Proceedings of the European Control Conference 2007
Kos, Greece, July 2-5, 2007

A contribution to the use of Hopfield neural networks for parameter estimation

Hugo Alonso, Teresa Mendonça, Paula Rocha

Abstract— **This paper presents a contribution to the use of Hopfield neural networks (HNNs) for parameter estimation. Our focus is on time-invariant systems that are linear in the parameters. We introduce a suitable HNN and present a weaker condition than the currently existing ones that guarantees the convergence of the parameterization estimated by the network to the actual parameterization. The application of our results is illustrated in a parameter estimation problem for a two carts system.**

Keywords Parameter estimation, Hopfield neural networks, Lyapunov stability theory.

I. Introduction

System identification plays undoubtedly an important role in different areas such as biomedicine, robotics and fluid dynamics applied to aircraft and motor vehicle industries. In fact, obtaining an accurate model for a process is important not only for the study of the process itself, but very often for the design of a suitable control strategy. In this context, consider the example of a patient undergoing general anesthesia, where a control action is used for drug delivery optimization [7]: the safety of the patient obviously depends on the reliability of the control strategy, but the latter is based on the identification of the patient dynamics.

Many approaches to system identification have been proposed, in particular the use of neural networks as black-box models (see, for instance, [8]). However, there are problems where gray-box models are preferred, since they can be applied not only for prediction but also for description purposes, offering in this case some insight into the underlying system dynamics. This motivated us to propose instead the use of neural networks as a tool designed to estimate the parameters of a gray-box model intended to fit the system data. In this paper, we consider Abe's formulation of a Hopfield neural network (HNN) [1]. Compared to the original Hopfield's formulation [4], this formulation is simpler, computationally less demanding, and hence more suitable for parameter estimation. These advantages have been pointed out in [2], where the use of such formulation in that context has been proposed. The authors proved that the parameterization estimated by their network asymptotically converges to the actual parameterization for a system assumed to be time-invariant and linear in the parameters. However, this result was obtained under very restrictive assumptions on the matrix (W_{ij}). In fact, these assumptions are so restrictive that they do not hold for the case study presented in that work, in spite of the good performance of the corresponding estimation process. This suggested us that the conditions of [2] could be relaxed, motivating this contribution.

The paper is organized as follows. In Section II, the problem of applying a HNN to parameter estimation is formulated and our contributions clarified. Section III presents both the way how we define the network estimator and its stability analysis. In Section IV, the application of our results is illustrated in a parameter estimation problem for a two carts system. Finally, we present the conclusions and future work.

II. Problem formulation

Our focus is on time-invariant systems that are linear in the parameters, *i.e.*, systems that can be represented in the form

$$\mathbf{y}(t) = \mathbf{A}(t)\theta, \tag{1}$$

for some $\mathbf{y} : [t_0, +\infty[\to \mathbb{R}^{m\times 1}$, $\mathbf{A} : [t_0, +\infty[\to \mathbb{R}^{m\times n}$, being $\theta \in \mathbb{R}^{n\times 1}$ the vector of the unknown parameters to estimate. It is assumed that $\mathbf{y}, \mathbf{A}$ are continuously differentiable and bounded functions, and that $\mathbf{y}(t), \mathbf{A}(t)$ are available at each time t, although $\mathbf{y}, \mathbf{A}$ are possibly explicitly unknown. Furthermore, it is reasonable to assume the knowledge about some $c > 0$ for which $\theta \in]-c, c[^n$. Our problem is that of defining a Hopfield neural network (HNN) that is able to generate a trajectory $\hat{\theta}(\cdot)$ such that $\forall t \geq t_0 \ \hat{\theta}(t) \in]-c, c[^n$ and $\lim_{t\to+\infty}\hat{\theta}(t) = \theta$ under mild assumptions. In [2], the proposed network generates a trajectory confined to $\hat{\theta}(t_0)+]-1, 1[^n$ and asymptotically convergent to θ under the assumption that $\forall t \geq t_0 \ \ker(\mathbf{A}(t)) = \{\mathbf{0}\}$. But if the latter condition holds, then the solution to the estimation problem given by $\theta = \left(\mathbf{A}^T(t)\mathbf{A}(t)\right)^{-1}\mathbf{A}^T(t)\mathbf{y}(t)$ can be obtained at any time $t \geq t_0$, even if $\mathbf{A}^T(t)\mathbf{A}(t)$ is ill-conditioned, in which case simpler standard methods can be applied to compute θ. In addition, a necessary condition for $\forall t \geq t_0 \ \ker(\mathbf{A}(t)) = \{\mathbf{0}\}$ to hold is that $m \geq n$, *i.e.*, the system should not be overparameterized. However, a successful application of a HNN to an estimation problem where $m < n$ is given in [2], which motivated us to find more general conditions. Here, we start by introducing a different HNN that has the advantage of requiring no prior knowledge on the choice of $\hat{\theta}(t_0)$, the initial estimate of θ, and which

Hugo Alonso (corresponding author) and Teresa Mendonça are with the Departamento de Matemática Aplicada, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal (e-mail: hugo.alonso@fc.up.pt,tmendo@fc.up.pt).

Paula Rocha is with the Departamento de Matemática, Universidade de Aveiro, Campo de Santiago, 3810-193 Aveiro, Portugal (e-mail: procha@mat.ua.pt).

ISBN: 978-960-89028-5-5

accommodates the value of c, a parameter that determines the size of the search space. Then, we show that θ is a globally uniformly asymptotically stable equilibrium point of this HNN at $t = t_0$[1] if the following holds: for all nondegenerate interval $I \subset [t_0, +\infty[$, $\bigcap_{t \in I} \ker(\mathbf{A}(t)) = \{\mathbf{0}\}$. It is clear that this is a weaker condition than the one presented in [2]; moreover, it does not imply an order relation between m and n. Finally, note that we assume no upper bound for the parameter c, which can be made arbitrarily large.

III. Hopfield neural networks for parameter estimation

In what follows, we first describe our adaptation of Abe's formulation of a time-invariant Hopfield neural network (HNN) to the parameter estimation problem. Then, we present the stability analysis of the proposed estimator.

A. Hopfield neural network estimator

Consider a HNN where the number of neurons equals the number of parameters to estimate, n, and the neuron dynamics is governed by the ordinary differential equation

$$\frac{dp_i}{dt}(t) = -\left(\sum_{j=1}^{n} W_{ij} f_j(p_j(t)) + I_i\right), \tag{2}$$

being p_i the total input to neuron i, f_j a continuous, nonlinear, bounded and strictly increasing function determining the output of neuron j, W_{ij} a parameter representing the weight associated with the connection from neuron j to neuron i, and I_i a parameter corresponding to the external input or bias of neuron i. The output of neuron j is regarded as the estimate of parameter j, *i.e.*,

$$\hat{\theta}_j = f_j \circ p_j.$$

In order to guarantee that the network trajectory is in the feasible region of the estimation problem, *i.e.*, that $\forall t \geq t_0 \ \hat{\theta}(t) \in]-c, c[^n$, we include c in the the usual definition of f_j to get

$$f_j(p_j(t)) = c \tanh\left(\frac{p_j(t)}{\beta}\right),$$

where $\beta > 0$ is a scaling parameter. The only thing left to define in (2) are the parameters W_{ij} and I_i. The goal is to make the vector of actual parameters θ an equilibrium point of the network at $t = t_0$; hence, it is natural to make the network dynamics depend on the system data, *i.e.*, to take W_{ij} and I_i as functions of $\mathbf{y}(t)$, $\mathbf{A}(t)$. First, let us look at the state space representation of the HNN,

$$\frac{d\hat{\theta}}{dt}(t) = -\frac{1}{c\beta}\mathbf{D}_c(\hat{\theta}(t))\left(\mathbf{W}(t)\hat{\theta}(t) + \mathbf{I}(t)\right), \tag{3}$$

where

$$\mathbf{D}_c(\hat{\theta}(t)) = \operatorname{diag}\left(\left(c^2 - \hat{\theta}_i^2(t)\right)_i\right),$$

$\mathbf{W}(t) = (W_{ij})(t) \in \mathbb{R}^{n \times n}$ and $\mathbf{I}(t) = (I_i)(t) \in \mathbb{R}^{n \times 1}$. Clearly, θ is an equilibrium point of the network at $t = t_0$ if and only if $\forall t \geq t_0 \ \mathbf{W}(t)\theta + \mathbf{I}(t) = \mathbf{0}$ (just note that $\mathbf{D}_c(\hat{\theta}(t))$ is invertible since $\forall t \geq t_0 \ \hat{\theta}(t) \in]-c, c[^n$). In order to achieve this condition, and given that $\forall t \geq t_0 \ \mathbf{A}^T(t)\mathbf{A}(t)\theta - \mathbf{A}^T(t)\mathbf{y}(t) = \mathbf{0}$ by (1), we take

$$\mathbf{W}(t) = \mathbf{A}^T(t)\mathbf{A}(t), \tag{4}$$

$$\mathbf{I}(t) = -\mathbf{A}^T(t)\mathbf{y}(t). \tag{5}$$

Henceforth, we shall refer to (3) together with (4), (5) as our HNN.

B. Stability analysis of the Hopfield neural network estimator

The goal of the stability analysis carried out in this section is to present a weaker sufficient condition, under which θ is a globally uniformly asymptotically stable equilibrium point of our HNN. Often in the literature only global attractiveness is proved, rather than global asymptotic stability. However, an equilibrium point can be attractive without being stable [9]. Stability means that the trajectory of the network remains close to θ if the initial estimate $\hat{\theta}(t_0)$ is sufficiently close, being therefore an important feature in practice, and thus worth to show.

Now, the approach to the stability analysis is made under the framework of Lyapunov stability theory. In the HNN literature, the function

$$E(t, \tilde{\theta}) = \frac{1}{2}\tilde{\theta}^T \mathbf{W}(t)\tilde{\theta} + \tilde{\theta}^T \mathbf{I}(t)$$

is usually taken as being Lyapunov in the time-invariant case, where $\mathbf{W}$ and $\mathbf{I}$ do not vary in time; however, this is not our case. Here, we cannot guarantee that E is Lyapunov, since $\frac{dE}{dt}$ evaluated along the network trajectory,

$$\begin{aligned}\frac{dE}{dt}(t, \hat{\theta}(t)) = &-\frac{1}{2}\hat{\theta}^T(t)\frac{d\mathbf{W}}{dt}(t)(2\theta - \hat{\theta}(t)) \\ &- \frac{1}{c\beta}(\mathbf{W}(t)(\theta - \hat{\theta}(t)))^T \mathbf{D}_c(\hat{\theta}(t))\mathbf{W}(t)(\theta - \hat{\theta}(t)),\end{aligned}$$

is not necessarily negative definite despite the negative definiteness of its second term. Hence, we are led to introduce a suitable Lyapunov function. Before doing so, note that instead of studying the stability behavior of θ as an equilibrium point of our HNN, we can study the stability behavior of the origin as an equilibrium point of a system obtained from the network through an appropriate change of variables, like the one determined from

$$\Delta(t) = \theta - \hat{\theta}(t).$$

The latter condition defines the estimation error, whose dynamics is

[1] Let $f : [t_0, +\infty[\times D \to \mathbb{R}^n$ be piecewise continuous in t and locally Lipschitz in $\mathbf{x}$ on $[t_0, +\infty[\times D$, where $D \subset \mathbb{R}^n$. The point $\mathbf{x}^\star \in D$ is an equilibrium point of the system $\frac{d\mathbf{x}}{dt} = f(t, \mathbf{x})$ at $t = t^\star \geq t_0$ if $\forall t \geq t^\star \ f(t, \mathbf{x}^\star) = \mathbf{0}$ [5].

$$\frac{d\Delta}{dt}(t) = f(t, \Delta(t)), \tag{6}$$

where

$$f(t, \Delta(t)) = -\frac{1}{c\beta}\mathbf{D}_c(\theta - \Delta(t))\mathbf{W}(t)\Delta(t)$$

was obtained minding that $\forall t \geq t_0 \;\; \mathbf{W}(t)\theta + \mathbf{I}(t) = \mathbf{0}$. It is clear that $\Delta^\star = \mathbf{0}$ is the equilibrium point of the estimation error dynamics (6) at $t = t_0$ which corresponds to the equilibrium point $\hat{\theta}^\star = \theta$ of our HNN at $t = t_0$. We are now ready to introduce a suitable Lyapunov function through the next lemma.

Lemma 1: The function $V : \theta +]-c, c[^n \to \mathbb{R}$ defined by

$$V(\Delta(t)) = -\frac{1}{2c}\sum_{i=1}^{n} \ln\left(\left(1 + \frac{\Delta(t)_i}{c - \theta_i}\right)^{c-\theta_i}\left(1 - \frac{\Delta(t)_i}{c + \theta_i}\right)^{c+\theta_i}\right) \tag{7}$$

is a Lyapunov function for the estimation error dynamics (6).

The next lemma states that in every concentric, closed sub-hypercube of $\theta +]-c, c[^n$ containing the origin, the Lyapunov function V is lower and upper bounded by suitable comparison functions. Its proof is based on that of Lemma 4.3 in [5].

Lemma 2: Consider the set

$$S_r = \{\Delta \in \theta +]-c, c[^n : \|\Delta - \theta\|_\infty \leq r\}$$

for some $r \in]\|\theta\|_\infty, c[$[2]. Then, there exist class $\mathcal{K}$ functions[3] γ_1 and γ_2 such that

$$\gamma_1(\|\Delta\|_\infty) \leq V(\Delta) \leq \gamma_2(\|\Delta\|_\infty)$$

for all $\Delta \in S_r$.

An important result to the proof our main contribution is given in the next lemma, which guarantees that the trajectory $\hat{\theta}(\cdot)$ generated by our HNN is unique for each initial estimate $\hat{\theta}(t_0)$ of the actual parameterization θ. This follows from Theorem 3.3 in [5], since f is Lipschitz and every trajectory lies entirely in a compact subset of $\theta +]-c, c[^n$.

Lemma 3: The initial-value problem defined by the estimation error dynamics (6) and the initial condition $\Delta(t_0) = \Delta_0$,

$$\frac{d\Delta}{dt}(t) = f(t, \Delta(t)), \quad \Delta(t_0) = \Delta_0, \tag{8}$$

has a unique solution over $[t_0, +\infty[$.

Corollary 1: If $\Delta(t_0) \neq \mathbf{0}$, then the solution $\Delta(t)$ to the initial-value problem (8) is such that $\forall t \geq t_0 \;\; \Delta(t) \neq \mathbf{0}$.

We are now ready to state and prove the main result of this paper.

[2]By definition, $\|\mathbf{v}\|_\infty = \max\{|v_i|\}$.

[3]A continuous function $\gamma : [0, a] \to [0, +\infty[$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and $\gamma(0) = 0$ [5].

Theorem 1: The equilibrium point $\Delta^\star = \mathbf{0}$ of the estimation error dynamics (6) is globally uniformly asymptotically stable if

$$\text{for all nondegenerate interval } I \subset [t_0, +\infty[, \quad \bigcap_{t \in I} \ker(\mathbf{A}(t)) = \{\mathbf{0}\}. \tag{9}$$

Proof: Consider the initial-value problem (8), which has a unique solution over $[t_0, +\infty[$ by Lemma 3. Let us assume that $\Delta_0 \in \Omega_\rho$, where

$$\Omega_\rho = \{\Delta \in S_r : V(\Delta) \leq \rho\}$$

for some $r \in]\|\theta\|_\infty, c[$, $\rho \in]0, \min_{\Delta \in \partial S_r}\{V(\Delta)\}$; moreover, suppose that $\Delta_0 \neq \mathbf{0}$, otherwise $\forall t \geq t_0 \;\; \Delta(t) = \mathbf{0}$, since $\Delta^\star = \mathbf{0}$ is an equilibrium point at $t = t_0$. It is possible to prove that, under (9), we have

$$\exists \delta > 0 : \forall t \geq t_0 \;\; V(\Delta(t + \delta)) < V(\Delta(t)). \tag{10}$$

To this end, it suffices to consider the negative semidefiniteness of $\frac{dV}{dt}$, the symmetry and positive semidefiniteness of $\mathbf{W}$, the fact that $\forall t \geq t_0 \;\; \ker(\mathbf{W}(t)) = \ker(\mathbf{A}(t))$, and finally use Corollary 1. Now, it is easy to see that, from (10), we have

$$\exists \delta > 0 : \forall t \geq t_0 \;\; V(\Delta(t + \delta)) \leq (1 - \lambda)V(\Delta(t))$$

for some $\lambda \in]0, 1[$. Similarly to the proof of Theorem 8.5 in [5], it can be shown that there exists a class $\mathcal{KL}$ function[4] σ such that

$$\forall t \geq t_0 \;\; V(\Delta(t)) \leq \sigma(V(\Delta_0), t - t_0). \tag{11}$$

Hence,

$$\begin{aligned} \forall t \geq t_0 \;\; \|\Delta(t)\|_\infty &\leq \gamma_1^{-1}(V(\Delta(t))) \\ &\leq \gamma_1^{-1}(\sigma(V(\Delta_0), t - t_0)) \\ &\leq \gamma_1^{-1}(\sigma(\gamma_2(\|\Delta_0\|_\infty), t - t_0)) \\ &\triangleq \mu(\|\Delta_0\|_\infty, t - t_0) \end{aligned} \tag{12}$$

with γ_1, γ_2 as in Lemma 2, and where the function μ is a class $\mathcal{KL}$ function by Lemma 4.2 in [5]. Making use of the fact that the norms $\|\cdot\|_2$ and $\|\cdot\|_\infty$ are equivalent in $\mathbb{R}^n$, from (12) it follows that $\Delta^\star = \mathbf{0}$ is uniformly asymptotically stable by Lemma 4.5 in [5]. Finally, note that $\lim_{r \to c^-} \partial S_r = \partial(\theta +]-c, c[^n)$, and therefore $\lim_{r \to c^-} \min_{\Delta \in \partial S_r}\{V(\Delta)\} = +\infty$ by (7). In this way, $\rho < \min_{\Delta \in \partial S_r}\{V(\Delta)\}$ can be made arbitrarily large so that the set Ω_ρ includes any initial state $\Delta_0 \in \theta +]-c, c[^n$. Thus, $\Delta^\star = \mathbf{0}$ is globally uniformly asymptotically stable. ■

[4]A continuous function $\sigma : [0, a] \times [0, +\infty[\to [0, +\infty[$ is said to belong to class $\mathcal{KL}$ if, for each fixed τ, the mapping $\sigma(\xi, \tau)$ belongs to class $\mathcal{K}$ with respect to ξ and, for each fixed ξ, the mapping $\sigma(\xi, \tau)$ is decreasing with respect to τ and $\sigma(\xi, \tau) \to 0$ as $\tau \to +\infty$ [5].

IV. CASE STUDY

In the following, we illustrate the application of our theoretical results to the parameter estimation problem for a two carts system. This system of some unknown parameters is often considered in a benchmark problem of robust adaptive feedback control (see, for instance, [3]).

Consider two carts joined by a spring and a damper as represented in Fig. 1. The application of physical laws yields the state space model

$$\dot{\mathbf{x}}(t) = \mathbf{M}\mathbf{x}(t) + \mathbf{N}u(t), \tag{13}$$

where

$$\mathbf{x}(t) = (x_1(t), x_2(t), \dot{x}_1(t), \dot{x}_2(t))^T,$$

$$\mathbf{M} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k}{m_1} & \frac{k}{m_1} & -\frac{b}{m_1} & \frac{b}{m_1} \\ \frac{k}{m_2} & -\frac{k}{m_2} & \frac{b}{m_2} & -\frac{b}{m_2} \end{pmatrix},$$

$$\mathbf{N} = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{m_1} \\ 0 \end{pmatrix},$$

and where $x_i, \dot{x}_i, \ddot{x}_i$ denote respectively the displacement, velocity and acceleration of cart i, u the force applied to cart 1, and m_i, k, b the parameters corresponding respectively to the mass of cart i, the spring constant, and the damper constant. Assume that the masses m_i are known, and that the constants k, b are the unknown parameters to estimate. Hence, (13) can be cast into the reduced form (1) by taking

$$\mathbf{y}(t) = \begin{pmatrix} m_1\ddot{x}_1(t) - u(t) \\ m_2\ddot{x}_2(t) \end{pmatrix},$$

$$\mathbf{A}(t) = \begin{pmatrix} -x_1(t) + x_2(t) & -\dot{x}_1(t) + \dot{x}_2(t) \\ x_1(t) - x_2(t) & \dot{x}_1(t) - \dot{x}_2(t) \end{pmatrix},$$

$$\theta = \begin{pmatrix} k \\ b \end{pmatrix}.$$

A successful parameter estimation is guaranteed if condition (9) on ker($\mathbf{A}$) holds, *i.e.*, for all nondegenerate interval $I \subset [0, +\infty[$, $\bigcap_{t\in I} \ker(\mathbf{A}(t)) = \{\mathbf{0}\}$. Start by noting that $\forall t \geq 0$ $\ker(\mathbf{A}(t)) \supsetneqq \{\mathbf{0}\}$, given that $\forall t \geq 0$ $\det(\mathbf{A}(t)) = 0$; in fact, $\forall t \geq 0$ $\ker(\mathbf{A}(t)) = \{\tilde{\theta} \in \mathbb{R}^2 : (x_1(t) - x_2(t))\tilde{\theta}_1 + (\dot{x}_1(t) - \dot{x}_2(t))\tilde{\theta}_2 = 0\}$. Therefore, it is easy to see that condition (9) holds if for all nondegenerate interval $I \subset [0, +\infty[$ and all $r_1, r_2 \in \mathbb{R}$, there exists $t \in I$ such that

$$x_1(t) - x_2(t) \neq r_1 \mathrm{e}^{r_2 t}. \tag{14}$$

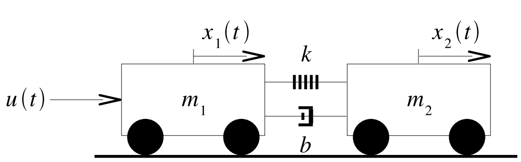

Fig. 1. Two carts system.

In the following, we shall study $x_1 - x_2$, and show that the latter condition holds for suitable initial values on the system state variables x_i, $\dot{x}_i$ and suitable forces u.

Let us start by seeing that any solution x_1, x_2 of (13) is also a solution of

$$\ddot{x}_1(t) - \ddot{x}_2(t) + \zeta_1(\dot{x}_1(t) - \dot{x}_2(t)) + \zeta_0(x_1(t) - x_2(t)) = \frac{1}{m_1}u(t),$$

where $\zeta_1 = b\frac{m_1+m_2}{m_1 m_2}$ and $\zeta_0 = k\frac{m_1+m_2}{m_1 m_2}$. The application of the (unilateral) Laplace transform $\mathcal{L}$ to both sides of the preceding equality yields after some manipulation

$$X_1(s) - X_2(s) = F(s)\,(U(s) + m_1((x_1(0) - x_2(0))(s + \zeta_1) + \dot{x}_1(0) - \dot{x}_2(0))),$$

where $X_i = \mathcal{L}[x_i]$, $U = \mathcal{L}[u]$, and

$$F(s) = \frac{1}{m_1(s^2 + \zeta_1 s + \zeta_0)}.$$

Given the form of the transfer function F, it can be seen that the condition concerning (14) holds for general initial values on the system state variables x_i, $\dot{x}_i$ and general forces u, like those represented by Bohl functions[5]. In this way, condition (9) holds, and our Hopfield neural network (HNN) is able to carry out a successful parameter estimation as long as $\mathbf{y}$,$\mathbf{A}$ are bounded and the value chosen for c is such that $\theta \in]-c, c[^2$. Note that, since $\zeta_i > 0$, all poles of the transfer function F have a negative real part; hence, the system is BIBO-stable, and $\mathbf{y}$,$\mathbf{A}$ are bounded for all bounded forces u. Therefore, a bounded u and a large value for c should be considered so that a successful parameter estimation can be carried out. Finally, mind that the value of the other time-invariant parameter of the HNN, β, does not affect the qualitative properties of the network. In what follows, we illustrate the performance of the HNN for a particular parameterization of the system.

Let us assume, for instance, that $m_1 = m_2 = 2\,[\mathrm{kg}]$, $k = 1\left[\frac{\mathrm{N}}{\mathrm{m}}\right]$ and $b = 0.1\left[\frac{\mathrm{Ns}}{\mathrm{m}}\right]$. Consider the HNN as defined in (3)-(5) with $c = 5$ and $\beta = 0.01$. Fig. 2 depicts the time-evolution of the estimated parameterization produced by this HNN for two different initial conditions $x_i(0)$, $\dot{x}_i(0)$ (columns) and two different forces u (rows): the first column refers to $x_i(0) = 0\,[\mathrm{m}]$, $\dot{x}_i(0) = 0\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$, the second to $x_i(0) = 0\,[\mathrm{m}], \dot{x}_1(0) = 1, \dot{x}_2(0) = 2\left[\frac{\mathrm{m}}{\mathrm{s}}\right]$, and the first row refers to $u(t) = 1+\mathrm{e}^{-t}\,[\mathrm{N}]$, the second to $u(t) = \cos(\pi t)\,[\mathrm{N}]$. In each of the four cases, the initial estimates of k and b were randomly generated according to the continuous uniform distribution $U(]0, c[)$. As expected, the estimation process is well succeeded in all cases. Finally, in order to assess the HNN performance when the system data is corrupted by noise, we repeated the simulations assuming that the measurement of x_1, the displacement of cart 1, is affected by Gaussian

[5]A Bohl function is a function whose Laplace transform is rational and strictly proper.

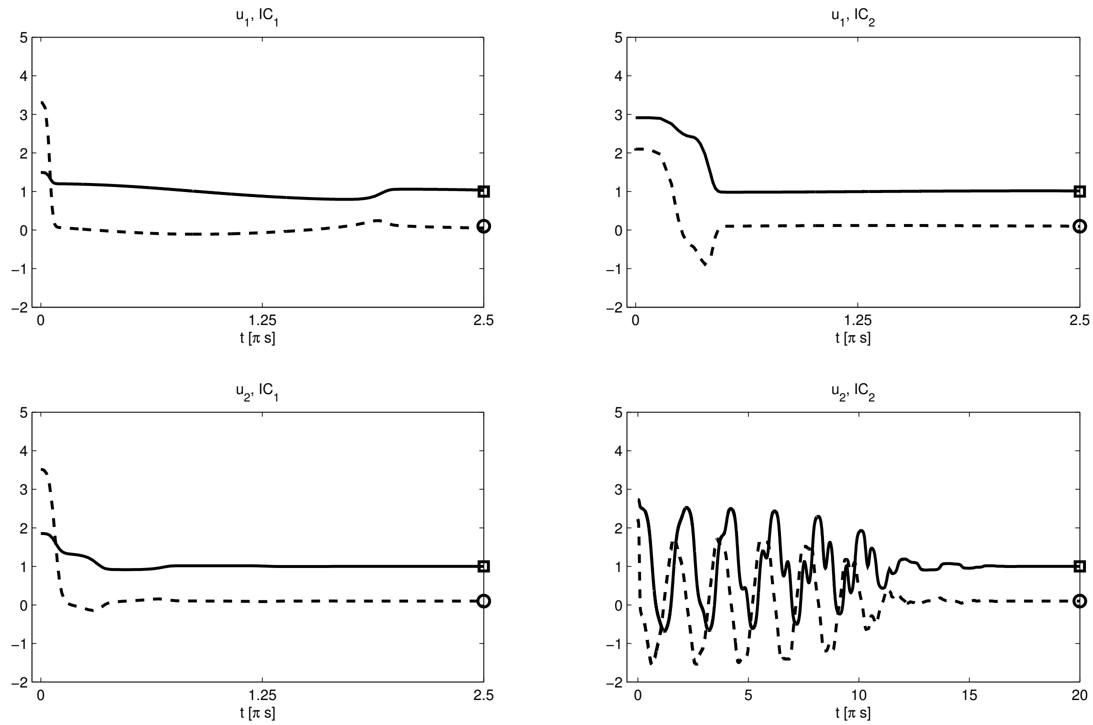

Fig. 2. Time-evolution of the estimated parameterization produced by the HNN for two different sets of initial conditions IC and two different forces u, where $\mathrm{IC}_1 = \{x_i(0) = 0\,[\mathrm{m}], \dot{x}_i(0) = 0\,[\frac{\mathrm{m}}{\mathrm{s}}]\}$, $\mathrm{IC}_2 = \{x_i(0) = 0\,[\mathrm{m}], \dot{x}_1(0) = 1, \dot{x}_2(0) = 2\,[\frac{\mathrm{m}}{\mathrm{s}}]\}$, and $u_1(t) = 1 + \mathrm{e}^{-t}$ [N], $u_2(t) = \cos(\pi t)$ [N]. The solid and dashed lines represent respectively the estimated values for k and b; the square and the circle represent respectively the actual values of k and b. Mind the different time interval used in the last simulation.

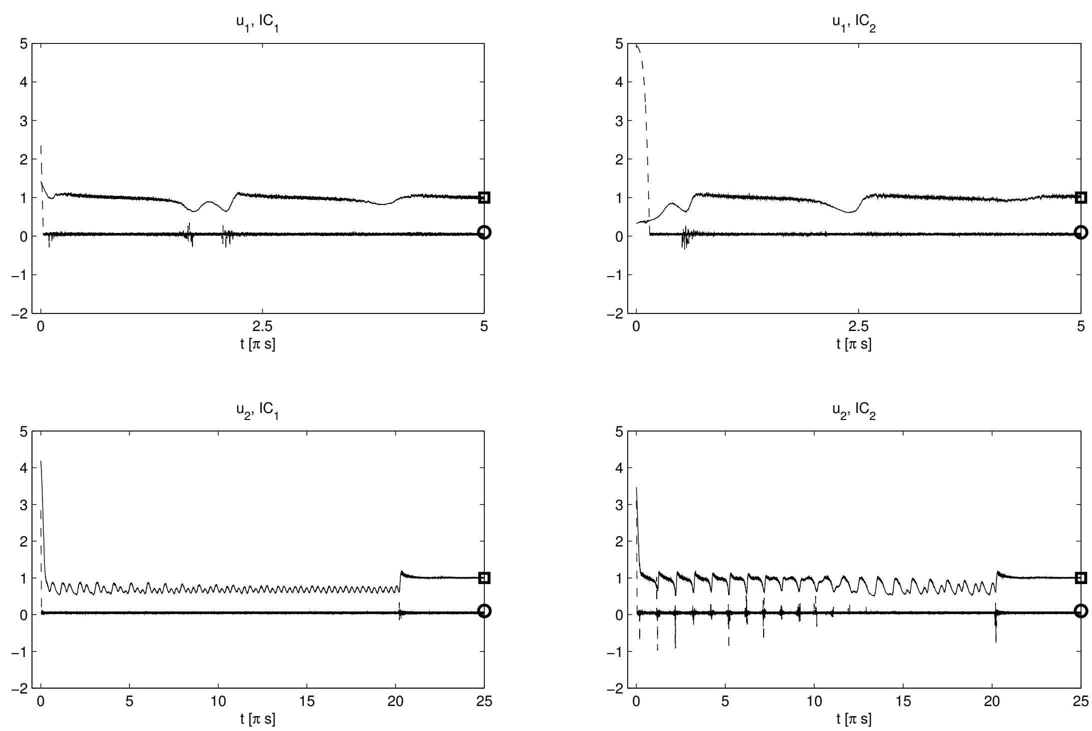

Fig. 3. Time-evolution of the estimated parameterization produced by the HNN for the two different sets of initial conditions IC and two different forces u previously considered, this time assuming that the measurement of x_1(and thus of $\dot{x}_1$, $\ddot{x}_1$) is affected by Gaussian noise. The solid and dashed lines represent respectively the estimated values for k and b; the square and the circle represent respectively the actual values of k and b. Mind the different time interval used in the last two simulations.

noise with mean $0\,[\mathrm{m}]$ and standard-deviation $0.05\,[\mathrm{m}]$, *i.e.*, the expected value of the measured displacement is x_1 and in 95% of the measurements the error is at most $0.1\,[\mathrm{m}]$. Note that $\dot{x}_1$, $\ddot{x}_1$, respectively the velocity and acceleration of cart 1, are also affected by the measurement noise since they are determined from x_1. Fig. 3 illustrates the simulation results, being clear that the parameterization estimated by the network converges in mean to the actual parameterization.

V. Conclusions and future work

In this paper, we considered the problem of parameter estimation and proposed the use of Hopfield neural networks (HNNs) to solve it. This is a preliminary work, where only time-invariant systems were considered. We assumed linearity in the parameters, a common assumption in system identification (take, for instance, the ARX structure, widely used to model system behavior); hence, our results are of general applicability. We presented a suitable HNN and a weaker sufficient condition under which the estimation error asymptotically converges to zero. Finally, we illustrated the application of our theory to the parameter estimation problem for a two carts system.

In the future, we plan to extend our results to time-variant systems. Moreover, we are interested in applying our theory to the identification of the neuromuscular blockade response to the infusion of *atracurium*, a muscle relaxant drug. This is a problem for which there exists a gray-box model that replicates very well the clinical data [6], but whose parameters have meaningful values that are difficult to estimate in practice.

VI. Acknowledgment

The first author would like to thank the Fundação para a Ciência e a Tecnologia (FCT) for the financial support (PhD grant) during the course of this project. This work was supported in part by FCT through the Unidade de Investigação Matemática e Aplicações (UIMA), Universidade de Aveiro, Portugal.

VII. References

[1] S. Abe, “Theories on the Hopfield neural networks,” in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 1, Washington, DC, USA, June 1989, pp. 557–564.

[2] M. Atencia, G. Joya, and F. Sandoval, “Parametric identification of robotic systems with stable time-varying Hopfield networks,” *Neural Computing and Applications*, vol. 13, pp. 270–280, 2004.

[3] S. Fekri, M. Athans, and A. Pascoal, “RMMAC: a novel robust adaptive control scheme - part II: performance evaluation,” in *Proceedings of the 43rd IEEE Conference on Decision and Control*, Atlantis, Paradise Island, Bahamas, December 2004, pp. 1140–1145.

[4] J. J. Hopfield, “Neurons with graded response have collective computational properties like those of two-state neurons,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 81, pp. 3088–3092, 1984.

[5] H. K. Khalil, *Nonlinear systems*, 3rd ed. New Jersey: Prentice Hall, 2002.

[6] P. Lago, T. Mendonça, and H. Azevedo, “Comparison of on-line autocalibration techniques of a controller of neuromuscular blockade,” in *Proceedings of the 4th IFAC Symposium on Modelling and Control in Biomedical Systems*, Karlsburg-Greifswald, Germany, March-April 2000, pp. 263–268.

[7] T. Mendonça and P. Lago, “PID control strategies for the automatic control of neuromuscular blockade,” *Control Engineering Practice*, vol. 6, no. 10, pp. 1225–1231, 1998.

[8] M. Norgaard, O. Ravn, N. K. Poulsen, and L. K. Hansen, *Neural networks for modelling and control of dynamic systems - a practitioner’s handbook*. London: Springer-Verlag, 2000.

[9] M. Vidyasagar, *Nonlinear systems analysis*. Philadelphia: SIAM, 2002.