International Journal of Innovative Computing, Information and Control
Volume **8**, Number **3(A)**, March **2012**

ICIC International ©2012 ISSN 1349-4198
pp. 1501–1521

NON-PARAMETRIC MODELING OF UNCERTAIN HYPERBOLIC PARTIAL DIFFERENTIAL EQUATIONS USING PSEUDO-HIGH ORDER SLIDING MODE OBSERVERS

Salvador Juárez Lopez[1], Oscar Camacho Nieto[1] and Jorge Isaac Chairez Oria[2]

[1]Centro de Investigación en Computación
National Polytechnic Institute
Av. Juan de Dios Bátiz, Delg. Gustavo Madero, México, D.F.
salvadorjuarez@sagitario.cic.ipn.mx; oscarc@cic.ipn.mx

[2]Bioprocess Department
Professional Interdisciplinary Unit of Biotechnology
National Polytechnic Institute
Av. Acueducto s/n, Barrio La Laguna, Col. Ticomán, México, D.F.
jchairez@ctrl.cinvestav.mx

Received November 2010; revised July 2011

Abstract. *There are many examples in science and engineering that may be described by a set of partial differential equations (PDEs). The modeling process of such phenomenons is in general a complex task. Moreover, there exist some sources of uncertainties around that mathematical representation that sometimes are difficult to be included in the obtained model. Neural networks appear to be a plausible alternative to get a non parametric representation of the aforementioned systems. It is well known that neural networks can approximate a large set of continuous functions defined on a compact set to an arbitrary accuracy. In this paper a strategy based on differential neural networks (DNNs) for the non parametric identification in a mathematical model described by hyperbolic partial differential equations is proposed. The identification problem is reduced to finding an exact expression for the weights dynamics using the DNN properties. The adaptive laws for weights ensure the convergence of the DNN trajectories to the hyperbolic PDE states. To investigate the qualitative behavior of the suggested methodology, here the no-parametric modeling problem for the wave equation is solved successfully. Some three dimension graphic representations are used to demonstrate the identification abilities achieved by the DNN designed in this paper.*

Keywords: Hyperbolic partial differential equations, High order sliding modes, Super-twisting, Numerical modelling

1. **Introduction.** The modern theory of identification regards to solve the problem of efficient retrieval of signal systems and dynamic properties based on measurements of available data. Basically, the class of linear and nonlinear systems whose dynamics depends linearly on the unknown parameters [22]. A general feature of these publications is that exact measurements of state vector space are available. Neural networks with universal approximation property and learning ability have proved to be a powerful tool to identify and control complex nonlinear dynamic systems with uncertainty parameters or structure [21, 25].

Exploiting the artificial neural network (ANN) natural ability to approximate nonlinear functions, the replacement of unknown system uncertainties by special adaptive models can be proposed. These numerical approximations (ANN) are defined by specific structures (continuous, discrete, etc.), but containing a number of unknown parameters (weights) that should be adjusted.

Based on the model selection, their free parameters can be modified using differential or difference equations. Many different schemes using differential forms to design ANN were proposed since 20 years ago. These constructions were based on the Hopfield structure. Today, they are recognized as differential neural networks or DNN for short. The focus of the DNN avoids the well-known problems that are common in conventional neural networks (global search minimization). Most of continuous ANN schemes use the controlled Lyapunov theory to transform the problem of numerical approximation into a robust adaptive nonlinear feedback [27].

When the mathematical model of the process under analysis is incomplete or partially known, the DNN approach provides an effective tool to address some problems in theory of control such as the parameter identification, state estimation, control for trajectory tracking. Special attention is paid to constructing differential neural identifier for dynamical systems with uncertainties that have limited information affected by external perturbations.

In contrast to many identifiers requiring a detailed mathematical description of the nonlinear systems, the ANN is efficient to deal with a large class of nonlinear systems that do not have a clearly defined model. The good performance of such identifiers [29] depends on its specific structure and the adaptive laws to adjust its parameters. In this paper, the DNN proposal given in [4, 12, 13] will be used. These DNNs have been successfully used in the identification of these unknown systems due to its massive parallelism, fast adaptation and learning capability quite success. The abilities showed by DNN have been used to approximate the right-hand side of uncertain ordinary differential equations. Indeed, the DNN can approximate such differential equations with high complexity [3].

This approximation skill may be used to solve the numerical identification of partial differential equations. The solution of the aforementioned problem is based on numerical methods such as the finite differences and finite elements methods. Therefore, the application of a mixed algorithm using a number of DNN to approximate the solution of each ODE produced by the numerical methods (finite differences in this paper) can be justified. Among the considerations that justify the use of numerical methods for solving certain types of ordinary differential equations in partial derivatives are: 1) the data from the real problems have always measurement errors, 2) the arithmetic work for the solution is limited to a finite number of significant figures resulting in rounding errors and 3) the numerical evaluation of analytical solutions is often a laborious task and computational inefficient, requiring a large number of iterations in the calculation and treatment of the data series, while generally numerical methods provide adequate numerical solutions in a simple and efficient way [2].

Indeed, several papers have showed the ANN's skills to approximate partial differential equations using the aforementioned method. In those papers, parabolic, hyperbolic and some other PDEs were successfully reproduced by a special class of continuous identifier base on special class of ANN, the so-called DNN identifier.

Even when the proposed technique is based on the well know DNN theory, the methodology introduced in the paper represents a novelty nonlinear identifier design for hyperbolic partial differential equations using a Lyapunov based method. Actually, this method connects the no-parametric approximation for uncertain second order systems defined in infinite dimensional space with the DNN scheme. One can easily conclude that the identifier designed here is useful for the dynamics considered as the object of the problem. Therefore, this design is actually a class of adaptive DNN based identifier for a class of uncertain PDE. Indeed, this class of problems has been attacked by a few researching groups as can be checked in the references section. This is the main contribution introduced in the paper.

2. **Identification of Hyperbolic Equations.**

The method used in this paper produces an approximation of uncertain hyperbolic PDE. As a result of the approach, the partial differential equation that describes the problem is replaced by a finite number of ordinary differential equations, written in terms of the values of the dependent variable at limited number of selected points. The value of the selected points are converted into the unknown, instead of the continuous spatial distribution of the dependent variable. The system of ODE must be solved and may involve a long number of arithmetic operations including the numerical integration for each ODE. The method can be solved in regressive or progressive manner.

Let us consider the uncertain hyperbolic partial differential equation

$$u_{tt}(x,t) = f\left(u(x,t), u_x(x,t), u_{xx}(x,t), v(x,t)\right) \tag{1}$$

here $u(x,t)$ is defined in a domain given by $x \in [0,1]$, $t \geq 0$, with boundary (Neumann and Dirichlet) and initial conditions given by:

$$\begin{aligned} &u_x(0,t) = 0 \\ &u(0,t) = u_0,\ u(x,0) = c \end{aligned} \tag{2}$$

The function $v(x,t) \in \Re$ can be considered as measurable external perturbation or a designed distributed control action. Indeed, the identifier based on DNN only uses the measurements at each point (x) for all times (t). In (1), and

$$\begin{aligned} u_x(x,t) &= \frac{\partial u(x,t)}{\partial x}, \quad u_{xx}(x,t) = \frac{\partial^2 u(x,t)}{\partial x^2} \\ u_t(x,t) &:= \frac{\partial u(x,t)}{\partial t}, \quad u_{tt}(x,t) := \frac{\partial^2 u(x,t)}{\partial t^2} \end{aligned}$$

System (1) armed with boundary and initial conditions (2) is driven in a Hilbert space H equipped with an inner product $(\cdot,\cdot)$.

Definition 2.1. *Let Ω be an open set in $\mathbb{R}^n$ and let $\nu \in C^m(\Omega)$. Define the norm of $\nu(x)$ as*

$$\|\nu\|_{m,p} := \sum_{0\leq|\alpha|\leq m} \left(\int_{\Omega} |D^{\alpha}\nu(x)|^p \, dx \right)^{1/p} \tag{3}$$

$(1 \leq p < \infty, D^{\alpha}\nu(x) := \dfrac{\partial^{\alpha}}{\partial x^{\alpha}}\nu(x))$. This is the Sobolev norm in which the integration is performed in Lebesgue sense. The completion of the space of function $\nu(x) \in C^m(\Omega)$: $\|\nu\|_{m,p} < \infty$ with respect to $\|\cdot\|_{m,p}$ is the Sobolev space $H^{m,p}(\Omega)$. For $p = 2$, the Sobolev space is a Hilbert space.

3. **Finite Differences Method and Mesh-Based Approximation.**

The finite difference method consists of a partial approximation of algebraic expressions involving the dependent variable values in a limited number of selected points. As a result of the approach, the partial differential equation describing the problem is replaced by a finite number of algebraic equations, written in terms of the values of the dependent variable at selected points. When the PDE is time dependent, the algebraic equations turn out to be ordinary differential equations [20].

The method produces the value of these selected points for the PDE solutions, instead of the continuous spatial distribution of the dependent variable. The system of algebraic or differential equations must be solved and may involve a long number of arithmetic operations.

Let us consider that $u := u(x,t)$ is the PDE state. In view of u is a function of x with finite and continuous derivatives, then by Taylor's theorem, one has

$$\begin{aligned} u(x+h,t) &= u(x,t) + hu_x(x,t) + \frac{1}{2}h^2u_{xx}(x,t) + \frac{1}{6}h^3u_{xxx}(x,t) + l_1(x,t) \\ u(x-h,t) &= u(x,t) - hu_x(x) + \frac{1}{2}h^2u_{xx}(x,t) - \frac{1}{6}h^3u_{xxx}(x,t) + l_2(x,t) \end{aligned} \tag{4}$$

here $l_1(x,t)$ and $l_2(x,t)$ represent terms containing powers of h or/and higher. Adding these expansions,

$$u(x+h,t) - u(x+h,t) = 2u_x(x,t) + h^2u_{xx}(x,t) + l(h)^4 \tag{5}$$

The expression $l(h^4)$ denotes the new set of terms containing the powers of order 4 or higher. Assuming these terms are small in relation to the smaller powers of h, it follows that:

$$u_{xx}(x,t)_{t\ \text{const}} \simeq \frac{1}{h^2}\left\{u(x+h,t) - 2u(x,t) + u(x-h,t)\right\} \tag{6}$$

with error de order h^2.

Subtracting the Equation (4) of Equation (5), and neglecting the terms of order h^3 may be obtained:

$$u_x(x,t)_{t\ \text{const}} \simeq \frac{1}{2h}\left\{u(x+h,t) - u(x-h,t)\right\} \tag{7}$$

with error de order h^2. Equation (5) approximates the slope of the tangent at the point These results were obtained using the supposition of t remains unchanged. This approximation called central difference approximation. This method is not usable for the approximation based on DNN.

One can also approximate the slope of the tangent at $(x,t,u(x,t))_{t\ \text{const}}$ by the slope of the line through the points $(x,t,u(x,t))$ and $(x+h,t,u(x+h,t))$, obtaining the backward difference approximation

$$u_x(x,t)_{t\ \text{const}} \simeq \frac{1}{h}\left\{u(x,t) - u(x-h,t)\right\} \tag{8}$$

or the slope of the line through the points $(x-h,t,u(x-h,t))$ and $(x,t,u(x,t))$, obtaining the forward difference approximation

$$u_x(x,t)_{t\ \text{const}} \simeq \frac{1}{h}\left\{u(x-h,t) - u(x,t)\right\} \tag{9}$$

In this paper, the backward difference will be used. This selection is easily relaxed because, the measuring availability of $u(x,t)$ is assumed to be valid.

So, it is necessary to construct a set (commonly called grid or mesh) that divides the sub-domain $x \in [0,1]$ in N equidistant sections (1) defined as x_i in such a way that $x_0 = 0$ and $x_N = 1$.

Using this mesh description, one can use the next definitions

$$u_i(t) := u_i(x,t); \quad u_{i,t}(x,t) = \left.\frac{du(x,t)}{dt}\right|_{x=x_i}; \quad u_{i,tt}(x,t) = \left.\frac{d^2u(x,t)}{dt^2}\right|_{x=x_i}$$

$$u_{i,x}(x,t) = \left.\frac{\partial u(x,t)}{\partial x}\right|_{x=x_i}; \quad u_{i,xx}(x,t) = \left.\frac{\partial^2 u(x,t)}{\partial x^2}\right|_{x=x_i}$$

Using the mesh description and applying the finite-difference representation, one has

$$u_{i,x}(x,t) \approx \frac{u_i(x,t) - u_{i-1}(x,t)}{\Delta x}, \quad u_{i,xx}(x,t) \approx \frac{u_{i,x}(x,t) - u_{i-1,x}(x,t)}{\Delta x}$$

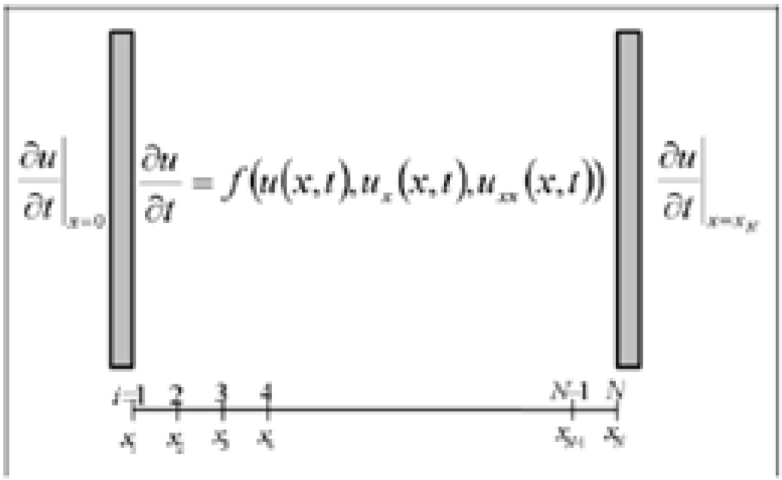

FIGURE 1. Discretization of the spacial domain. This scheme is used to solve the finite difference method to approximate the hyperbolic partial difference equation used in this paper.

The mesh based approximation of the nonlinear PDE (1) can be represented as follows:

$$u_{i,t}(t) = \dot{u}_i(t) = \Theta_i\left(u_i(t), u_{i-1}(t), u_{i-2}(t), v_i\left(x,t\right)\right) + \tilde{f}\left(u_i(t), u_{i-1}(t), u_{i-2}(t), v_i\left(x,t\right)\right)$$

where

$$\Theta_i\left(u_i(t), u_{i-1}(t), u_{i-2}(t)\right) := f_i\left(u(x_i,t), \frac{u_i\left(x,t\right) - u_{i-1}\left(x,t\right)}{\Delta x}, \frac{u_{i,x}(x,t) - u_{i-1,x}(x,t)}{\Delta x}, v_i\left(x,t\right)\right)$$

Evidently, this is a numeric method to approximate the PDE solution. However, this approximation can be formulated if f is perfectly known. This is not the case for the case considered in this paper.

4. DNN Hyperbolic Identification Using a Pseudo Sliding Mode Observer.

The system presented in (1) is a class of generalized second order system defined in the Hilbert space H. By the finite-diferences representation, each subsystem at the point i may be represented as

$$\begin{aligned} \dot{u}_{1,i}\left(t\right) &= u_{2,i}\left(x,t\right)_{x=x_i} \\ \dot{u}_{2,i}\left(t\right) &= f_i(x,v,t)_{x=x_i} \end{aligned} \tag{10}$$

That may be represented as follows

$$\begin{aligned} \dot{\mu}_i(t) &= F_i(x,v,t)_{x=x_i} \\ F_i(x,t) &:= \begin{bmatrix} u_{2,i}(x,t)_{x=x_i} \\ f_i(x,v,t)_{x=x_i} \end{bmatrix} \end{aligned} \tag{11}$$

where $\mu_i\left(t\right) = \left[u_{1,i}\left(t\right), u_{2,i}\left(t\right)\right]^{\intercal} \in \Re^2$ is the state vector and $v_i\left(t\right) \in \Re^1$ is the control action applied into the system at time t.

The solution of the last unperturbed differential equation is understood in the Filippov sense [11], that is, the second equation in (10) where appears $F_i(x,t)$ is replaced by an equivalent differential inclusion

$$\frac{d}{dt}\mu_i\left(x,v,t\right) \in \bar{F}_i\left(x,t,v\right)$$

This assumption is given to consider the possible application of discontinuous perturbations or controllers $v_i(t)$. Even when those signals wont be discontinuous, the class of solutions understandable in Filippov sense generalize the continuous case.

In view of the continuity almost everywhere of $F_i(x,v,t)$, the set-valued $\bar{F}_i(x_0,v_0) = \left[\bar{F}_i(\cdot)\right]_{i=1,n}$ is the convex closure of $F_i(x,v,t)$ of the set of all limits of $F_i(x_a,v_a)$ as $[x_a,v_a] \to [x_0,v_0]$ where $[x_0,v_0]$ is the set of all continuity points of $F_i(x,v,t)$ for any $x_a \in X \subset \Re^2$ and $v_a \in V^{adm}$ where V^{adm} is the set of all admissible nonlinear controllers given by

$$V^{adm} := \left\{ v\colon \|v_t\|^2 \le r_1 \|\mu(t)\|^2_{\Lambda_\mu} < \infty \right\}$$

This set V^{adm} may include different controllers designs such as the linear feedback, conventional and high order sliding modes, integral controllers, etc. Here $r_1 \in \Re^+$ and Λ_μ is a positive definite matrix ($\Lambda_\mu = \Lambda_\mu^\intercal$, $\Lambda_\mu > 0$) with adequate dimensions ($\Lambda_\mu \in \Re^{2\times 2}$).

5. **Finite Differences and DNN Approximation for Uncertain PDE.** Using a neural network to approximate unknown nonlinear functions $f_i(x,v,t)_{x=x_i}$ has been considered as a very important tool to solve many uninformative problems within system theory. Nevertheless, the presence of error modelling $\tilde{f}_i$ has been considered to relax the design conditions. The latter term is associated to the unavailable information to construct the numerical reproduction of $f_i(x,v,t)_{x=x_i}$.

Therefore, the following equation is valid

$$f_i(x,v,t)_{x=x_i} = f_{0,i}(x,t) + \tilde{f}_i(x,t) \tag{12}$$

The last decomposition is based on the approximation capabilities of neural networks. Here it should be noticed that $f(\cdot,\cdot) \in \Re^n$ always could be presented (by the Stone-Weisstrass and the Kolmogorov theorems [5]) as the composition of nominal $f_0(x_t,u_t \mid \Omega) : \Re^{n+m} \to \Re^n$ and a modeling error $\tilde{f}_t : \Re^n \to \Re^n$ term (as is usual when a model-free approximation is applied). The nominal part f_0 will be approximated using a nonlinear description based on neural network theory [14, 26] using general basis to reproduce the assumed unknown nonlinear function. Many possible sophisticated suggestions can be made here to design any suitable basis for the numerical approximation.

The main idea behind the application of DNN to approximate the PDE's solution is to use a class of finite-difference method but for uncertain nonlinear functions. In the introduction section, it has been established the contribution of approximation theory to construct numerical models (usually called no-parametric) based on a general approximation. The following paragraphs give a general description about how to produce a suitable approximation for each f_i.

5.1. **General theory used to approximate uncertain function.** To give a smooth approach to the proposal introduced in this paper, lets consider a continuous function $h_0(\cdot)$ defined in a Hilbert space. By the ideas presented in [9], $h_0(\cdot)$ can be rewritten in terms of the H basis (Ψ_{ij}) as

$$h_0(z,\theta^*) = \sum_i \sum_j \theta^*_{ij} \Psi_{ij}(z)$$

$$\theta^*_{ij} = \int_{-\infty}^{+\infty} h_0(z)\,\Psi_{ij}(z)\,dz,\ \forall i,j \in \mathbb{Z}$$

where $\{\Psi_{ij}(z)\}$ are functions constituting a basis in H. Last expression is referred to as a perfect function series expansion of $h_0(z,\theta^*)$. The proposed neural network corresponds to h_0. The main idea behind the neural network design is to use a similar construction but using a special class of basis functions. Nevertheless, one can not assume that θ^*_{ij} are known. So, one must propose the function aproximation $\hat{h}_0(z,\theta)$ designed to be adjusted using a time varying structure.

Based on this series expansion, the adjustable NN can take the following mathematical structure

$$\begin{aligned} &\hat{h}_0(z,\theta(t)) := \sum_{i=M_1}^{M_2}\sum_{j=N_1}^{N_2}\theta_{ij}(t)\Psi_{ij}(z) = \Theta(t)\Pi(z) \\ &\Theta(t) = [\theta_{M_1N_1}(t),\ldots,\theta_{M_1N_2}(t),\ldots\theta_{M_2N_1}(t),\ldots\theta_{M_2N_2}(t)]^\intercal \\ &\Pi(z) = [\Psi_{M_1N_1},\ldots,\Psi_{M_1N_2},\ldots\Psi_{M_2N_1},\ldots\Psi_{M_2N_2}]^\intercal \end{aligned} \tag{13}$$

that can be used to approximate a nonlinear function $h_0(z,\theta^*) \in H$ with an adequate selection of integers M_1, M_2, N_1, $N_2 \in \mathbb{Z}^+$. Following the Stone Weiestrass Theorem [5], if

$$\epsilon(M_1,M_2,N_1,N_2,t) = h_0(x,\theta^*) - \hat{h}_0(x,\theta(t))$$

is the NN approximation error. Then for any arbitrary positive constant ϵ there are some constants $M_1, M_2, N_1, N_2 \in \mathbb{Z}$ such that for all $z \in Z \subset \Re$.

$$\sup_t \|\epsilon(M_1,M_2,N_1,N_2,t)\|_2 \le \epsilon \tag{14}$$

Remark 5.1. *Appropriate selection of functions $\Psi_{ij}(\cdot)$ is an important task to construct an adequate approximation of nonlinear functions. Many functions have been reported in literature* [24] *that have remarkable results to approximate nonlinear unknown functions. Which one is the most suitable basis in practical application depends on each particular design specifications.*

Remark 5.2. *M_1, M_2, N_1, N_2 parameters in neural network design are closely related to the quality approximation $\epsilon(M_1,M_2,N_1,N_2,t)$. The NN has been demonstrated to be effective to reproduce uncertain nonlinear functions satisfying the Lipschitz condition.*

5.2. **DNN approximation for hyperbolic PDE.** Following the ideas presented above, by simple adding and subtracting the necessary terms, one can represent the hyperbolic PDE as

$$\begin{aligned} u_{tt}(x,t) =& A\mu(x,t) + \mathring{V}_1(x)\bar{\sigma}(x)\mu(x,t) + \mathring{V}_2(x)\bar{\varphi}(x)\mu_x(x,t) \\ &+ \mathring{V}_3(x)\bar{\gamma}(x)\mu_{xx}(x,t) + \mathring{V}_4(x)\bar{\eta}(x)v(x,t) + \tilde{f}(x,t) \end{aligned} \tag{15}$$

here $\tilde{f}(x,t) \in R$ represents the modeling error and is defined explicitly by

$$\tilde{f}(x,t) := f(u(x,t), u_x(x,t), u_{xx}(x,t)) - u_{tt}(x,t)$$

The vectors $A \in R^{1\times 2}$, $\mathring{V}_1(\cdot) \in R^{n_1}$, $\mathring{V}_2(\cdot) \in R^{n_2}$, $\mathring{V}_3(\cdot) \in R^{n_3}$, $\mathring{V}_4(\cdot) \in R^{n_4}$ are constants and the set of functions $\bar{\sigma}(x) \in R^{n_1}$, $\bar{\varphi}(x) \in R^{n_2}$, $\bar{\gamma}(x) \in R^{n_3}$, $\bar{\eta}(x) \in R^{n_4}$ forming a basis obey the following sector conditions:

$$\begin{aligned} &\|\bar{\sigma}(x) - \bar{\sigma}(x\prime)\| \le L_{\bar{\sigma}}\|x - x\prime\| \qquad \|\bar{\varphi}(x) - \bar{\varphi}(x\prime)\| \le L_{\bar{\varphi}}\|x - x\prime\| \\ &\|\bar{\gamma}(x) - \bar{\gamma}(x\prime)\| \le L_{\bar{\gamma}}\|x - x\prime\| \qquad \|\bar{\eta}(x) - \bar{\eta}(x\prime)\| \le L_{\bar{\eta}}\|x - x\prime\| \end{aligned}$$

Also it can be shown they are bounded in x, i.e.,

$$\|\bar{\sigma}(\cdot)\| \le \sigma^+,\ \|\bar{\varphi}(\cdot)\| \le \varphi^+,\ \|\bar{\gamma}(\cdot)\| \le \gamma^+,\ \|\bar{\eta}(\cdot)\| \le \eta^+$$

5.3. **DNN approximation based on the finite differences.** Following the DNN constructions and applying the same representation to (15), one gets for each $i \in [1, N]$:

$$u_{i,tt}(x,t) = f_{0,i}(x,t) + \tilde{f}_i(x,t)$$

here the term $f_{0,i}(x,t)$ is usually referred to as the nominal dynamics or the DNN approximation. This structure obeys the basic regressor form described in [22]. Therefore, the $f_{0,i}(x,t)$ has the form

$$\begin{aligned} f_{0,i}(x,t) := {} & A\mu(x,t)|_{x=x_i} + V_1(x)\,\bar{\sigma}(x)\mu(x,t)|_{x=x_i} + V_2(x)\,\bar{\varphi}(x)\mu_x(x,t)|_{x=x_i} \\ & + V_3(x)\,\bar{\gamma}(x)\mu_{xx}(x,t)|_{x=x_i} + V_4(x)\,\bar{\eta}(x)v(x,t)|_{x=x_i} \end{aligned}$$

On the other hand, the term

$$\tilde{f}_i(x,t) := \left.\tilde{f}(x,t)\right|_{x=x_i}$$

is the so-called modeling error representing the distance between the approximation produced by the DNN and the real PDE trajectories at each point within the space domain $[0,1]$. By a simple mathematical algorithm, one has

$$\tilde{f}_i(x,t) = R_i\left(u_i(t), u_{i-1}(t), u_{i-2}(t)\right) - f_{0,i}(x,t)$$

Hereafter, it will be assumed that the modeling error terms satisfy the followings assumptions:

Assumption: The modelling error is absolutely bounded in Ω :

$$\left\|\tilde{f}_i(x,t)\right\|^2 \le \tilde{f}_{1,i}\left\|\mu_i(t)\right\|^2 \tag{16}$$

Direct application of finite difference method to (12) leads to

$$\begin{aligned} f_{0,i}(x(t),t) := {} & A_i\mu_i(x,t) + \Big[\mathring{V}_{1,i}(x)\,\bar{\sigma}_i(x_i) + (\Delta x)^{-1}\mathring{V}_2(x_i)\,\bar{\varphi}(x_i) \\ & + (\Delta x)^{-2}\mathring{V}_3(x_i)\,\bar{\gamma}(x_i)\Big]\mu_i(x,t) - \Big[(\Delta x)^{-1}\mathring{V}_2(x_i)\,\bar{\varphi}(x_i) \\ & + 2(\Delta x)^{-2}\mathring{V}_3(x_i)\,\bar{\gamma}(x_i)\Big]\mu_{i-1}(x,t) + \Big[(\Delta x)^{-2}\mathring{V}_3(x_i)\,\bar{\gamma}(x_i)\Big]\mu_{i-2}(x,t) \\ & + \mathring{V}_4(x_i)\,\bar{\eta}(x_i)v(x_i,t) \end{aligned}$$

Indeed by the approximation theory introduced above, one may represent this expression as

$$f_{0,i}(x(t),t) := \Omega_{0,i}^{\intercal}(t)\,\Pi_i(t)$$

These new variables are defined as

$$\begin{aligned} \Omega_{0,i}^{\intercal}(t) &:= \begin{bmatrix} A_i & \mathring{W}_{i,1}(x) & \mathring{W}_{i,2}(x) & \mathring{W}_{i,3}(x) & \mathring{W}_{i,4}(x) \end{bmatrix} \\ \Pi_i^{\intercal}(t) &:= \begin{bmatrix} \mu(x_i,t) & \sigma(x_i)\mu(x_i,t) & \varphi(x_i)\mu(x_{i-1},t) & \gamma(x_i)\mu(x_{i-2},t) & \eta(x_i)v(x_i,t) \end{bmatrix} \end{aligned}$$

This is the so-called finite difference DNN approximation of the uncertain hyperbolic partial differential equation.

Considering the structure given in (10) and using the DNN description, PDE may be approximated by the following second order DNN identifier

$$\begin{aligned} \dot{u}_{1,i}(t) &= u_{2,i}(x,t) \\ \dot{u}_{2,i}(t) &= \Omega_{0,i}^{\intercal}(t)\,\Xi_i(t) + \tilde{f}_i(x,t) \end{aligned} \tag{17}$$

This form will be used to show the convergence of the DNN identifier.

5.4. **Non-parametric identifier.** In this case, following the neural network theory, the nominal section is proposed as

$$\begin{aligned}
&\bar{f}_{0,i}\left(x\left(t\right),t\right):=\Omega_{i}^{\intercal}\left(t\right)\bar{\Pi}_{i}\left(z\right)\\
&\bar{\Theta}_{i}^{\intercal}\left(t\right):=\left[\begin{array}{ccccc} A_{i} & V_{1}\left(x,t\right)|_{x=x_{i}} & V_{2}\left(x,t\right)|_{x=x_{i}} & V_{3}\left(x,t\right)|_{x=x_{i}} & V_{4}\left(x,t\right)|_{x=x_{i}}\end{array}\right]\\
&\bar{\Pi}_{i}^{\intercal}\left(t\right):=\left[\begin{array}{c} \mu(x,t)|_{x=x_{i}} \quad \bar{\sigma}(x)\mu(x,t)|_{x=x_{i}}\\ \bar{\varphi}(x)\mu_{x}(x,t)|_{x=x_{i}} \quad \bar{\gamma}(x)\mu_{xx}(x,t)|_{x=x_{i}} \quad \bar{\eta}(x)v(x,t)|_{x=x_{i}}\end{array}\right]
\end{aligned}\tag{18}$$

where $V_j\left(x,t\right)\in\Re^{n_j}$ are constant matrices which are the so-called best-fitted weights and are defined as follows

$$V_j\left(x,t\right)|_{x=x_i}:=V_j\left(x_i,t\right)$$

The activation functions $\bar{\Pi}_i\left(z\right)$ are constituted by the usual sigmoid functions defined as

$$S_r(x_t):=a_r\left(1+b_r\exp\left(-\sum_{j=1}^{n}c_jx_{j,t}\right)\right)^{-1},\quad r=[1,n_j]$$

These functions satisfy the following sector conditions

$$\begin{aligned}
&\left|S_r\left(x_t^1\right)-S_r\left(x_t^2\right)\right|^2\le l_{S_r}\left\|x_t^1-x_t^2\right\|^2\\
&x_t^1,x_t^2\in\mathbb{R}^l,\ l\ge 1\ l_{S_r}\in\mathbb{R}^+
\end{aligned}$$

The upper bound (16) is guaranteed if the uncertain nonlinear system (17) can be approximated by a possible adaptive algorithm. This property is usually referred to as that system is identifiable [14].

Direct application of finite difference method to (18) leads to

$$\begin{aligned}
\bar{f}_{0,i}\left(x\left(t\right),t\right):=\;&A_i\mu_i(x,t)+\left[V_{1,i}\left(x,t\right)\bar{\sigma}_i(x_i)+\left(\Delta x\right)^{-1}V_2\left(x_i,t\right)\bar{\varphi}(x_i)\right.\\
&\left.+\left(\Delta x\right)^{-2}V_3\left(x_i,t\right)\bar{\gamma}(x_i)\right]\mu_i(x,t)-\left[\left(\Delta x\right)^{-1}V_2\left(x_i,t\right)\bar{\varphi}(x_i)\right.\\
&\left.+2\left(\Delta x\right)^{-2}V_3\left(x_i,t\right)\bar{\gamma}(x_i)\right]\mu_{i-1}(x,t)\\
&+\left[\left(\Delta x\right)^{-2}V_3\left(x_i,t\right)\bar{\gamma}(x_i)\right]\mu_{i-2}(x,t)\\
&+V_4\left(x_i,t\right)\bar{\eta}(x_i)v(x_i,t)
\end{aligned}$$

Indeed by the approximation theory introduced above, one may represent this expression as

$$\bar{f}_{0,i}\left(x\left(t\right),t\right):=\Omega_i^{\intercal}\left(t\right)\bar{\Pi}_i\left(t\right)$$

These new variables are defined as

$$\begin{aligned}
&\Omega_i^{\intercal}\left(t\right):=\left[\begin{array}{ccccc}A_i & W_{i,1}\left(x,t\right) & W_{i,2}\left(x,t\right) & W_{i,3}\left(x,t\right) & W_{i,4}\left(x,t\right)\end{array}\right]\\
&\bar{\Pi}_i^{\intercal}\left(t\right):=\left[\begin{array}{ccccc}\hat{\mu}(x_i,t) & \sigma(x_i)\hat{\mu}(x_i,t) & \varphi(x_i)\hat{\mu}(x_{i-1},t) & \gamma(x_i)\hat{\mu}(x_{i-2},t) & \eta(x_i)v(x_i,t)\end{array}\right]\\
&W_{i,j}\left(x,t\right):=W_j\left(x,t\right)|_{x=x_i}
\end{aligned}$$

This structure is the DNN identifier for the uncertain partial differential equation based on the finite difference method.

As one can understand, the real PDE identifier will be constituted by N identifiers working each one at the specific point x_i. Evidently, the final approximation will be obtained by the usual interpolation algorithm used in the reconstruction of the final solution. Using a similar method to that described in the previous section, one has:

$$\begin{aligned}
&\frac{d}{dt}\bar{u}_{1,i}\left(t\right)=\bar{u}_{2,i}\left(x,t\right)\\
&\frac{d}{dt}\bar{u}_{2,i}\left(t\right)=\bar{f}_{0,i}\left(x\left(t\right),t\right)
\end{aligned}\tag{19}$$

6. **Adaptive DNN Pseudo-Observer.** One should note that measurement conditions just allow to obtain the value of $\bar{u}_{1,i}(t)$. Therefore, the proposed identifier introduced in (19) cannot be developed. So, in this paper is proposed the application of the so-called DNN pseudo-observer. This idea comes from the similarity that structure showed in (19) has with second order nonlinear systems. Among others, mechanical systems are good examples of such systems. These mechanical systems where the nonlinear dynamics $f_i(x, v, t)_{x=x_i}$ can be explicitly described using the Euler-Lagrange method. Besides, many electromechanical devices such as the induction motor (IM) or high-power generators obey a nonlinear dynamics with similar structure.

The state estimation of this kind of systems has attracted a large amount of research efforts [8], specially when there are not mechanical sensors for the measurement of speed or position. This is a consequence of the difficulty to measure simultaneously both variables $\bar{u}_{1,i}(t)$ and $\bar{u}_{2,i}(t)$ resulting in the natural proposition to construct adaptive observers. Moreover, the complex nonlinear structure makes the state estimation a real challenge for designers. These inconveniencies have been solved by many adaptive observer proposals: back-stepping, conventional sliding mode, adaptive observers, optimal techniques, etc.

Observers are dynamic systems from the information of a plant (known model structure and input and output variables available from measurements), estimated variables (known as states) or parameters are not known or directly measurable. There is a wide range of systems for which no observers can be designed using the standard theory, since it requires the Lipschitz condition for the existence and uniqueness of solutions implies that the functions defining the system are continuous at all values states. Among the systems that are defined by discontinuous or multivalued functions, and therefore do not meet the Lipschitz condition, including some as common as mechanical stiction or hysteresis, as well as hybrids, which are now widely studied. For these systems work in the design of observers has been low, partly due to the mathematical complexity inherent in the inability to use the traditional theory.

In the literature on nonlinear observers are various uses of discontinuous nonlinear design. For example, in sliding mode observers by introducing discontinuous injection terms in order to improve performance. There are also some published works on the design of observers for systems that explicitly include discontinuous or multivalued nonlinearities. The approach mentioned in these publications strongly restricts the type of nonlinear systems with discontinuous or multivalued, since it requires the uniqueness of solutions, for which, the nonlinearities must be monotonous. This is a very strong requirement, since the uniqueness of solutions in systems described by differential inclusions is rather the exception than the rule.

The theory of differential inclusions developed in several decades, to suggest an appropriate and unified systems that include discontinuous or multivalued nonlinearities. Differential inclusion in the map that defines the dynamics of the system assigns each element of the domain is not an element of the codomain, but a subset of it, in what is called a multivalued function. Traditional differential equations are incorporated in this formulation, considering each image as a set of one element. Since the existence (but not uniqueness) of solutions of differential inclusions requires the image of each element of the domain is convex and compact, discontinuous or multivalued functions, often encountered in practice, must be adjusted to meet such conditions and can be treated by the theory of differential inclusions [18, 19].

Differential neural observers are studied in the approach of sliding mode (SM) is used to obtain the algebraic learning procedure for online identification of nonlinear plant (design model) fully available to states. The essential feature of the SM technique is the application of discontinuous feedback laws to achieve and maintain the closed-loop dynamics of a given variety in the space of states (for switching from the surface), with some desired properties for paths system [30, 34]. This method offers many advantages over other identification and control techniques with a good transient behavior, the need for a reduced amount of information compared to traditional control techniques, not as a model with capacity for disturbance rejection, insensitive to plant nonlinearity or parameter variations, a remarkable stability and performance robustness.

In general there are some nice results to design several possible observers for the class of nonlinear systems given in (17), even using the high-order sliding mode (HOSM) technique [15]. However, many of them requires the complete knowledge on $f_i(x, v, t)_{x=x_i}$ or it should be admitted that $\|f_i(x, v, t)_{x=x_i}\|$ is bounded (there exists a constant $f_i^+ \in \Re^+$ such that $\|f_i(x, v, t)_{x=x_i}\| \leq f_i^+$, $\forall i \in [1, N]$). Some papers have shown the ability of HOSM to achieve an excellent reproduction of the unknown state $\bar{u}_{2,i}(t)$ using just the f_i^+ information [6]. On the other hand, the nice properties of the SM method to be invariant to some sort of uncertainties could be undesirable because if there exists the possibility to design any control function using these estimated states, the domain where the control action is valid may reach big values that can not be realizable by the actuators and moreover, to provoke overheating in the power amplifiers for example (which are devoted to manage the DC motors, steppers motors and others) [10, 33].

Some attractive features of SOSM compared to the classical first-order sliding modes are widely recognized: higher accuracy motions, chattering reduction , finite-time convergence for systems with relative degree two [17, 19], etc. In most cases, sliding modes are obtained by the injection of a non-linear discontinuous term, depending of the output error. This design may be used to construct robust controlling or observing algorithms. The discontinuous injection must be designed in such a way that system trajectories are enforced to remain in a submanifold contained in the estimation error space (the so-called sliding surface). For both, the control and the observation problem, the resulting motion is referred to as the sliding mode [32]. This discontinuous term enables the rejection of external matched disturbances [31].

Sliding modes observers are widely used because can provide finite-time convergence, robustness with respect to perturbations and uncertainties estimation [1, 7]. A new generation of observers based on the so-called second-order sliding-modes has been recently developed [16]. In [17], robust exact differentiators were performed. That observer based on the so called super twisting algorithm ensures finite time convergence to the real trajectories without filtration or numeric derivation.

In [7], a second order sliding mode observer based on a modification of the super-twisting algorithm is proposed to observe a large class of mechanical systems. A discrete version of such observer (via the Euler Scheme) is also presented; its finite time convergence is proved by means of majoring curves. In the same sense, in [18], it is shown that finite differences are applicable to the on-line estimation of arbitrary-order derivatives in homogeneous discontinuous control.

Another important contribution in the field of the stability analysis for SOSM observers was made by [23]. In that paper, a strong Lyapunov function for a class of algorithms of SOSM is obtained. Additionally, a modified version of the super-twisting algorithm is implemented adding a proportional term in its structure. This term helps the estimation process.

6.1. **Pseudo-observer structure.** Following the non-parametric state identification methods, the adaptive observation scheme is composed by an adaptive reproduction of the nominal unknown section $f_0(\cdot, \cdot \mid \cdot)$ and a set of corrective terms using the available information for the uncertain system, that is the output signal. The proposal given here deals with both adaptive sections using the DNN theory to reproduce the vague mathematical description for the nonlinear system and the high order sliding mode method to avoid the chattering presence on the observer trajectories. The state estimator description uses the approaches developed in [6, 29]. Those techniques showed, independently, a great capability to reconstruct the unknown states for an uncertain nonlinear system affected by perturbations. However, both have some inconveniencies associated to, firstly, for the DNN the difficulty associated with the fast convergence between the measurable states, the corresponding estimates and second for the inability to provide a good approximation for the structure of the doubtful section of the nonlinear system.

A new structure mixing the abilities from these pair of methods is the basement of the observer:

$$\begin{aligned} &\frac{d}{dt}\hat{u}_{1,i}\left(t\right) := \hat{u}_{2,i}\left(t\right) + \beta_1\lambda\left(\tilde{u}_{1,i}\left(t\right)\right)\operatorname{sign}\left(\tilde{u}_{1,i}\left(t\right)\right) \\ &\frac{d}{dt}\hat{u}_{2,i}\left(t\right) := \Omega_i^{\intercal}\left(t\right)\bar{\Pi}_i\left(t\right) + \beta_2\operatorname{sign}\left(\tilde{u}_{1,i}\left(t\right)\right) \\ &\tilde{u}_{1,i}\left(t\right) := u_{1,i}\left(t\right) - \hat{u}_{1,i}\left(t\right) \end{aligned} \tag{20}$$

Here, the weights matrices (Ω_i) provide the time varying adaptive behavior to this class of observers. This observer uses a training method that is executed on-line to adjust the weights to improve the current representation of (17) by DNN. This mixed alternative gives a second level of robustness under external perturbations and modelling uncertainties. As it was previously stated, the matrices Ω_i are responsible to reproduce the unknown dynamics.

The solution of (20) should be understood in the Filippov sense. That means the pseudo-observer generates a set of trajectories depending on the definition applied for the sign function. In this paper, the sign function is defined by:

$$\operatorname{sign}\left(z\right) := \begin{cases} +1 & \text{if } x > 0 \\ [-1, +1] & \text{if } x = 0 \\ -1 & \text{if } x < 0 \end{cases}$$

Therefore, the strong stability concept will be used to show that all trajectories in (20) converge to the set of uncertain ODE presented in (17).

6.2. **Learning laws for identifier weights.** For each $i = 0, \cdots, N$ define the vector-functions defining the error between the trajectories produced by the model and the DNN-identifier as well as their derivatives with respect to x for each i.

$$\tilde{u}_i\left(t\right) := \hat{u}_i\left(t\right) - u_i\left(t\right), \tag{21}$$

Let the weights matrices satisfy the following nonlinear matrix differential equations

$$\dot{W}_{i,k}\left(t\right) := \Phi_{i,k}\left(\tilde{u}_i\left(t\right), \tilde{W}_{i,k}\left(t\right)\right) \tag{22}$$

where

$$\begin{aligned} &\Phi_{k,i}(\tilde{u}_{1,i}(t), \tilde{W}_{k,i}(t)) := -k_{k,i}\Sigma_3^{\intercal}C^{\intercal}M_2\zeta(t)\Xi_{k,i}^{\top}(\hat{u}, t) - \\ &- k_{k,i}\Sigma_3^{\intercal}\Lambda_c\Sigma_3\tilde{W}_{k,i}\left(x, t\right)\Xi_{k,i}(\hat{u}, t)\Xi_{k,i}^{\top}(\hat{u}, t) - \alpha\tilde{W}_{k,i}\left(x, t\right) \\ &\Sigma_3 = NP_iM_3, \quad M_2 := [0 \;\; 1 \;\; 0]^{\top}, \quad M_3 := [0 \;\; 0 \;\; 1]^{\top} \\ &\Xi_{1,i}(t) := \sigma(x_i)\hat{u}_i(t), \quad \Xi_{2,i}(t) := \varphi(x_i)\hat{u}_{i-1}(t) \\ &\Xi_{3,i}(t) := \gamma(x_i)\hat{u}_{i-2}(t), \quad \Xi_{4,i}(t) := \eta(x_i)v(x_i, t) \end{aligned}$$

with positive constants $k_{i,k} > 0$ $\left(k = \overline{1,4}\right)$ and P_i $(i = \overline{0,N})$ which are positive definite and symmetric solutions $((P_i)^\intercal = P_i > 0)$ of the algebraic Riccati equations defined as follows

$$Ric\,(P_i) := P_i M A_i + [MA_i]^\intercal P_i + P_i R_i P_i + Q_i^P = 0 \tag{23}$$

where

$$\begin{aligned}
R_i^P &:= M\mathring{W}_{i,1}(x)\Lambda_1^{-1}\mathring{W}_{i,1}^\intercal M^\intercal \\
&\quad + M\mathring{W}_{i,2}(x)\Lambda_2^{-1}\mathring{W}_{i,2}^\intercal M^\intercal + M\mathring{W}_{i,3}(x)\Lambda_3^{-1}\mathring{W}_{i,3}^\intercal M^\intercal \\
&\quad + M\mathring{W}_{i,4}(x)\Lambda_4^{-1}\mathring{W}_{i,4}^\intercal M^\intercal + \frac{1}{2}N\Lambda_5 N^\intercal \\
&\quad + M\Lambda_f M^\intercal + \lambda_{\max}\left\{\Lambda_f^{-1}\right\}\tilde{f}_{1,i} \\
Q_i^P &:= \alpha_\sigma I_{n\times n} + \alpha_\nu I_{n\times n} + d^2 O^\intercal I_{n\times n} O \\
M &:= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^\intercal
\end{aligned}$$

The learning laws (22) have been obtained using the Lyapunov methodology as it will be explained in the main theorem of this paper as well as in the appendix. One must note that learning laws derived by the methodology suggested in this paper depends only on t. Once the training method has been completed, one can turn off the learning law while the identifier structure is fixed with the weights values generated after the training.

Remark 6.1. *The Special class of Riccati equation*

$$PA + A^\intercal P + PRP + Q = 0$$

has a unique positive solution P if and only if [29] *the following four conditions given below are fulfilled: 1) Matrix A is stable; 2) Pair $\left(A, R^{1/2}\right)$ is controllable; 3) Pair $\left(Q^{1/2}, A\right)$ is observable, and 4) Matrices $(A,\ Q,\ R)$ should be selected in such a way to satisfy the following inequality*

$$\frac{1}{4}\left(A^\intercal R^{-1}\text{-}R^{-1}A\right) R \left(A^\intercal R^{-1}\text{-}R^{-1}A\right)^\intercal + Q \le A^\intercal R^{-1} A$$

which restricts the largest eigenvalue of R guarantying the existence of a unique positive solution.

Theorem 6.1. *Consider the non linear model (1), given by the system of PDE's with uncertainties (perturbations) in the states and the outputs, under the border conditions (2). Let also suppose that parameters in the DNN-identifier given by (10) are adjusted by the learning laws (22). If positive definite matrices Q_i^P provide the existence of positive solutions P^i $(i = 0, \cdots, N)$ to the Riccati Equation (23), and for any positive scalar $d \in \Re^+$ such that the following LMI has solution for some $H_i > 0$, $H_i^\intercal = H_i \in \mathbb{R}^{n\times n}$*

$$\left(P_i \begin{bmatrix} -\beta_1 & 0 & 1 \\ 0 & -d & 0 \\ -r_1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -\beta_1 & 0 & 1 \\ 0 & -d & 0 \\ -r_1 & 0 & 0 \end{bmatrix}^\intercal P_i \right) \le H_i \tag{24}$$

then the origin

$$\hat{u}_i(t) - u_i(t) = 0,\ \tilde{W}_{i,k}(t) = 0$$

is an equilibrium point that is strongly asymptotically stable. The detailed proof is given in Appendix.

6.3. Lyapunov-based strategy to proof the identifier convergence. The Lyapunov function candidate is

$$V(t) = \sum_{i=1}^{N} \|\zeta_i(t)\|_{P_i}^2 + \sum_{i=1}^{N}\sum_{j=1}^{4}\left[k_{i,j}\mathrm{tr}\left\{\tilde{W}_{i,j}^T(t)\,\tilde{W}_{i,j}(t)\right\}\right]$$

The vector $\zeta_i(t)$ is defined as follows

$$\zeta_i(t) := \begin{bmatrix} |\tilde{u}_{1,i}(t)|^{1/2}\,\mathrm{sign}(\tilde{u}_{1,i}(t)) \\ \tilde{u}_{1,i}(t) \\ \tilde{u}_{2,i}(t) \end{bmatrix}$$

In [23], it has been show the characteristics of this class of strict Lyapunov functions. This function is actually differentiable, therefore a non-smooth option of Lyapunov theory is required. Indeed, one must note that usual generalized gradient can not be used in here. Besides, the first element on $\zeta_i(t)$ cannot be differentiated using the product rule.

7. Simulation Results. This section shows the results for the identification algorithm in two different systems. We have selected the utilization of two different uncertain hyperbolic PDE. These both systems fulfill the necessary conditions required to achieving the results for the identification using the designed pseudo-sliding mode observer based on the DNN methodology.

1. Below, the numerical simulations show the qualitative illustration for a benchmark system. Let us consider the simplified problem of a vibrating string. We represent the position of a point, at a instant t, by a continuous real function $u : [0; L] \times [0; 1)$, where $[0; L]$ represents the string in the reference frame. Let's consider the following distributed parameter system representing the vibrating string model

$$\begin{aligned} &u_{tt}(x,t) = 0.01u_{xx}(x,t) + 0.01\sin(x,t) \\ &u_x(0,t) = 0, \quad u_x(1,t) = 0 \end{aligned}$$

This model assumes that you have access to discrete measures of the state $\sin(x,t)$ along its entire domain. This model will be used just to generate the data required to test the identifier based on DNN. The previous boundary conditions correspond to the situation where the end of the string at $x = 1$ is pinned at the end $x = 0$ is free. The zero-slope boundary condition at $x = 0$ has the physical meaning of no force being applied at that end. The parameters used within the identifier for the simulations were selected as follows:

$$A = \begin{bmatrix} 0 & 1 \\ -88.51 & -95.62 \end{bmatrix}, \quad P = \begin{bmatrix} 0.004 & 0 \\ 0 & 0.0028 \end{bmatrix}$$

$$S = \begin{bmatrix} 5^{-7} & 0 \\ 0 & 5^{-7} \end{bmatrix}, \quad T = \begin{bmatrix} 2^{-7} & 0 \\ 0 & 2^{-7} \end{bmatrix}$$

$$W_{1,i}(0) = W_{2,i}(0) = W_{3,i}(0) = \begin{bmatrix} 31.78 \\ 39.2 \\ 23.8 \end{bmatrix}$$

DNN identifier generates the trajectories which are very close to the real trajectories of the system (see Figure 2). The identifier state produced by the DNN identifier is shown in Figure 3. The trajectories depicted in that figure were generated by the pseudo-observer proposed in this paper. One can see the closeness between both Figures 2 and 3 that is a consequence of the proposed variable structure identifier. The error between trajectories produced by the model and the proposed identifier is

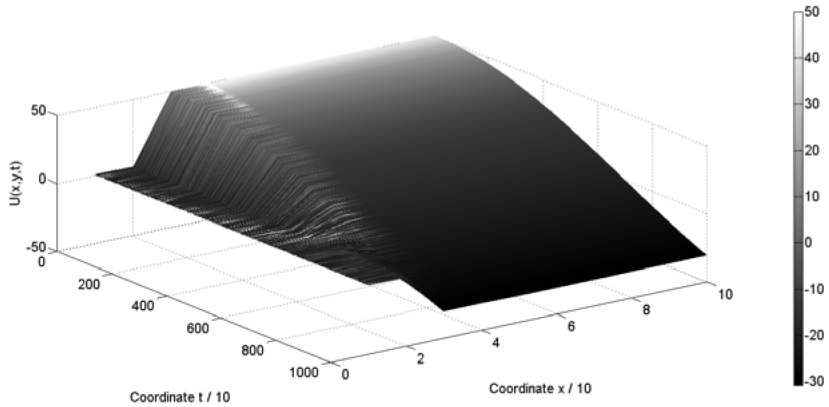

Figure 2. Trajectories of the hyperbolic partial differential equation obtained by the finite difference method. These trajectories were generated using the parameters defined in the first numerical problem without perturbations.

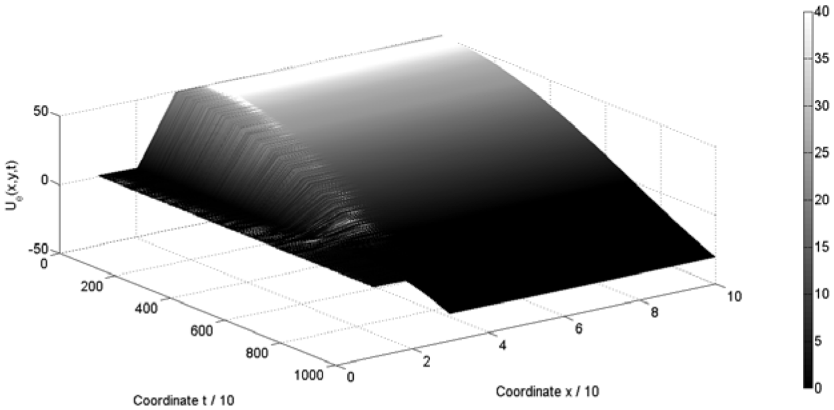

Figure 3. Trajectories of the DNN identifier obtained by the finite difference method. These trajectories were generated using the parameters defined in the identifier considering the problem.

close to zero almost for all x and all t that shows the efficiency of the identification process provided by the suggested DNN algorithm. This similarity demonstrates the possibility to approximate with high accuracy the solution of an uncertain hyperbolic partial differential equation. Here one can see the convergence to a small zone near to zero which is defined by the uncertainties considered in the hyperbolic partial differential equation.

2. The second system is a little bit more complex than the previous one.
This was used to demonstrate that is possible with the pseudo-observer stability and to identify any other system through DNN.

$$u_{tt}(x,t) = Du_{xx}(x,t) + Vu_x(x,t) + \Gamma u(x,t) + A\sin(x,t)$$

It assumes that you have access to discrete measures of the state $\sin(x,t)$, u_{xx}, u_x along its entire domain and so does u. This model will be used just to generate the data to test the identifier based on DNN. It is assumed that initial conditions have being fixed just like in the previous example. The parameters used in the simulation

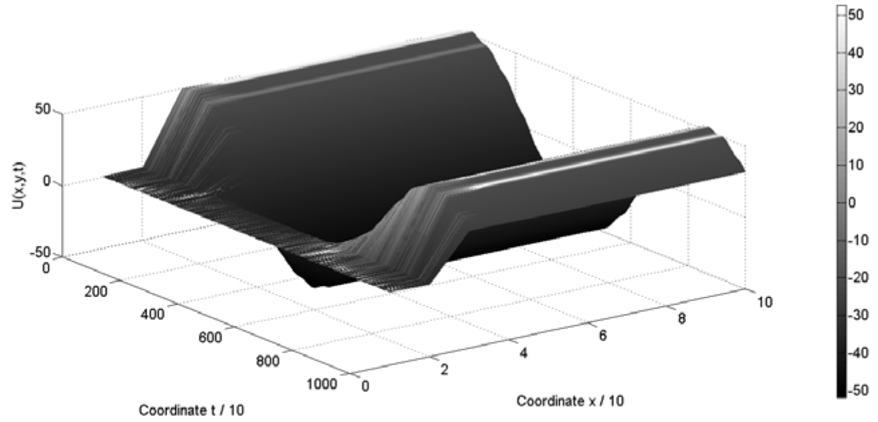

FIGURE 4. Trajectories of the hyperbolic partial differential equation obtained by the finite difference method. These trajectories were generated using the parameters defined in the second numerical problem with perturbations.

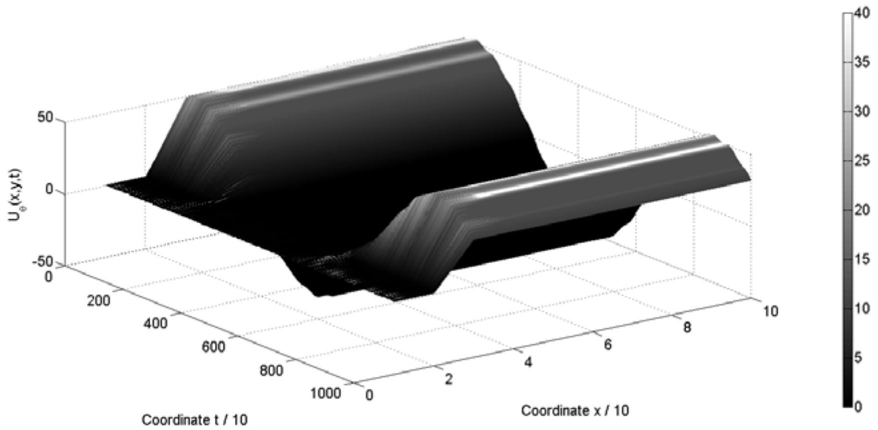

FIGURE 5. Trajectories of the DNN identifier obtained by the finite difference method. These trajectories were generated using the parameters defined in the identifier considering the problem with perturbations.

are:

$$D = .0051, \quad A = 0.0012$$
$$V = 0.021, \quad \Gamma = 0.02$$

The following numerical results were achieved using the previous parameters. These trajectories are used just as data generator. No information regarding the model is used in the identifier definition. Considering the effect as the external perturbation $A\sin(x,t)$ and the natural oscillating trajectory associated to this system, the numerical simulation leads to the following portrait (Figure 4). The trajectory produced by the identifier is showed in Figure 5. One can see the similarities between both trajectories: the model and the identifier. Even when they are really similar, the grayscale defining the trajectories demonstrates small variations which are natural consequences of the adaptive scheme used by the identifier to produce its trajectory.

8. **Conclusions.** This paper has shown that there are some methods, such as the DNN, to approximate an uncertain system. This methodology is useful when only a few parameters are available for study. The parameters known such as the boundary conditions,

are used to approximate the value of the neighbor point. This process is repeated several times until the center of the space of study is reached. This way the nature of the system is known even though the model is not known. It has been proved that DNN are useful to know the behavior of an uncertain system and even abnormalities could be discovered through this method. Due to the fact that not all the mathematical models of diseases and pathologies are known, an application of this methodology to model new systems is a useful way of research areas of medicine not explored yet. The suggested approach solves the problem of non parametric identification of uncertain nonlinear described by hyperbolic partial differential equations. Asymptotic convergence for the identification error has been demonstrated applying a Lyapunov-like analysis using a special class of Lyapunov functional. Besides, the same analysis leads to the generation of the corresponding conditions for the upper bound of the weights involved in the identifier structure. Identifier structure is based on a pseudoobserver constructed within the linear observers framework. Numerical example showing the beam dynamics demonstrates the workability of this new methodology based on continuous neural networks.

Acknowledgement. Authors want to acknowledge the support given by the projects IPN-SIP-COFAA-CIC20113090, ICYTDF243/2010 and 277/2010 and CONACYTI0110/127/08 (Mod.Ord. 38/08) granted by the National Polytechnic Institute, the Science and Technology Institute of Federal District and the National Council of Science and Technology respectively.

REFERENCES

[1] J. Barbot, M. Djemai and T. Boukhobza, *Sliding Mode Observers in Sliding Mode Control in Engineering*, Springer-Verlag, 2002.
[2] T. Bentrcia, Y. Tamrabet and M. Chahdi, A low complexity finite-differences neural network combined approach for the modelling of p-n junction electrical properties, *Journal of Electron Devices*, vol.7, pp.203-210, 2010.
[3] M. Biey, M. Gilli and P. Checco, Complex dynamic phenomena in space-invariant cellular neural networks, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol.49, no.3, pp.340-345, 2002.
[4] I. Chairez, R. Fuentes, A. Poznyak, T. Poznyak, M. Escudero and L. Viana, Neural network identification of uncertain 2D partial differential equations, *The 6th International Conference on Electrical Engineering, Computing Science and Automatic Control, CCE*, 2009.
[5] N. E. Cotter, The stone-weierstrass theorem and its application to neural networks, *IEEE Transactions on Neural Networks*, vol.1, no.4, pp.290-295, 1990.
[6] J. Davila, L. Fridman and A. Levant, Second-order sliding-modes observer for mechanical systems, *IEEE Transactions on Automatic Control*, vol.50, no.11, pp.1785-1789, 2005.
[7] J. Davila, L. Fridman and A. Poznyak, Observation and identification of mechanical systems via second order sliding modes, *International Workshop on Variable Structure Systems*, vol.9, pp.232-327, 2006.
[8] A. Derdiyok, Z. Yan, M. Guven and V. Utkin, A sliding mode speed and rotor time constant observer for induction machines, *The 27th Annual Conference of the IEEE on Industrial Electronics Society*, vol.2, pp.1400-1405, 2001.
[9] B. Deylon, A. Juditsky and A. Benveniste, Accuracy analisys for wavelet approximations, *IEEE Transactions on Neural Networks*, vol.6, pp.332-348, 1995.
[10] J. Fei and F. Chowdhury, Robust adaptive sliding mode control for triaxial gyroscope, *International Journal of Innovative Computing, Information and Control*, vol.6, no.6, pp.2439-2448, 2010.
[11] A. Filippov, *Differential Equations with Discontinuous Right-Hand Side*, Kluwer, 1998.
[12] R. Fuentes, A. Poznyak, I. Chairez and T. Poznyak, Neural numerical modeling for uncertain distributed parameter systems, *International Joint Conference on Neural Networks*, pp.909-916, 2009.
[13] R. Fuentes, A. Poznyak, T. Poznyak and I. Chairez, Neural numerical modeling for uncertain distributed parameter systems, *Proc. of the International Joint Conference of Neural Networks*, 2009.
[14] S. Haykin, *Neural Networks – A Comprehensive Foundation*, IEEE Press, New York, 1994.

[15] A. Levant, Sliding order and sliding accuracy in sliding mode control, *International Journal of Control*, vol.58, no.6, pp.1247-1263, 1993.
[16] A. Levant, High order sliding modes, in *Sliding Mode Control in Engineering*, Marcel Dekker, Inc., 2002.
[17] A. Levant, Principles of 2-sliding mode designing, *Automatica*, vol.43, pp.823-830, 2005.
[18] A. Levant, Finite differences in homogeneous discontinuous control, *IEEE Transactions on Automatic Control*, vol.52, no.7, pp.1208-1217, 2007.
[19] A. Levant and L. Fridman, High order sliding modes, in *Sliding Mode in Control in Engineering*, Marcel Dekker, Inc., 2002.
[20] R. Leveque, Finite difference methods for ordinary and partial differential equations: Steady-state and time-dependent problems (classics in applied mathematics), *Society for Industrial and Applied Mathematics*, 2007.
[21] F. Lewis, A. Yesildirek and K. Liu, Multilayer neural-net robot controllerwith guaranteed tracking performance, *IEEE Transactions Neural Networks*, vol.7, pp.987-1000, 2001.
[22] L. Ljung, *System Identification: Theory for the User*, Prentice Hall, Englewood Cliff, NJ, 1987.
[23] J. Moreno and M. Osorio, Lyapunov approach to second-order sliding mode controllers and observers, *The 47th IEEE Conference on Decision and Control*, 2008.
[24] K. Narendra and S. Mukhopadhyay, Adaptive control using neural networks and approximation models, *IEEE Transactions on Neural Networks*, vol.8, no.3, pp.475-485, 1997.
[25] O. Nelles, *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*, Springer-Verlag, Berlin Heildeberg-New York, 2001.
[26] O. Omidvar and D. L. Elliott, *Neural Systems for Control*, Academic Press, 1997.
[27] V. Petridis and S. Petridis, Construction of neural network based lyapunov functions, *International Joint Conference on Neural Networks*, 2006.
[28] A. Poznyak, *Advanced Mathematical Tools for Automatic Control Engineers Volume 1: Deterministic Techniques*, Elsevier, 2008.
[29] A. Poznyak, E. Snchez and W. Yu, *Differential Neural Networks for Robust Nonlinear Control (Identification, State Estimation and Trajectory Tracking)*, World Scientific, 2001.
[30] T. Sato, Sliding mode control with proportional-integral compensation and application to an inverted pendulum system, *International Journal of Innovative Computing, Information and Control*, vol.6, no.2, pp.519-528, 2010.
[31] C. Tan and C. Edwards, An lmi approach for designing sliding mode observers, *International Journal of Control*, vol.74, no.16, pp.1559-1568, 2001.
[32] V. Utkin, *Sliding Modes in Control and Optimization*, Springer-Verlag, 1992.
[33] W. Xiang and F. Chen, Sliding mode control strategies for the hyperchaotic mck system, *ICIC Express Letters*, vol.3, no.3(A), pp.283-288, 2009.
[34] M. Zhang, Z. Yu, H. Huan and Y. Zhou, The sliding mode variable structure control based on composite reaching law of active magnetic bearing, *ICIC Express Letters*, vol.2, no.1, pp.59-63, 2008.

Appendix. Let us consider the identification error $[u_{1,i}(t) - \hat{u}_{1,i}(t), u_{2,i}(t) - \hat{u}_{2,i}(t)]$ that obeys the following dynamics

$$\begin{aligned}
\frac{d}{dt}\tilde{u}_{1,i}(t) &= \tilde{u}_{2,i}(t) - \beta_1 \lambda(\tilde{u}_{1,i}(t)) \operatorname{sign}(\tilde{u}_{1,i}(t)) \\
\frac{d}{dt}\hat{u}_{2,i}(t) &= f_{0,i}(x(t), t) - \bar{f}_{0,i}(x(t), t) + \tilde{f}_i(x,t) - \beta_2 \operatorname{sign}(u_{1,i}(t) - \hat{u}_{1,i}(t))
\end{aligned}$$

The convergence of the identification error will be based on a special Lyapunov function [23, 29]. This function is defined as:

$$V(t) = \sum_{i=1}^{N} V_i(t)$$

$$V_i(t) := \sum_{i=1}^{N} \|\zeta_i(t)\|_{P_i}^2 + \sum_{j=1}^{4} \left[k_{i,j} \operatorname{tr}\left\{ \tilde{W}_{i,j}^T(t)\, \tilde{W}_{i,j}(t) \right\} \right]$$

The vector $\zeta_i(t)$ is defined as follows

$$\zeta_i(t) := \begin{bmatrix} |\tilde{u}_{1,i}(t)|^{1/2}\operatorname{sign}(\tilde{u}_{1,i}(t)) \\ \tilde{u}_{1,i}(t) \\ \tilde{u}_{2,i}(t) \end{bmatrix}$$

Following the ideas given in [23], one can easily show that

$$\dot{\zeta}_i(t) := \begin{bmatrix} \frac{1}{2}|\tilde{u}_{1,i}(t)|^{-1/2}\frac{d}{dt}\tilde{u}_{1,i}(t) \\ \frac{d}{dt}\tilde{u}_{1,i}(t) \\ \frac{d}{dt}\tilde{u}_{2,i}(t) \end{bmatrix}$$

Therefore, the time derivative of each specific Lyapunov-like function has the following structure

$$\dot{V}_i(t) = 2\zeta_i^{\mathsf{T}}(t)P_i\dot{\zeta}_i(t) + \sum_{j=1}^{4}\left[k_{i,j}\operatorname{tr}\left\{\dot{W}_{i,j}^{T}(t)\tilde{W}_{i,j}(t)\right\}\right]$$

By direct substitution in the first term of the right hand side, one has

$$2\zeta_i^{\mathsf{T}}(t)P_i\dot{\zeta}_i(t) := \zeta_i^{\mathsf{T}}(t)P_i\begin{bmatrix} \frac{1}{2}|\tilde{u}_{1,i}(t)|^{-1/2}\left[\tilde{u}_{2,i}(t) - \beta_1\lambda(\tilde{u}_{1,i}(t))\operatorname{sign}(\tilde{u}_{1,i}(t))\right] \\ \tilde{u}_{2,i}(t) - \beta_1\lambda(\tilde{u}_{1,i}(t))\operatorname{sign}(\tilde{u}_{1,i}(t)) \\ G_i(t) - \beta_2\operatorname{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix}$$

where

$$\begin{aligned} G_i(t) := {} & A_i\tilde{\mu}(x_i,t) + \mathring{W}_{i,1}(x)\sigma(x_i)\tilde{\mu}(x_i,t) + \tilde{W}_{i,1}(x,t)\sigma(x_i)\hat{\mu}_i(x_i,t) \\ & + \mathring{W}_{i,2}(x)\varphi(x_i)\tilde{\mu}(x_{i-1},t) + \tilde{W}_{i,2}(x,t)\varphi(x_i)\hat{u}(x_{i-1},t) \\ & + \mathring{W}_{i,3}(x)\gamma(x_i)\tilde{\mu}(x_{i-2},t) + \tilde{W}_{i,3}(x,t)\gamma(x_i)\hat{u}(x_{i-2},t) \\ & + \mathring{W}_{i,4}(x)\eta(x_i)v(x_i,t) + \tilde{W}_{i,4}(x,t)\eta(x_i)v(x_i,t) + \tilde{f}_i(x,t) \end{aligned}$$

One can see that

$$\begin{aligned} & \zeta_i^{\mathsf{T}}(t)P_i\begin{bmatrix} \frac{1}{2}|\tilde{u}_{1,i}(t)|^{-1/2}\left[\tilde{u}_{2,i}(t) - \beta_{1,i}\lambda(\tilde{u}_{1,i}(t))\operatorname{sign}(\tilde{u}_{1,i}(t))\right] \\ \tilde{u}_{2,i}(t) - \beta_{1,i}\lambda(\tilde{u}_{1,i}(t))\operatorname{sign}(\tilde{u}_{1,i}(t)) \\ G_i(t) - \beta_{2,i}\operatorname{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix} \\ & = \frac{1}{2}|\tilde{u}_{1,i}(t)|^{-1/2}\zeta_i^{\mathsf{T}}(t)P_i\begin{bmatrix} -\beta_{1,i} & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} \lambda(\tilde{u}_{1,i}(t))\operatorname{sign}(\tilde{u}_{1,i}(t)) \\ \tilde{u}_{1,i}(t) \\ \tilde{u}_{2,i}(t) \end{bmatrix} \\ & + \zeta_i^{\mathsf{T}}(t)P_i\begin{bmatrix} 0 \\ \tilde{u}_{2,i}(t) - \beta_{1,i}(t)\lambda(\tilde{u}_{1,i}(t))\operatorname{sign}(\tilde{u}_{1,i}(t)) \\ G_i(t) - \beta_{2,i}(t)\operatorname{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix} \end{aligned} \tag{25}$$

Here, the following identity has been used

$$\zeta_i(t) := N\left(\zeta_i(t) + C^{\mathsf{T}}\tilde{u}_{1,i}(t)\right)$$

$$N = \left(I + C^{\mathsf{T}}C\right)^{-1}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

The last term in the previous equations can be rearranged as

$$\zeta_i^{\mathsf{T}}(t)P_i\begin{bmatrix} 0 \\ 0 \\ G_i(t) - \beta_2\operatorname{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix} := \zeta_i^{\mathsf{T}}(t)P_iM\left[G_i(t) - \beta_2\operatorname{sign}(\tilde{u}_{1,i}(t))\right]$$

To analyze this last term, lets use the matrix inequality [28]

$$X^{\intercal}Y + Y^{\intercal}X \le X^{\intercal}\Lambda X + Y^{\intercal}\Lambda Y$$

where

$$X, Y \in \Re^{n\times m}$$
$$\Lambda \in \Re^{n\times n}, \quad \Lambda^{\intercal} = \Lambda > 0$$

By direct application of this inequality, one has

$$\begin{aligned}
&\zeta_i^{\intercal}(t)\, P_i M \left[G_i(t) - \beta_2 \mathrm{sign}(\tilde{u}_{1,i}(t))\right] \\
&\le \zeta_i^{\intercal}(t)\, P_i M A_i \tilde{\mu}(x_i,t) + \tilde{\mu}^{\intercal}(x_i,t) A_i^{\intercal} M^{\intercal} P_i \zeta_i(t) \\
&\quad + \zeta_i^{\intercal}(t)\, P_i M \mathring{W}_{i,1}(x)\, \Lambda_1^{-1} \mathring{W}_{i,1}^{\intercal} M^{\intercal} P_i \zeta_i(t) + \tilde{\mu}^{\intercal}(x_i,t)\sigma^{\intercal}(x_i)\Lambda_1\sigma(x_i)\tilde{\mu}(x_i,t) \\
&\quad + \zeta_i^{\intercal}(t)\, P_i M \mathring{W}_{i,2}(x)\, \Lambda_2^{-1} \mathring{W}_{i,2}^{\intercal} M^{\intercal} P_i \zeta_i(t) + \tilde{\mu}^{\intercal}(x_{i-1},t)\varphi^{\intercal}(x_i)\Lambda_2\varphi(x_i)\tilde{\mu}(x_{i-1},t) \\
&\quad + \zeta_i^{\intercal}(t)\, P_i M \mathring{W}_{i,3}(x)\, \Lambda_3^{-1} \mathring{W}_{i,3}^{\intercal} M^{\intercal} P_i \zeta_i(t) + \tilde{\mu}^{\intercal}(x_{i-2},t)\gamma^{\intercal}(x_i)\Lambda_3\gamma(x_i)\tilde{\mu}(x_{i-2},t) \\
&\quad + \zeta_i^{\intercal}(t)\, P_i M \mathring{W}_{i,4}(x)\, \Lambda_4^{-1} \mathring{W}_{i,4}^{\intercal} M^{\intercal} P_i \zeta_i(t) + v^{\intercal}(x_i,t)\eta^{\intercal}(x_i)\Lambda_4\eta(x_i)v(x_i,t) \\
&\quad + \zeta_i^{\intercal}(t)\, P_i M \Lambda_f M^{\intercal} P_i \zeta_i(t) + \tilde{f}_i^{\intercal}(x,t)\Lambda_f^{-1}\tilde{f}_i(x,t)\zeta_i^{\intercal}(t)\, P_i M \left[\tilde{f}_i(x,t) - \beta_2 \mathrm{sign}(\tilde{u}_{1,i}(t))\right] \\
&\quad + \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,1}(x,t)\, \sigma(x_i)\hat{\mu}_i(x_i,t) + \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,2}(x,t)\, \varphi(x_i)\hat{u}(x_{i-1},t) \\
&\quad + \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,3}(x,t)\, \gamma(x_i)\hat{u}(x_{i-2},t) + \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,4}(x,t)\, \eta(x_i)v(x_i,t)
\end{aligned}$$

Using this information, one gets by the simple inclusion of term $\frac{1}{2}d\,|\tilde{u}_{1,i}(t)|^{-1/2}\,\zeta_i^{\intercal}(t)\,P_i N\tilde{u}_{1,i}(t)$

$$\begin{aligned}
2\zeta_i^{\intercal}(t)\, P_i \dot{\zeta}_i(t) \le{}& \frac{1}{2}|\tilde{u}_{1,i}(t)|^{-1/2}\,\zeta_i^{\intercal}(t) \left(P_i \begin{bmatrix} -\beta_1 & 0 & 1 \\ 0 & -d & 0 \\ -r & 0 & 0 \end{bmatrix} + \begin{bmatrix} -\beta_1 & 0 & 1 \\ 0 & -d & 0 \\ -r & 0 & 0 \end{bmatrix}^{\intercal} P_i \right) \zeta_i(t) \\
&+ \zeta_i^{\intercal}(t) \left[P^i M A_i + [M A_i]^{\intercal} P^i + P^i R^i P^i + Q_P^i\right] \zeta_i(t) \\
&+ \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,1}(x,t)\, \sigma(x_i)\hat{\mu}_i(x_i,t) + \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,2}(x,t)\, \varphi(x_i)\hat{u}(x_{i-1},t) \\
&+ \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,3}(x,t)\, \gamma(x_i)\hat{u}(x_{i-2},t) + \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,4}(x,t)\, \eta(x_i)v(x_i,t)
\end{aligned}$$

By the assumption on the existence of positive definite solutions of (23), this inequality is changed by

$$\begin{aligned}
\dot{V}_i(t) \le{}& -\frac{1}{2}|\tilde{u}_{1,i}(t)|^{-1/2}\,\zeta_i^{\intercal}(t)\, H_i \zeta_i(t) + \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,1}(x,t)\, \sigma(x_i)\hat{\mu}_i(x_i,t) \\
&+ \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,2}(x,t)\, \varphi(x_i)\hat{u}(x_{i-1},t) + \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,3}(x,t)\, \gamma(x_i)\hat{u}(x_{i-2},t) \\
&+ \zeta_i^{\intercal}(t)\, P_i M \tilde{W}_{i,4}(x,t)\, \eta(x_i)v(x_i,t) + \sum_{j=1}^{4}\sum_{i=1}^{N}\left[k_{i,j}\mathrm{tr}\left\{\dot{W}_{i,j}^{T}(t)\,\tilde{W}_{i,j}(t)\right\}\right]
\end{aligned}$$

where H_i satisfy the inequality given in (24). Finally, if using the adjustment laws given in (22) one has

$$\begin{aligned}
\dot{V}_i(t) &\le -\frac{1}{2}\alpha_{\min}\left\{P_i^{-1/2} H_i P_i^{-1/2}\right\} |\tilde{u}_{1,i}(t)|^{-1/2}\,\zeta_i^{\intercal}(t)\, P_i \zeta_i(t) \\
&\le -\frac{1}{2}\alpha_{\min}\left\{P_i^{-1/2} H_i P_i^{-1/2}\right\} \lambda_{\max}^{-1}\{P_i\}\sqrt{V_i(t)}
\end{aligned}$$

This last inequality can demonstrated following the procedure given in [23] and taking into account that $|\tilde{u}_{1,i}(t)|^{1/2} \le \|\zeta_i(t)\| \le \lambda_{\max}\{P_i\}\sqrt{V_t}$. Since the solution of the differential equation

$$\dot{v}_0(t) = -\alpha\sqrt{v_0\,(t)},\ \ v_0(0) > 0$$

is given by

$$v_0\,(t) = \left(\sqrt{v_0(0)} - \frac{\alpha}{2}t\right)^2$$

it follows from the comparison principle that $V_i\,(t) \leq v\,(t)$ whenever $V_i\,(0) \leq v\,(0)$. The proof is completed.