o quadrado, (40) implica

$$\begin{aligned} \dot{V} \leq & -\|\tilde{x}\| \left[ \alpha_1 \|\tilde{x}\| + c_{f \min} \|\tilde{W}_f\|_F^2 + \sum_{i=1}^p c_{gi \min} \|\tilde{W}_{gi}\|_F^2 + \psi_0 \tilde{\psi}^2 + (c_{\rho \min} - 1/2 - \alpha_3^2/2) \tilde{\rho}^2 \right. \\ & \left. - \psi_1^2/4\psi_0 \right] \end{aligned} \quad (41)$$

e considere o tempo no qual a identificação é feita,  $t \in \mathfrak{R}^+$ , como dois conjuntos disjuntos  $\Omega_1$  e  $\Omega_2$ , tais que

$$\Omega_1 = \{[t_i, t_{i+1}] \mid \tilde{x}(t) = 0, t \in [t_i, t_{i+1}]\} \quad (42)$$

$$\Omega_2 = \mathfrak{R}^+ \setminus \Omega_1 \quad (43)$$

Então, se  $t \in \Omega_1$ , com base em (41), tem-se que  $\tilde{W}_f$ ,  $\tilde{W}_{gi}$ ,  $\tilde{\psi}$  e  $\tilde{\rho}$  são limitados.

Se por outro lado,  $t \in \Omega_2$ , quaisquer uma das condições:  $\|\tilde{x}\| > \frac{\psi_1^2}{4\alpha_1\psi_0}$ ,  $\|\tilde{W}_f\|_F > \frac{\psi_1}{2\sqrt{\psi_0 c_{fmin}}}$ ,  $\|\tilde{W}_{gi}\|_F > \frac{\psi_1}{2\sqrt{\psi_0 c_{gimin}}}$ ,  $|\tilde{\psi}| > \frac{\psi_1}{2\psi_0}$  ou  $\|\tilde{\rho}\| > \frac{\psi_1}{2\sqrt{\psi_0(c_{\rho min} - 1/2 - \alpha_3^2/2)}}$ , tornam  $\dot{V} < 0$  fora do conjunto compacto  $B_{\gamma_0} = \{\tilde{\chi} \in B_R \mid \|\tilde{\chi}\| \leq \gamma_0\}$ , onde  $B_{\gamma_0} \in B_R$  por (17). Uma vez que  $V$  é positiva definida e  $\dot{V}$  é negativa definida fora da região  $B_{\gamma_0}$ , conclui-se que  $V$  é limitada por um  $\bar{V}$  constante e conseqüentemente  $\tilde{x}$ ,  $\tilde{W}_f$ ,  $\tilde{W}_{gi}$ ,  $\tilde{\psi}$  e  $\tilde{\rho}$  são limitados. Em outras palavras, o fato de que  $B_{\gamma_0}$  é limitada implica que unicamente poderá ocorrer  $\dot{V} \geq 0$  para  $\tilde{x}$ ,  $\tilde{W}_f$ ,  $\tilde{W}_{gi}$ ,  $\tilde{\psi}$  e  $\tilde{\rho}$  finitos e dentro da região  $B_{\gamma_0}$ , pois em  $B_{\gamma_0}^C$  temos  $\dot{V} < 0$  e  $V$  tem que parar de crescer para valores finitos de  $\tilde{x}$ ,  $\tilde{W}_f$ ,  $\tilde{W}_{gi}$ ,  $\tilde{\psi}$  e  $\tilde{\rho}$ , implicando que  $\tilde{x}$ ,  $\tilde{W}_f$ ,  $\tilde{W}_{gi}$ ,  $\tilde{\psi}$  e  $\tilde{\rho}$  são limitados. Objetivando-se provar agora que o erro estendido  $\tilde{\chi}$  é *ultimately bounded*, escreva (26) como  $V = \tilde{\chi}^T T \tilde{\chi}$ , sendo  $T$  uma matriz diagonal, cujos máximo e mínimo autovalores  $T_M$  e  $T_m$  são fornecidos pela hipótese 3, e defina o máximo e mínimo valores de  $V$  sobre os contornos de  $B_{\gamma_0}$  e  $B_R$ , respectivamente, como  $V_M = \max_{\|\tilde{\chi}\|=\gamma_0} V = \gamma_0^2 T_M$  e  $V_m = \min_{\|\tilde{\chi}\|=R} V = R^2 T_m$ . Defina também os conjuntos  $\Omega_M = \{\tilde{\chi} \mid V \leq V_M\}$  e  $\Omega_m = \{\tilde{\chi} \mid V \leq V_m\}$ . Então, uma vez que  $\Omega_M \subset \Omega_m$ , conclui-se que  $\tilde{\chi}$ , e portanto,  $\tilde{x}$ ,  $\tilde{W}_f$ ,  $\tilde{W}_{gi}$ ,  $\tilde{\psi}$  e  $\tilde{\rho}$  são *ultimately bounded*.

Com a finalidade de provar a convergência assintótica do erro de predição, considere os conjuntos disjuntos  $\Omega_x$  e  $\Omega_x^C$ , onde

$$\Omega_x = \{\tilde{x} \mid \|\tilde{x}\| \leq z_0 \exp(-z_1 t)\} \quad (44)$$

Se  $\tilde{x} \in \Omega_x^C$  (ou equivalentemente  $\|\tilde{x}\| > z_0 \exp(-z_1 t)$ ), tem-se

$$\tilde{x}^T K l > \|\tilde{x}\|^3 + \psi_1 \hat{\psi} \|\tilde{x}\| \quad (45)$$

pois  $\frac{2\psi_1 \hat{\psi} \|\tilde{x}\|^2}{\|\tilde{x}\| + z_0 \exp(-z_1 t)} > \psi_1 \hat{\psi} \|\tilde{x}\|$ . Então o erro de predição converge para zero. De fato, usando-se (45), (38) implica

$$\begin{aligned} \dot{V} < -\|\tilde{x}\| \left[ \alpha_1 \|\tilde{x}\| + c_{fmin} \|\tilde{W}_f\|_F^2 + \sum_{i=1}^p c_{gimin} \|\tilde{W}_{gi}\|_F^2 + \psi_0 \tilde{\psi}^2 + (c_{\rho min} - \alpha_3/2 - 1) \tilde{\rho}^2 \right. \\ \left. - \psi_0 \psi^{*2} - \alpha_0 - \psi_1 \tilde{\psi} + \psi_1 \hat{\psi} \right] \end{aligned} \quad (46)$$

que, usando-se (39), resulta

$$\dot{V} < -\alpha_1 \|\tilde{x}\|^2 \quad (47)$$

Adicionalmente note que, de (26) e (47),  $V$  é limitada inferiormente e não crescente com o tempo. Então, com base em (47), obtém-se

$$\lim_{t \rightarrow \infty} \int_0^{\infty} \|\tilde{x}(\tau)\|^2 d\tau = (V(0) - V_{\infty}) / \alpha_1 < \infty \quad (48)$$

onde  $\lim_{t \rightarrow \infty} V(t) = V_\infty < \infty$ . Note ainda que  $\|\tilde{x}\|^2$  é uniformemente contínua, uma vez que  $\tilde{x}$ ,  $\tilde{W}_f$ ,  $\tilde{W}_{gi}$ ,  $\tilde{\psi}$ ,  $\tilde{\rho}$  e  $l$  são limitados e, com base em (13), segue que  $\dot{\tilde{x}}$  é também limitado. Portanto, aplicando o lema de Barbalat [12], conclui-se que

$$\lim_{t \rightarrow \infty} \tilde{x}(t) = 0, \quad \forall \tilde{x} \in \Omega_x^C \quad (49)$$

Se  $\tilde{x} \in \Omega_x$ , obviamente

$$\lim_{t \rightarrow \infty} \tilde{x}(t) = 0, \quad \forall \tilde{x} \in \Omega_x \quad (50)$$

Conseqüentemente, com base em (49) e (50), segue-se que  $\lim_{t \rightarrow \infty} \tilde{x}(t) = 0$ . □

**Comentário 5:** Note que as condições (24) e (25) não podem ser verificadas *a priori*, uma vez que em (24) e (25) são desconhecidos os pesos “ótimos” e limitantes para os erros de aproximação e distúrbios. Contudo, esta deficiência não é peculiar exclusivamente ao esquema proposto, senão à maioria de algoritmos para identificação baseados em RNAs. Por exemplo, os trabalhos em [3]-[7] utilizam a projeção de parâmetros, *switching- $\sigma$*  e zona-morta, que requerem informação prévia sobre o sistema e erro de modelagem. Convém ressaltar que esta deficiência, vide [12] para maiores detalhes, também tem sido reportada no caso de estimação paramétrica *on-line* de sistemas lineares.

**Comentário 6:** No caso em que  $z_1 = 0$  em (19), o algoritmo proposto também garante convergência exponencial do erro de predição para uma bola em torno da origem  $B_{z_0} = \{\tilde{x} \mid \|\tilde{x}\| \leq z_0, z_0 > 0\}$ . Mais exatamente, se  $\|\tilde{x}\| > z_0$ , então (38) implica

$$\begin{aligned} \dot{V} &\leq -\alpha_1 \|\tilde{x}\|^2 - c_{f \min} z_0 \|\tilde{W}_f\|_F^2 - z_0 \sum_{i=1}^p c_{gi \min} \|\tilde{W}_{gi}\|_F^2 - \psi_0 z_0 \tilde{\psi}^2 + z_0 (c_{\rho \min} - 1/2 - \alpha_3^2/2) \tilde{\rho}^2 \\ &\leq -2\eta V \end{aligned} \quad (51)$$

e por conseguinte  $\|\tilde{x}(t)\| \leq \eta_0 \exp(-\eta t)$ , onde  $\eta_0 = \sqrt{V(0)/\lambda_{\min}(P)}$  e

$$\eta = \min[\alpha_1/2\lambda_{\max}(P), c_{f \min} z_0 \alpha_f, c_{gi \min} z_0 \alpha_{gi}, \psi_0 z_0 \alpha_\psi, z_0 (c_{\rho \min} - 1/2 - \alpha_3^2/2) \alpha_\rho, i = 1, \dots, p].$$

**Comentário 7:** Os resultados apresentados nesta seção, baseados em RNAPLs, podem ser modificados para usar RNAs com uma camada escondida (SLHNNs): usando-se a expansão de Taylor com respeito aos pesos para representar as componentes de alta ordem da SLHNN [20], e empregando a metodologia desenvolvida.

## 6 Simulações

Nesta seção são apresentadas simulações com a finalidade de ilustrar os resultados teóricos e mostrar a aplicação do algoritmo proposto. Para tanto, considere o processo descrito por

$$\dot{x} = \begin{bmatrix} x_2 + u \\ -\sqrt{x_1} + 0.5x_2^2 \end{bmatrix} + \begin{bmatrix} 0 \\ \sin(x_1^3 + x_2^2) \end{bmatrix} [2 + \sin(t) + \cos(5t) + \exp(-t)] + \begin{bmatrix} 0 \\ \sin(4t) \end{bmatrix} \quad (52)$$

onde  $x = [x_1, x_2]^T$  e  $u$  é um sinal quadrado de amplitude 1 e frequência 0.25.

Note que o parâmetro em (52) não satisfaz a hipótese 2, entretanto, o algoritmo proposto pode ser naturalmente aplicado neste processo, uma vez que (52) pode ser escrita como

$$\dot{x} = \begin{bmatrix} x_2 + u \\ -\sqrt{x_1} + 0.5x_2^2 \end{bmatrix} + \begin{bmatrix} 0 \\ \sin(x_1^3 + x_2^2) \end{bmatrix} [2 + \exp(-t)] + \begin{bmatrix} 0 \\ \sin(4t) + \sin(x_1^3 + x_2^2) [\sin(t) + \cos(5t)] \end{bmatrix} \quad (53)$$

onde agora  $\rho_1(t) = 2 + \exp(-t)$  satisfaz a hipótese 2.

O modelo para identificação (12) e as leis de adaptação (18)-(23) propostas foram implementadas com

$$\begin{aligned} \hat{x}(0) = [0,0]^T, \quad \hat{W}_f(0) = 0, \quad \hat{W}_{g1}(0) = 0, \quad \hat{\psi}(0) = 1, \quad \hat{\rho}(0) = -1, \quad \xi = \hat{x}, \quad z_0 = z_1 = 0.5, \quad L = 2I, \quad K = I, \\ W_{f0} = 0, \quad W_{g10} = 0, \quad \pi_f = \pi_{g1} = [s(\hat{x}_1), s(\hat{x}_2), s(u)]^T, \quad s(\cdot) = 5/[1 + \exp(-0.5(\cdot))], \quad (54) \\ \alpha_f = \alpha_{g1} = \alpha_\psi = \alpha_\rho = 1, \quad C_f = C_{g1} = 0.1I, \quad \psi_0 = 0.1, \quad \psi_1 = 5 \text{ e } c_{\rho1} = 10. \end{aligned}$$

O desempenho na estimação de (52) é apresentado nas Figuras 1 e 2. Note que as simulações confirmam os resultados teóricos, isto é, o algoritmo proposto é estável e o erro de predição converge assintoticamente para zero. Adicionalmente, de (54), não é necessário usar ganhos elevados ou estruturas complexas, como usual na literatura [3]-[11], para assegurar erros de predição residuais arbitrariamente pequenos.

## 7 Conclusões

Neste artigo foi proposta, no contexto da identificação *on-line* de sistemas não-lineares incertos, uma metodologia para projeto de leis de adaptação para os pesos de redes neurais parametrizáveis linearmente. Para assegurar estabilidade e convergência do algoritmo proposto foi incorporado, similarmente a [13], um termo forçante ou de realimentação no modelo para identificação, o que, conjuntamente com argumentos usuais de Lyapunov e uma técnica de limitação adaptativa, permitiu provar a estabilidade dos sinais de erro e convergência do erro residual de predição para zero, mesmo na presença de erros de aproximação e distúrbios. Com a finalidade de reduzir o conservadorismo de trabalhos prévios, como por exemplo [3]-[11] e [13]-[14], foi assumido para o projeto do algoritmo que os parâmetros do sistema incerto podem ser variantes no tempo. Além disso, foram minoradas as hipóteses de aplicação restritivas em [13]-[14]. Um exemplo de simulação que considerou a identificação neural de um sistema não-linear incerto com parâmetros variantes no tempo e sujeito a distúrbios foi considerado para mostrar a aplicação e desempenho do esquema proposto.

## Agradecimentos

Os primeiro autor agradece à FAPESP, processos 02/13829-9 e 03/11199-0, e o segundo autor agradece ao CNPq/Pronex, processo 662015/1998-3, pelo suporte financeiro.

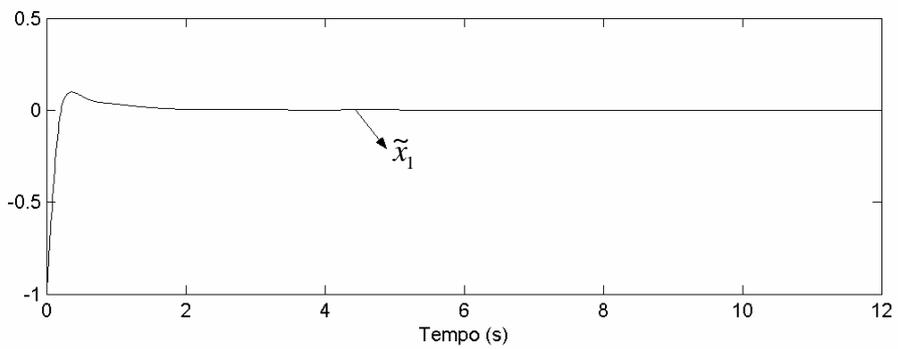
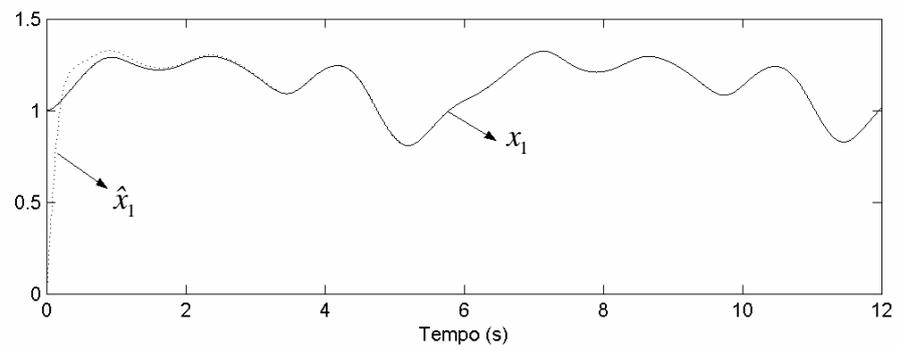


Figura 1- Desempenho na predição de  $x_1$ .

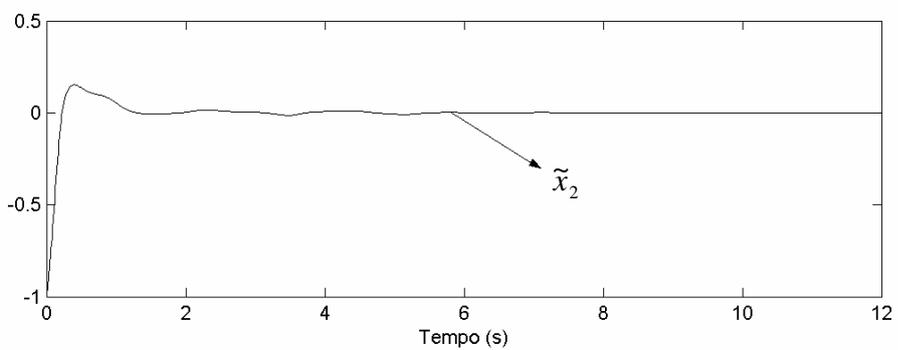
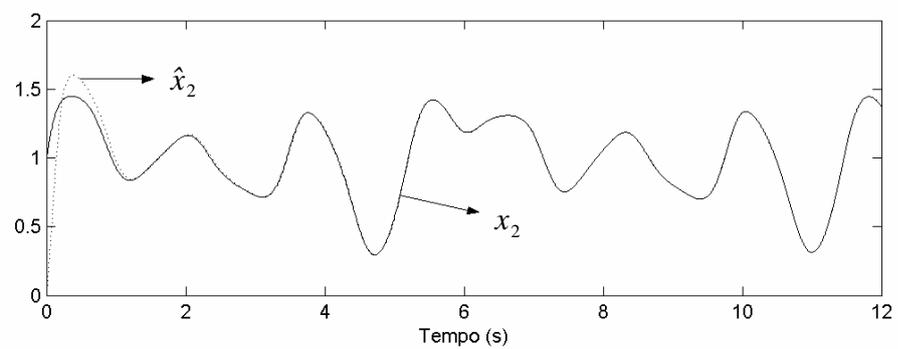


Figura 2- Desempenho na predição de  $x_2$ .

## Referências bibliográficas

- [1] R. Johansson, *System modeling and identification*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1993.
- [2] R.K. Mehra, "Nonlinear system identification: selected survey and recent trends, *5th IFAC Symp. Identification and Systems Parameter Estimation*, p.77-83, 1979.
- [3] M. M. Polycarpou and P. A. Ioannou, "Identification and control of nonlinear systems using neural network models: design and stability analysis," Tech. Rep. 91-09-01, Dept. Elect. Eng. Syst., Univ. Southern Calif., Los Angeles, 1991.
- [4] E. B. Kosmatopoulos, M. M. Polycarpou, M. A. Christodoulou, and P. A. Ioannou, "High-order neural network structures for identification of dynamical systems," *IEEE Trans. Neural Networks*, vol. 6, pp. 422-431, Mar. 1995.
- [5] W. Yu, A. S. Poznyak, and X. Li, "Multilayer dynamical neural networks for non-linear system on-line identification," *International Journal of Control*, vol. 74, no.18, pp. 1858-1864, 2001.
- [6] Q. Song and M. J. Grimble, "Robust training algorithms of multilayered neural networks for identification of nonlinear dynamical systems," *Proceedings of the American Control Conference*, pp. 2419-2423, 1997.
- [7] S. Jagannathan and F. L. Lewis, "Identification of nonlinear dynamical systems using multilayered neural networks," *Automatica*, vol. 32, no. 12, pp. 1707-1712, 1996.
- [8] J. A. R. Vargas, *Identificação de sistemas dinâmicos via redes neurais artificiais*, dissertação de mestrado, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, Brasil, 1997.
- [9] W. Yu and X. Li, "Some new results on system identification with dynamic neural networks," *IEEE Trans. Neural Networks*, vol. 12, pp. 412-417, Feb. 2001.
- [10] W. Yu, "Passivity analysis for dynamic multilayer neuro identifier," *IEEE Trans. Circuits and Systems I: Fundamental Theory and Applications*, vol. 50, Iss.1, pp. 173-178, Jan. 2003.
- [11] W. Yu, "nonlinear system identification using discrete-time recurrent neural networks with stable learning algorithm," *Information sciences*, vol. 158, pp. 131-137, 2004.
- [12] P.A. Ioannou and J. Sun, *Robust adaptive control*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1996.
- [13] J.A.R. Vargas, E.M. Hemerly and V. I. Gervini, "Convergência assintótica do erro de estimação em algoritmos usando aproximadores parametrizáveis linearmente," *Anais do XV Congresso Brasileiro de Automática*, sessão 0502, Gramado/RS, Brasil, setembro 2004.
- [14] J.A.R. Vargas and E.M. Hemerly, "Algoritmo para identificação assintótica via APLS e sem conhecimento prévio de limitantes," aceito para publicação na edição especial da Revista *Controle & Automação – Automação Inteligente da Sociedade Brasileira de Automática*, 2006.
- [15] M. M. Polycarpou, "Stable adaptive neural control scheme for nonlinear systems," *IEEE Trans. Automatic Control*, vol. 41, no. 3, pp. 447-451, Mar. 1996.
- [16] S. S. Ge, C. C. Hang, T. H. Lee, and T. Zhang, *Stable adaptive neural network control*, Kluwer academic publishers, 2002.
- [17] R. M. Sanner and J. J. E. Slotine, "Stable recursive identification using radial basis function networks," *Proc. 1992 ACC*, Chicago, IL, pp. 1829-1833, June 1992.
- [18] Q. Zhang and A. Benveniste, "Wavelet networks," *IEEE Trans. Neural Networks*, vol. 3, pp. 889-898, 1992.
- [19] L. X. Wang, *Adaptive fuzzy system and control: design and stability analysis*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1994.
- [20] F. L. Lewis, A. Yesildirek, and K. Liu, "Multilayer neural-net robot controller with guaranteed tracking performance," *IEEE Trans. Neural Networks*, vol. 7, pp. 388-399, Jan. 1996.