



**INSTITUTO POLITÉCNICO NACIONAL**  
**CENTRO DE INVESTIGACIÓN EN COMPUTACIÓN**

**Identificación No Paramétrica Por Redes Neuronales  
Diferenciales De Un Sistema Incierto Descrito En  
Ecuaciones Diferenciales Parciales Hiperbólicas, Utilizando  
Un Pseudo-Observador Por Modos Deslizantes De Alto  
Orden**

**TESIS**

Para obtener el grado de  
**Doctorado en Ciencias De La Computación**

**PRESENTA:**

**M. en C. Salvador Juárez López**

**DIRECTORES DE TESIS:**

Dr. Jorge Isaac Chairez Oria  
Dr. Oscar Camacho Nieto



MÉXICO D.F.

MAYO 2012





# INSTITUTO POLITÉCNICO NACIONAL SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

## ACTA DE REVISIÓN DE TESIS

En la Ciudad de México, D.F. siendo las 12:00 horas del día 5 del mes de septiembre de 2011 se reunieron los miembros de la Comisión Revisora de la Tesis, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación del:

**Centro de Investigación en Computación**

para examinar la tesis titulada:

**"IDENTIFICACIÓN NO PARAMÉTRICA POR REDES NEURONALES DIFERENCIALES DE UN SISTEMA INCIERTO DESCRITO EN ECUACIONES DIFERENCIALES PARCIALES HIPERBÓLICAS UTILIZANDO UN PSEUDO-OBSERVADOR POR MODOS DESLIZANTES DE ALTO ORDEN"**

Presentada por el alumno:

**JUÁREZ**

Apellido paterno

**LÓPEZ**

Apellido materno

**SALVADOR**

Nombre(s)

Con registro:

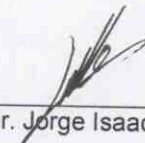
B	0	4	1	2	5	1
---	---	---	---	---	---	---


aspirante de: **DOCTORADO EN CIENCIAS DE LA COMPUTACIÓN**

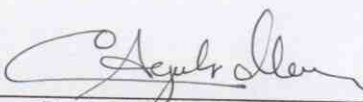
Después de intercambiar opiniones los miembros de la Comisión manifestaron **APROBAR LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.


### LA COMISIÓN REVISORA

Directores de tesis

  
Dr. Jorge Isaac Chairez Oria


  
Dr. Oscar Camacho Nieto

  
Dr. Carlos Aguilar Ibáñez

  
Dr. Cornejo Yáñez Márquez

  
Dr. Floriberto Ortiz Rodríguez

PRESIDENTE DEL COLEGIO DE PROFESORES

  
Dr. Luis Alfonso Villa Vargas



DIRECCION





# INSTITUTO POLITÉCNICO NACIONAL

## SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

### CARTA CESIÓN DE DERECHOS

En la Ciudad de México, D.F. el día 10 del mes de febrero del año 2012, el que suscribe **M. en C. Salvador Juárez López** alumno del programa de **Doctorado en Ciencias de la Computación**, con número de registro **B041251** adscrito al **Centro de Investigación en Computación**, manifiesta que es autor intelectual del presente trabajo de Tesis bajo la dirección de los doctores: **Dr. Oscar Camacho Nieto** y **el Dr. Jorge Isaac Chairez Oria** y cede los derechos del trabajo titulado: *Identificación No Paramétrica Por Redes Neuronales Diferenciales De Un Sistema Incierto Descrito En Ecuaciones Diferenciales Parciales Hiperbólicas Utilizando Un Pseudo-Observador Por Modos Deslizantes De Alto Orden*, al **Instituto Politécnico Nacional** para su difusión, con fines académicos y de investigación.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y directores del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección [salvadorjuarez@sagitario.cic.ipn.mx](mailto:salvadorjuarez@sagitario.cic.ipn.mx). Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

M. en C. SALVADOR JUÁREZ LÓPEZ



# **AGRADECIMIENTOS**

“EL AMOR Y EL DESEO SON LAS ALAS DEL ESPÍRITU DE LAS GRANDES HAZAÑAS”

J. WOLFGANG GOETHE.

## **A DIOS**

ANTE TODO, DEBO AGRADECERTE A TI MI PADRE JESÚS POR TODAS LAS  
BENDICIONES QUE ME HAS OTORGADO.

”BENDITO SEAS PADRE JESÚS DE NAZARET”.

## **A MIS PADRES**

GRACIAS POR SER MI EJEMPLO DE VIDA Y SER PARTE FUNDAMENTAL DE ESTE  
LOGRO.

## **A MIS GRANDES AMORES**

GRACIAS POR SER EL MOTOR QUE IMPULSA MIS SUEÑOS.

## **A MI FAMILIA**

### **AMIGOS Y HERMANOS**

GRACIAS POR TODO EL APOYO QUE ME HAN BRINDADO SIEMPRE. USTEDES  
SABEN QUE EN VERDAD ESTOY AGRADECIDO CON DIOS POR TENERLOS A MI  
LADO.





## **A MIS ASESORES**

Dr. Jorge Isaac Chairez Oria

Dr. Oscar Camacho Nieto

GRACIAS POR COMPARTIR SUS CONOCIMIENTOS Y MOSTRARME SUS GRANDES  
DOTES DE SERES HUMANOS.

## **A MI JURADO**

Dr. Carlos Fernando Aguilar Ibáñez

Dr. Cornelio Yáñez Márquez

Dr. Floriberto Ortiz Rodríguez

Dr. Amadeo José Argüelles Cruz

GRACIAS POR SU APORTACIÓN A ESTE TRABAJO, POR SUS COMENTARIOS Y LA  
RETROALIMENTACIÓN QUE REALIZARON CON EL PROFESIONALISMO QUE LOS  
CARACTERIZA Y ASÍ ENRIQUECERLO.

## **A CONACYT**

AL CONSEJO NACIONAL DE CIENCIA Y TECNOLOGÍA POR EL APOYO OTORGADO  
DURANTE MI FORMACIÓN.

## **AL CIC-IPN**

POR BRINDARME LOS APOYOS NECESARIOS PARA MI FORMACIÓN.



# Resumen

Existen un sin número de problemas de las ciencias en general y la ingeniería en particular, que pueden ser modelados matemáticamente mediante ecuaciones diferenciales parciales. El proceso del modelado de fenómenos físicos es en general una tarea compleja. Por otra parte, existen fuentes de incertidumbre que se suman a la tarea compleja de dichos modelados que son representaciones matemáticas. Las redes neuronales parecen ser una alternativa plausible para obtener una representación no paramétrica de los sistemas antes mencionados. Es conocido que las redes neuronales pueden aproximarse a un gran conjunto de funciones continuas definidas en un conjunto compacto con una precisión arbitraria. En este trabajo se propone una estrategia basada en redes neuronales diferenciales, para la identificación no paramétrica de un modelo matemático descrito mediante ecuaciones diferenciales parciales de tipo hiperbólico. El problema de la identificación se reduce a encontrar una expresión exacta de la dinámica de pesos utilizando las propiedades de redes neuronales diferenciales. Las leyes de adaptación para los pesos aseguran la convergencia de las trayectorias de las redes neuronales diferenciales para los estados de la ecuación diferencial parcial de tipo hiperbólico. Para investigar el comportamiento cualitativo de la metodología propuesta, se desarrollará un pseudo observador el cual tendrá como función la identificación de los parámetros de algunos sistemas presentados en este trabajo. Hay que mencionar que existen dos análisis diferentes, el primero tendrá ganancias constantes y el segundo se llevará a cabo utilizando ganancias variables, y en este último caso se visualizará claramente que existe una mejor identificación mucho más eficiente cuando se consideran las ganancias variables a través de las redes neuronales diferenciales.



# Abstract

There are many examples in science and engineering which may be described by a set of partial differential equations. The modeling process of such phenomena is in general, a complex task. Moreover, there exist some sources of uncertainties around that mathematical representation. 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 (DNN) 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 DNNs properties. The adaptive laws for weights ensure the convergence of the trajectories of the differential neural networks for the states to the partial differential equation of hyperbolic type. To investigate the qualitative behavior of the suggested methodology, develop a pseudo observer which function will be to identify the parameters of some systems presented in this paper. It should be mentioned that there are two different tests, the first will have steady income and the second was carried out using variable gains, and the latter is clearly displayed that there is a better identification much more efficient when considering the variable gains through differential neural networks.



# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	3
1.2. Planteamiento del problema . . . . .	4
1.3. Objetivos . . . . .	5
1.3.1. Principal . . . . .	5
1.3.2. Particulares . . . . .	5
1.4. Organización del trabajo . . . . .	6
<b>2. Antecedentes</b>	<b>7</b>
2.1. Ecuaciones Diferenciales Parciales . . . . .	7
2.1.1. Ecuaciones diferenciales parciales de tipo hiperbólico . . . . .	9
2.1.2. Espacio de Sobolev . . . . .	12
2.2. Métodos de Solución de EDP . . . . .	13
2.2.1. Método de Separación de Variables . . . . .	14
2.2.2. Métodos Numéricos de Solución de EDP . . . . .	15
2.2.3. Método de diferencias finitas y aproximación por mallas . . . . .	16
2.2.4. Método de elemento finito . . . . .	19
2.3. Redes Neuronales . . . . .	28
2.3.1. Memorias Hopfield . . . . .	29
2.3.2. Redes Neuronales Diferenciales . . . . .	33
2.3.3. Aplicación de la Técnica de RND en la aproximación de EDP . . . . .	34
2.4. Modos deslizantes . . . . .	34
2.4.1. Inclusiones diferenciales . . . . .	36
2.4.2. Modos deslizantes convencionales (MD) . . . . .	38
2.4.3. Modos deslizantes respecto a funciones con restricciones . . . . .	46





2.4.4. Un enfoque de control de MDAO a través de Lyapunov . . . . .	52
<b>3. Desarrollo del Trabajo</b>	<b>57</b>
3.1. Identificación por redes neuronales diferenciales hiperbólicas utilizando un pseudo observador de modos deslizantes. . . . .	57
3.1.1. Identificación de Ecuaciones Hiperbólicas . . . . .	57
3.1.2. Identificación con RND hiperbólica utilizando un pseudo observador de modos deslizantes . . . . .	59
3.1.3. Diferencias finitas y aproximación de RND para EDP inciertas . . . . .	60
3.1.4. Teoría general utilizada para aproximar funciones inciertas . . . . .	61
3.1.5. Aproximación con RND para EDPH . . . . .	62
3.1.6. Aproximación de RND basado en diferencias finitas . . . . .	63
3.1.7. Identificador no paramétrico . . . . .	64
3.1.8. Pseudo-observador adaptativo con la RND . . . . .	65
3.1.9. Estructura del Pseudo Observador . . . . .	68
3.1.10. Leyes de aprendizaje para identificar los pesos . . . . .	70
3.1.11. Estrategia basada en Lyapunov para probar la convergencia del identificador . . . . .	72
3.2. Identificación por redes neuronales diferenciales hiperbólicas utilizando un pseudo observador de modos deslizantes con ganancias variables . . . . .	72
3.2.1. Estrategia basada en Lyapunov para probar la convergencia del identificador . . . . .	74
<b>4. Resultados Numéricos</b>	<b>75</b>
4.1. Primer sistema . . . . .	75
4.2. Segundo Sistema . . . . .	78
4.3. Tercer sistema: Ganancias variables . . . . .	80
<b>5. Conclusiones</b>	<b>84</b>
<b>6. Apéndice</b>	<b>88</b>
<b>7. Trabajo a Futuro</b>	<b>92</b>



# Índice de Figuras

- Fig. 1 Cuerda Vibrante.
- Fig. 2 Problema con Valores en la Frontera (Ecuación de Onda).
- Fig. 3 Representación Diferencial Parcial.
- Fig. 4 Caja Gris.
- Fig. 5 Inclusiones Diferenciales.
- Fig. 6 Función Signo.
- Fig. 7 Modos Deslizantes-Twisting (Trayectoria de Fase).
- Fig. 8 Algoritmo Sub optimal (Trayectoria de Fase).
- Fig. 9 Algoritmo Super Twisting (Trayectoria de Fase).
- Fig. 10 Diagrama de la Ecuación Diferencial Parcial Hiperbólica Original 1.
- Fig. 11 Diagrama de la Ecuación Diferencial Parcial 1 Reconstruida Por Redes Neuronales.
- Fig. 12 Diagrama del Error de la Ecuación Diferencial Parcial 1.
- Fig. 13 Diagrama de la Ecuación Diferencial Parcial Hiperbólica Original 2.
- Fig. 14 Diagrama de la Ecuación Diferencial Parcial 2 Reconstruida Por Redes Neuronales.
- Fig. 15 Diagrama del Error de la Ecuación Diferencial Parcial 2.
- Fig. 16 Diagrama de la Ecuación Diferencial Parcial Hiperbólica Original con Ganancias Variables
- Fig. 17 Diagrama de la Ecuación Diferencial Parcial Hiperbólica Reconstruida por RND con G.V.
- Fig. 18 Diagrama del Error de la Ecuación Diferencial Parcial Hiperbólica con G.V.



# Lista de símbolos

$\in$	Pertenece o es elemento de
$A^\top$	Traspuesta de una matriz
$A^{-1}$	Inversa de una matriz
$tr A$	Traza de una matriz
$\det A$	Determinante de una matriz
$ x $	Valor absoluto de un número
$\ x\ $	Norma euclídeana de un vector, es decir $\ x\  = \sqrt{x^\top x} = \sqrt{\sum_{i=1}^n x_i^2}$
$\ x\ _A$	Norma ponderada de un vector, es decir $x^\top A x \leq 1$
$ $	Tal que
$\langle \cdot, \cdot \rangle$	Producto interno
$R(R^+)$	Conjunto de números reales no negativos
$R^{n \times m}$	Conjunto de matrices reales con n renglones y m columnas
LMI	Desigualdad Matricial Lineal
$\frac{\partial}{\partial t}$	Derivada parcial respecto al tiempo
$u_t$	Derivada parcial de $u$ respecto al tiempo
$C^1$	Espacio continuo en su primera derivada
$W$	Pesos de las redes neuronales



# Publicaciones

1. Non-Parametric Modelling of Uncertain Hyperbolic Partial Differential Equations Using Pseudo-High Order Sliding Mode Observers. *International Journal of Innovative Computing, Information and Control* (Volume 8, Number 3(A), March 2012).
2. Adaptive Gain Identifier for Uncertain Hyperbolic Partial Differential Equations Based on Differential Neural Networks and Second Order Sliding Mode Algorithm. *Neural Computing and Applications*, Springer (2011 Submitted).





# Capítulo 1

## Introducción

La teoría moderna de identificación a través de redes neuronales, se ocupa del problema de la extracción eficiente de señales y las propiedades dinámicas del sistema basándose en las mediciones de los datos disponibles. La identificación de parámetros para las diferentes clases de sistemas no lineales ha sido ampliamente estudiado en las últimas tres décadas[8][29]. Básicamente, son la clase de sistemas lineales y no lineales cuya dinámica depende linealmente de los parámetros desconocidos. Una característica general de estos trabajos es: que las mediciones exactas del vector de espacio de estados están disponibles. Las redes neuronales con su propiedad de aproximación universal y su capacidad de aprendizaje, han demostrado ser una poderosa herramienta para identificar y controlar los sistemas dinámicos complejos no lineales con parámetros o estructuras inciertas [28][30]. Explotar la capacidad natural de las redes neuronales artificiales (RNA) para aproximar funciones no lineales, permite la sustitución de las incertidumbres del sistema desconocido por modelos especiales adaptables. Estas aproximaciones numéricas de las RNA se definen por determinadas estructuras (discretas o continuas), que contienen una serie de parámetros desconocidos (pesos) que pueden ser ajustados.

Dependiendo del modelo de red neuronal, sus parámetros libres pueden ser modificados utilizando ecuaciones diferenciales. En la década de los noventa se propusieron diversos métodos que utilizan formas diferenciales para el diseño de las RNA. Estas construcciones se basaron en la estructura de Hopfield [10]. Hoy en día, se conocen como redes neuronales diferenciales (RND). El enfoque de las RND evita los problemas comunes de las redes neuronales convencionales como por ejemplo minimizar la búsqueda global. La mayoría de los métodos continuos de las RNA utilizan la teoría de control de Lyapunov para transformar

el problema de aproximación numérica a una retroalimentación no lineal adaptable. Cuando el modelo matemático de un proceso es incompleto o conocido parcialmente, el enfoque de las RND ofrece una herramienta eficaz para hacer frente a problemas en la teoría de control moderna como es: la identificación, la estimación de estados, la planeación de trayectorias, entre otros. En sistemas dinámicos con incertidumbres que tienen poca información y que están bajo los efectos de perturbaciones externas, se requiere poner atención especial en la construcción del identificador neuronal diferencial. En contraste, para muchos identificadores que requieren una descripción matemática detallada del sistema no lineal, las RNA's son suficientemente eficientes para hacer frente a una amplia clase de sistemas no lineales y que no tienen un modelo claramente definido. Para tener un buen desarrollo de cada identificador se requiere de una estructura específica y de leyes adaptativas para ajustar sus parámetros.

Las virtudes de las RND justifican su uso para aproximar por la derecha a la ecuación diferencial ordinaria incierta, la RND puede aproximarse a tales ecuaciones diferenciales con alta complejidad [9][5]. Ésta aproximación puede ser utilizada para resolver la identificación numérica de ecuaciones diferenciales parciales (EDP), y está basada en métodos numéricos como son: el método en diferencias finitas y el método de elementos finitos. Por lo tanto, la aplicación de un algoritmo que resuelve las ecuaciones diferenciales parciales en ecuaciones diferenciales ordinarias a través de las RND se justifica. Las consideraciones que justifican el uso de métodos numéricos para resolver las EDP en ecuaciones diferenciales ordinarias (EDO) son: 1) los datos de los problemas reales tienen siempre errores de medición, 2) el trabajo aritmético para la solución está limitado a un número finito de cifras significativas; resultado de errores de redondeo y 3) la evaluación numérica de las soluciones analíticas es a menudo, una tarea laboriosa y computacionalmente ineficiente, que requieren un gran número de iteraciones en el cálculo y el tratamiento de los datos, además de que en general, los métodos numéricos proporcionan soluciones numéricas adecuadas, y de una manera simple y eficiente.

Se ha mencionado que las RND son capaces de identificar los parámetros para resolver sistemas que no estén definidos claramente[11], estos sistemas se pueden resolver por medio de algunos métodos entre los que destaca el método de modos deslizantes (MD) que puede ser presentado sucintamente como la aplicación de una señal de control conmutando a alta frecuencia que consigue llevar el estado del sistema a una superficie denominada superficie de deslizamiento y una vez en ella mantenerlo ante posibles perturbaciones externas. Dicha superficie de deslizamiento es definida por el diseñador con el objeto de que el estado cumpla las especificaciones deseadas. La principal ventaja del control por MD es que aporta robustez

ante perturbaciones, tanto internas como externas. Los observadores neuronales diferenciales que se estudian en el enfoque de MD. Se utilizan para obtener el procedimiento algebraico de aprendizaje en línea para la identificación de la planta no lineal (diseño del modelo) totalmente a disposición de los estados. La característica esencial de la técnica de MD es la aplicación de las leyes discontinuas para lograr y mantener la dinámica de ciclo cerrado de una variedad determinada en el espacio de estados, para la conmutación de la superficie, con algunas características deseadas para el sistema. Este control ofrece muchas ventajas sobre otro tipo de identificación [29][19] y sobre las técnicas de control, obteniéndose un buen comportamiento transitorio. La necesidad de una cantidad reducida de información en comparación con las técnicas tradicionales de control, permite la capacidad de rechazo a perturbaciones, insensible a las variaciones de plantas como la no-linealidad, y una notable estabilidad y robustez en rendimiento.

El presente trabajo desarrolla un nuevo enfoque de identificación con redes neuronales diferenciales, utilizando sistemas de ecuaciones diferenciales parciales hiperbólicas, las cuales representan a diversos sistemas mecánicos que contienen oscilaciones, además, del control de modos deslizantes de alto orden (MDAO) que es eficiente para controlar plantas dinámicas complejas de orden dos o superiores, y que operan bajo condiciones de incertidumbre. Lo anterior, se justifica con la propiedad de las RND para ser modeladas por ecuaciones diferenciales complejas y que son capaces de resolver ecuaciones diferenciales parciales a través de un pseudo observador en la estructura de la red.

La principal aportación de este trabajo es la solución eficiente y eficaz de un sistema mecánico con la generación de un conjunto de herramientas (modelos, algoritmos, entre otros) que logran la convergencia de las siguientes áreas: modos deslizantes de alto orden, ecuaciones diferenciales parciales hiperbólicas, diferencias finitas y redes neuronales diferenciales.

## 1.1. Motivación

Entre los diversos métodos de control de sistemas no lineales que existen en la actualidad, sólo algunos son capaces de realizar algún tipo de control en este tipo de sistemas, entre ellos se encuentran el control por modos deslizantes de alto orden. Los modos deslizantes pueden aparecer en un sistema dinámico representado por ecuaciones diferenciales ordinarias. El término modo deslizante apareció primero en el contexto de los sistemas relevadores. Puede presentarse en el control como una función del interruptor de estado de un sistema en alta frecuencia; este movimiento es llamado modo deslizante. Formalmente, el modo deslizante

puede aparecer no solo en un sistema con control discontinuo sino en cualquier sistema dinámico con discontinuidades en las ecuaciones de movimiento. Sin embargo para resolver este tipo de sistemas, que se conforman de ecuaciones diferenciales ordinarias y parciales, se necesita que los parámetros se logren obtener, y esto se puede realizar a través de las redes neuronales. Dependiendo del modelo de red neuronal, se sugiere que sus parámetros libres puedan ser modificados por ecuaciones diferenciales. La mayoría de los métodos utilizan formas diferenciales para diseño de las redes neuronales artificiales. Estas construcciones se basan en la estructura de Hopfield, lo que actualmente se conocen como redes neuronales diferenciales. El enfoque de estas redes evita los problemas conocidos que son comunes en las redes neuronales convencionales.

Los métodos de redes neuronales artificiales se utilizan comúnmente para lograr un control utilizando la teoría de Lyapunov, que logre transformar el problema de aproximación numérica a través de una retroalimentación no lineal adaptable. Es por ello que se aprovechará este enfoque para lograr la identificación aplicando el control por modos deslizantes, mediante el uso de una función de Lyapunov que pueda ser útil para obtener un pseudo observador, que dependerá del sistema para que pueda ser estimado y a través de la estructura de la red neuronal implementarlo para que facilite la identificación de los parámetros desconocidos de los sistemas mostrados en este trabajo.

## 1.2. Planteamiento del problema

En la vida cotidiana existen demasiados ejemplos aplicados en la ciencia y la ingeniería que pueden ser descritos por un conjunto de ecuaciones en derivadas parciales. El proceso de modelar estos sistemas resultan ser un cálculo complejo.

Se hace énfasis en que existen fuentes de incertidumbre en torno a la representación matemática definida por ecuaciones diferenciales parciales, y es ahí donde las redes neuronales parecen ser una alternativa plausible para obtener una representación paramétrica de los sistemas descritos. Es bien sabido que las redes neuronales pueden aproximarse a un gran conjunto de funciones continuas definidas en un conjunto compacto con una precisión arbitraria. Se hace alusión que en la mayoría de los trabajos con redes neuronales diferenciales que se han llevado a cabo, utilizan ecuaciones diferenciales ordinarias, y no consideran las ecuaciones diferenciales parciales sin mencionar las ecuaciones de tipo hiperbólico.

En este trabajo se propone una estrategia basada en redes neuronales diferenciales para la identificación no paramétrica de un modelo matemático descrito por ecuaciones diferenciales

parciales de tipo hiperbólico que es incierto. El problema de la identificación se reduce a encontrar una expresión exacta de la dinámica de pesos utilizando las propiedades de redes neuronales diferenciales. Las leyes de adaptación para los pesos garantizará la convergencia de las trayectorias en las redes diferenciales para los estados de la ecuación diferencial parcial de tipo hiperbólico. Para investigar el comportamiento cualitativo de la metodología propuesta, como es el uso de un observador de modos deslizantes de alto orden, y directamente utilizar el algoritmo super twisting, el cual trabaja con sistemas de ecuaciones de orden 2, además, de utilizar una función de Lyapunov para generar un pseudo observador dentro de la estructura de la red neuronal diferencial, que resuelve el problema del modelado no paramétrico. En el presente trabajo se obtendrá la recuperación con éxito, de los parámetros de los diversos sistemas propuestos por medio de ecuaciones diferenciales parciales de tipo hiperbólico que contienen incertidumbres o perturbaciones, y con el apoyo de la red neuronal diferencial el cual incluye dentro de su estructura un pseudo observador aplicando el esquema del algoritmo super twisting, teniendo ganancias de tipo constante y variable.

## **1.3. Objetivos**

### **1.3.1. Principal**

Diseñar un identificador basado en redes neuronales diferenciales para ecuaciones diferenciales parciales de tipo hiperbólico, implementando un pseudo observador por modos deslizantes de orden superior.

### **1.3.2. Particulares**

1. Diseñar un algoritmo para desarrollar un identificador por redes neuronales diferenciales para ecuaciones diferenciales parciales.
2. Diseñar las leyes de aprendizaje de la propuesta de redes neuronales para ecuaciones diferenciales parciales.
3. Diseñar el esquema de pseudo-observador para la identificación de sistemas descritos en ecuaciones diferenciales parciales de tipo hiperbólico.
4. Desarrollar el algoritmo de identificación por redes neuronales diferenciales usando el pseudo-observador empleando el algoritmo super-twisting.

5. Determinar los valores de las ganancias del esquema de super twisting para el pseudo observador desarrollado.
6. Desarrollar el algoritmo de super twisting modificado con término lineal para el pseudo observador utilizando el esquema basado en redes neuronales diferenciales.
7. Determinar los valores de las ganancias que serán variables dentro del esquema de super twisting para el pseudo observador desarrollado.
8. Desarrollar el algoritmo de super twisting modificado con un término lineal para el pseudo observador, utilizando el esquema basado en redes neuronales diferenciales y ganancias variables.
9. Desarrollar los métodos numéricos para resolver el algoritmo de identificación basado en redes neuronales diferenciales, y los pseudo observadores propuestos para ecuaciones diferenciales parciales de tipo hiperbólico aplicando ganancias constantes y variables.

## 1.4. Organización del trabajo

Este trabajo se organiza de la siguiente manera, en el capítulo dos se mencionan las ecuaciones diferenciales ordinarias y parciales, así como los diversos métodos de solución para las ecuaciones diferenciales parciales. Se citan los espacios de Hilbert, Banach y Sobolev, a continuación se muestran las redes neuronales: estáticas y dinámicas, concluyendo con las redes neuronales diferenciales y para finalizar este capítulo se menciona el control por modos deslizantes convencionales así como los de alto orden como son: el twisting y el supertwisting. En el capítulo tres se realiza el desarrollo del trabajo de investigación, el cual inicia con la identificación de funciones hiperbólicas, la identificación con RND hiperbólicas utilizando un pseudo observador de modos deslizantes, el identificador no paramétrico, adaptación del pseudo observador con RND, la estructura del pseudo observador, leyes de aprendizaje para los pesos del identificador, identificación de RND hiperbólicas con ganancias constantes y variables. En el capítulo cuatro se muestran los resultados numéricos de los diversos sistemas utilizados en el presente trabajo a través del software matlab. En el capítulo cinco se muestran las conclusiones. En el capítulo seis se encuentra el apéndice, en el cual se demuestran los teoremas. En el capítulo 7 se presenta el trabajo a futuro.

# Capítulo 2

## Antecedentes

En este capítulo, se presenta primero un esquema de las ecuaciones diferenciales parciales, las ecuaciones diferenciales parciales de tipo hiperbólico, así como sus diferentes métodos de solución como son los analíticos y numéricos. Posteriormente se brindan conceptos básicos concernientes a las redes neuronales, empezando por las memorias asociativas Hopfield y las redes neuronales diferenciales. Al finalizar este capítulo se muestran los diversos tipos de control por modos deslizantes convencionales, twisting, super twisting y sub optimal, los cuales sirven para analizar sistemas no lineales que contienen perturbaciones denominadas chattering.

### 2.1. Ecuaciones Diferenciales Parciales

El vector compuesto por las variables de estados de una ecuación diferencial se encuentra en el espacio de Banach ( $B$ ). Si  $B$  es de dimensión finita ( $B \in \mathbb{R}^n, n < \infty$ ), el sistema se denomina sistema con parámetros concentrados y se menciona eso porque no es necesario considerar la distribución espacial de sus parámetros debido a que se puede considerar que están concentrados en un punto y se caracteriza por un sistema de ecuaciones diferenciales ordinarias. Si  $B$  es de dimensión infinita, el sistema se denomina con parámetros distribuidos pues en ellos es necesario tomar en cuenta la distribución espacial de sus parámetros y se caracteriza por sistemas de ecuaciones diferenciales parciales (EDP).

Una ecuación diferencial parcial o ecuación en diferencias, es una relación entre una función  $f$  de varias variables independientes  $x, y, z, t, \dots$  y las derivadas parciales de  $f$  respecto a esas variables. Las soluciones de EDP's están definidos naturalmente en espacios de Sobolev

$(W^{m,p})$ , donde las derivadas se entienden en un sentido débil para hacer el espacio completo, el de Banach [2].

Es así que un espacio de Sobolev es un espacio Banach o un espacio Hilbert de funciones con suficientes derivadas para algunas aplicaciones de dominio y equipado con una norma que mide tanto el tamaño como la suavidad de una función. La forma en la que las EDP's se presentan habitualmente, es a través de la modelización de fenómenos de la ciencia y tecnología, en los que se describe la dinámica a lo largo del tiempo de determinada variable (también a veces denominada estado) que puede representar objetos de lo más diversos, que van desde la posición de un satélite en el espacio hasta la dinámica de un átomo, pasando por los índices bursátiles o el grado en que una enfermedad afecta a la población. En otras palabras, los modelos dinámicos o de evolución reproducen nuestra propia concepción del mundo en un espacio tridimensional que evoluciona y cambia en el tiempo.

Cuando el estado o variable de un modelo o sistema de evolución es finito-dimensional, el modelo más natural es un sistema de EDO, cuya dimensión coincide precisamente con el número de parámetros necesarios para describir dicho estado. Así, por ejemplo, para posicionar una partícula en el espacio necesitamos de tres variables dependientes del tiempo, y para describir su dinámica un sistema de tres ecuaciones diferenciales. Pero en muchas ocasiones, como es el caso sistemáticamente en el contexto de la mecánica de medios continuos, la variable de estado es infinito-dimensional. Esto ocurre por ejemplo cuando se pretende describir la deformación de cuerpos elásticos o la temperatura de un cuerpo sólido, en los que la deformación o temperatura de cada uno de los puntos de ese medio continuo constituye, una variable o incógnita del sistema.

Las ecuaciones parabólicas y las hiperbólicas, representadas respectivamente por la ecuación del calor y la de ondas, son los modelos clásicos en el contexto de las EDP de evolución. Sus características matemáticas son distintas. Mientras que la ecuación del calor permite describir fenómenos altamente irreversibles en tiempo en los que la información se propaga a velocidad infinita, y la ecuación de ondas es el prototipo de modelo de propagación a velocidad finita completamente reversible en el tiempo. A continuación se describen las ecuaciones diferenciales parciales de tipo hiperbólico que son el tipo de ecuaciones que se tratarán en este trabajo.



### 2.1.1. Ecuaciones diferenciales parciales de tipo hiperbólico

Estas ecuaciones aparecen en cualquier análisis matemático de los fenómenos que comprenden la propagación de ondas en un medio continuo. Por ejemplo, los estudios de las ondas acústicas, ondas en el agua y ondas electromagnéticas, todos se basan en esta ecuación. Quizá la situación más fácil de visualizar en estas ecuaciones, ocurre en la investigación de las vibraciones mecánicas. Supóngase que una cuerda elástica de longitud  $l$  se tensa con firmeza entre dos soportes al mismo nivel horizontal, de modo que el eje  $x$  se quede a lo largo de la cuerda. Se tiene que la cuerda elástica es una cuerda de guitarra, un tirante o quizá un cable de energía eléctrica. Imagine que la cuerda se pone en movimiento de modo que vibre en un plano vertical y se denote por  $u(x, t)$  el desplazamiento vertical experimentado por la cuerda en el punto  $x$ , en el instante  $t$ . Si se desprecian los efectos de amortiguamiento, como la resistencia del aire, y si la amplitud del movimiento no es demasiado grande, entonces  $u(x, t)$  satisface la ecuación parcial

$$a^2 u_{xx} = u_{tt} \quad (2.1)$$

en el dominio  $0 < x < l, t > 0$ . La expresión (2.1) se conoce como ecuación de onda. El coeficiente constante  $a^2$  que aparece en (2.1) se define como

$$a^2 = \frac{T}{\rho} \quad (2.2)$$

donde  $T$  es la tensión en la cuerda y  $\rho$  es la masa por unidad de longitud del material de la misma. Por lo tanto, las unidades de  $a$  son m/s que es la velocidad (Fig.1).

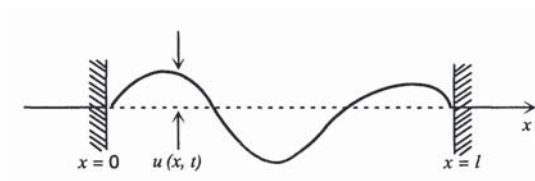


Fig. 1 Cuerda Vibrante

Para describir por completo el movimiento de la cuerda, es necesario especificar condiciones iniciales y en la frontera adecuadas para el desplazamiento  $u(x, t)$ . Se supone que los extremos permanecen fijos y, por lo tanto, las condiciones en la frontera son

$$u(0, t) = 0, u(l, t) = 0, t \geq 0 \quad (2.3)$$

En virtud en que la ecuación diferencial (2.1) es de segundo orden con respecto a  $t$ , es loable que sea necesario prescribir dos condiciones iniciales; éstas son la posición inicial de la cuerda y su velocidad inicial

$$u_t(x, 0) = g(x), 0 \leq x \leq l \quad (2.4)$$

donde  $f$  y  $g$  son funciones dadas. Para que las expresiones (2.3 y 2.4) sean coherentes también es necesario que

$$f(0) = f(l) = 0, g(0) = g(l) = 0 \quad (2.5)$$

Entonces el problema matemático es determinar la solución de la ecuación de onda (2.1) para que también satisfaga las condiciones en la frontera, y las condiciones iniciales. Este problema es con valor inicial en la variable de tiempo  $t$  y con valores en la frontera en la variable espacial  $x$ . Alternativamente, puede considerarse como un problema con valores en la frontera en la franja semifinita  $0 < x < l, t > 0$  del plano (ver Fig.2)

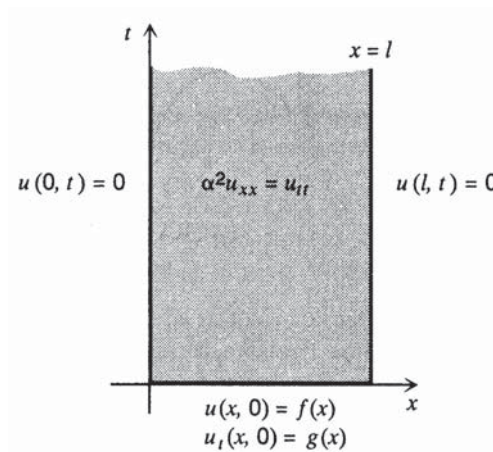


Fig. 2 Ecuación de Ondas

generalizando el término (2.1), se tiene

$$a^2(u_{xx} + u_{yy}) = u_{tt}$$

Esta ecuación surgirá al considerar el movimiento de una delgada hoja elástica, como el parche de un tambor. De manera semejante, la ecuación de onda en tres dimensiones es

$$a^2(u_{xx} + u_{yy} + u_{zz}) = u_{tt}$$

además, deben generalizarse de manera adecuada las condiciones en la frontera y condiciones iniciales.

Se ha proporcionado la ecuación de onda debido a que es un sistema natural que se define como ecuación diferencial hiperbólica. A continuación se mencionan las características y propiedades de las EDPH, que son parte importante en este trabajo.

Se muestran resultados importantes de las EDPH, como es el método aproximado de solución del problema de Cauchy, el cual se conoce como el método de diferencias finitas.

Para la ecuación

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \quad (2.6)$$

con las condiciones iniciales

$$\begin{aligned} u(0, x, y) &= \tau(x, y) \\ u'_t(0, x, y) &= \omega(x, y) \end{aligned} \quad (2.7)$$

supongase que las funciones iniciales  $u(0, x, y) = \tau(x, y)$  y  $u'_t(0, x, y) = \omega(x, y)$  tienen derivadas continuas hasta el cuarto orden y que están definidas en un cuadrado  $G$

$$a < x < b; c < y < d$$

Se trazan tres familias de planos paralelos en el espacio  $(t, x, y)$ :

$$\begin{aligned} t &= k\Delta, k = 0, 1, 2, \dots \\ x &= m\delta, y = n\delta \end{aligned}$$

aquí  $\delta$  y  $\Delta$  son ciertos números positivos. Los números  $m$  y  $n$  recorren ciertos valores enteros sucesivos tales que siempre:

$$a < m\delta < b; c < n\delta < d$$

para simplificar la exposición se supone:

$$a = m_1\delta, b = m_2\delta, c = n_1\delta, d = n_2\delta$$

sustituyendo la ecuación (2.6)  $u''_{tt}(k\Delta; m\delta, n\delta)$  por:

$$\frac{u[(k-1)\Delta, m\delta, n\delta] + u[(k+1)\Delta, m\delta, n\delta] - 2u(k\Delta; m\delta, n\delta)}{\Delta^2} \quad (2.8)$$

$u''_{xx}(k\Delta; m\delta, n\delta)$  por

$$\frac{u[k\Delta, (m+1)\delta, n\delta] + u[k\Delta, (m-1)\delta, n\delta] - 2u(k\Delta; m\delta, n\delta)}{\delta^2} \quad (2.9)$$

y  $u''_{yy}(k\Delta; m\delta, n\delta)$

$$\frac{u[k\Delta, m\delta, (n+1)\delta] + u[k\Delta, m\delta, (n-1)\delta] - 2u(k\Delta; m\delta, n\delta)}{\delta^2} \quad (2.10)$$

Es fácil comprobar que si  $u(t, x, y)$  tiene derivadas continuas hasta de segundo orden para  $y$  suficientemente pequeño, los errores que implica esa sustitución son pequeños. Después de la sustitución la ecuación diferencial (2.6) se convierte en una ecuación de diferencias que se denota por  $(k, m, n)$ . Dándole a  $(k, m, n)$  distintos valores admisibles, obteniendo un sistema de ecuaciones en diferencias. La solución de este sistema se designará por  $\bar{u}$ . De acuerdo con las condiciones iniciales se plantea:

$$\begin{aligned} \bar{u}(0, m\delta, n\delta) &= \tau(m\delta, n\delta) \\ \frac{\bar{u}(\Delta, m\delta, n\delta) - \bar{u}(0, m\delta, n\delta)}{\Delta} &= \omega(\bar{u}(0, m\delta, n\delta)) \end{aligned}$$

entonces las condiciones iniciales determinarán  $\bar{u}(0, m\delta, n\delta)$  y  $\bar{u}(\Delta, m\delta, n\delta)$ , en todos los puntos nodos para los cuales los correspondientes puntos  $(0, m\delta, n\delta)$  pertenecen a la región  $G$ .

Se debe tener en consideración que los sistemas que están conformados por ecuaciones diferenciales parciales de tipo hiperbólico, son similares a los sistemas que se analizan con modos deslizantes de alto orden, bajo el algoritmo de super twisting, es por ello, que en este trabajo se utilizarán los modos deslizantes con este tipo de ecuaciones.

### 2.1.2. Espacio de Sobolev

Es un espacio vectorial de funciones equipadas con una norma que es combinación de las normas  $L^p$  de la propia función [2], así como sus derivadas, hasta un orden determinado. Su importancia radica en el hecho de que las soluciones de ecuaciones diferenciales parciales están naturalmente en espacios de Sobolev más que en espacios clásicos de funciones continuas y con las derivadas entendidas en el sentido clásico.

Hay muchos criterios para la continuidad de funciones matemáticas. El criterio básico de continuidad es el relacionado con la diferenciabilidad (porque las funciones que son diferenciables son también continuas) y aún una mayor noción de continuidad es que si la función es diferenciable, también será continua (estas funciones se dice, son de la clase  $C^1$ ). Las funciones diferenciables son importantes en muchas esferas, y en particular para las ecuaciones diferenciales. Sin embargo, el espacio  $C^1$  (o  $C^2$ , etc) no es exactamente el espacio correcto

para estudiar las soluciones de las ecuaciones diferenciales. Los espacios de Sobolev son la sustitución moderna para estos espacios en los cuales se buscan soluciones de EDP.

El espacio de Sobolev  $W^{m,p}$  se define como el subconjunto de  $L^p$  de tal manera que la función  $f$  y sus derivadas débiles hasta de orden  $m$  tienen una norma finita  $L^p$ , para  $p \geq 1$  dada. Se debe tener cuidado al definir las derivadas en un sentido propio. En el caso unidimensional es suficiente suponer que la  $(m-1)$ -ésima derivada de la función  $f$ ,  $f^{(m-1)}$ , es diferenciable y es igual en casi todas partes para la integral de Lebesgue en sus derivadas.

Sea  $f \in C^m$ , la norma sobre  $f$  está definida por:

$$\|f\|_{m,p} := \left( \sum_{i=0}^m \|f^{(i)}\|^p \right)^{1/p} = \left( \sum_{i=0}^m \int |f^{(i)}(t)|^p dt \right)^{1/p}, \quad 1 \leq p < \infty \quad (2.11)$$

La completitud de  $u \in C^m: \|u\|_{m,p} < \infty$  con respecto a  $\|\cdot\|_{m,p}$  es el espacio de Sobolev  $W^{m,p}$ [35].

El espacio  $W^{m,p}$  equipado con la norma completa es un espacio de Banach.

Los espacios de Sobolev con  $p = 2$  son especialmente importantes por su conexión con las series de Fourier y porque ellos forman un espacio de Hilbert ( $H$ ). Una notación especial para cubrir este caso es:

$$H^m = W^{m,2} \quad (2.12)$$

Además el espacio  $H^m$  admite un producto interno, como el espacio  $H^0 = L^2$ . De hecho, el producto interno  $H^m$  se define en términos del producto interno  $L^2$ :

$$\langle u, v \rangle_{H^m} = \sum_{n=-\infty}^{\infty} \langle D^n u, D^n v \rangle_{L^2} \quad (2.13)$$

El espacio  $H^m$  se transforma en espacio de Hilbert con este producto interno [4].

La norma admitida en el espacio de Sobolev, es la utilizada para definir la estructura con la que se trabaja en la aproximación de sistemas con parámetros distribuidos descritos en EDP's, presentada en el presente trabajo.

## 2.2. Métodos de Solución de EDP

Una solución de una ecuación diferencial parcial es cualquier función que la verifica idénticamente. Las EDP's son una descripción idealizada de sistemas no lineales, utilizadas en

la representación matemática de fenómenos naturales que sacrifican la utilización de métodos convencionales de resolución analítica de dichas ecuaciones, en favor de la fiabilidad del modelo que representan. La desventaja es que no es fácil obtener sus soluciones exactas. Las soluciones de las EDP's pueden calcularse mediante métodos analíticos o aproximarse mediante métodos numéricos. Para ciertas clases de EDP's se conocen las soluciones analíticas, pero en general es necesario buscar soluciones aproximadas.

### 2.2.1. Método de Separación de Variables

En este método se supone que una solución se puede expresar como un producto de funciones desconocidas [35], cada una de las cuales depende de sólo una de las variables independientes. El éxito del método estriba en la posibilidad de escribir la ecuación resultante de modo que uno de sus miembros dependa solamente de una variable mientras que el otro miembro dependa de las demás variables, de tal manera que cada miembro sea una constante. Repitiendo el proceso se determinan las funciones desconocidas. La superposición de estas soluciones se puede utilizar para encontrar la solución verdadera.

El procedimiento que se sigue se presenta a continuación:

- Dada una ecuación diferencial parcial de una función de  $n$  variables, es útil advertir la solución de la forma:  $F = F_1(x_1)F_2(x_2)...F_n(x_n) = f_1(x_1) + f_2(x_2) + \dots + f_n(x_n)$  que lleva a la EDP a un conjunto de EDO's. Usualmente cada variable independiente crea una constante de separación que no puede ser determinada sólo de la ecuación. Cuando funciona bien esta técnica llamamos a la ecuación; ecuación diferencial parcial separable.
- Lo que se espera para una solución  $u(x; t) = v(x)w(t)$  o  $u(x; t) = v(x) + w(t)$  con las variables separables y sustituyendo en la EDP tal que  $v(x)w'(t) - v'(x)w(t) = 0$ . Así se espera que la solución  $u$  nunca sea cero (esto sucede en algunas ocasiones) y dividiendo por ella, se llega a:

$$\frac{w'(t)}{w(t)} = \frac{v'(x)}{v(x)} = -\lambda \text{ (alguna constante)} \quad (2.14)$$

Como una función de  $t$  al igual que una función de  $x$  debe ser constante para  $u(x; t) = v(x)w(t)$  lo mismo que para  $u(x; t) = v(x) + w(t)$ .

A pesar de la facilidad que representa el seguir el procedimiento anterior, éste método no es factible para resolver EDP's en un ordenador.

### 2.2.2. Métodos Numéricos de Solución de EDP

Esta técnica consiste en particionar el dominio espacial antes de construir el modelo matemático, y promediar espacialmente las variables descriptivas del sistema en cada uno de los subdominios finitos. Esto resulta en una concentración de los parámetros originalmente distribuidos en el espacio en nuevos parámetros que asumen valores uniformes en cada subdominio, y en la desaparición de la dependencia espacial de las variables descriptivas, las que ahora son evaluadas solamente en el escenario temporal.

Como consecuencia de todo esto, la dinámica del sistema resulta ahora descrita por EDO's en el tiempo. Si bien en muchos casos da buenos resultados, este método puede ser insostenible por la cantidad de EDO's que produce (tantas por subdominio como balances se hagan en él), o por no reflejar importantes propiedades del sistema, o simplemente porque a priori no se dispone de criterios físicos para realizar una partición adecuada. De los métodos de aproximación numérica disponibles para resolver ecuaciones diferenciales, los más utilizados son el método de diferencias finitas y el método de elementos finitos que suelen requerir de la definición de una malla (discretización del dominio), donde las funciones se aproximarán a nivel local. En esta sección se abordaran los métodos de: diferencias finitas y de elemento finito, pues son metodologías que se utilizan para aproximar numéricamente las soluciones de sistemas con parámetros distribuidos basada en RND. En la estructura de la red se utiliza el método de diferencias finitas, que es el principio básico manejado.

#### Método de Runge-Kutta

Un método relativamente sencillo y suficientemente exacto. El error local por truncamiento de éste método es proporcional a  $h^5$ . Por tanto, es dos ordenes de magnitud más exacto que otro como el mejorado de Euler y el de Taylor con tres términos, y tres ordenes de magnitud mejor que la fórmula simple de Euler. La fórmula de Runge-Kutta comprende un promedio ponderado de valores de  $f(t, y)$  tomados en diferentes puntos del intervalo  $t_n \leq t \leq t_{n+1}$ ; se expresa por

$$y_{n+1} = y_n + \frac{h}{6}(k_{n1} + 2k_{n2} + 2k_{n3} + k_{n4}) \quad (2.15)$$

en donde

$$\begin{aligned}
 k_{n_1} &= f(t_n, y_n) \\
 k_{n_2} &= f\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_{n_1}\right) \\
 k_{n_3} &= f\left(t_n + \frac{1}{2}h, y_n + \frac{1}{2}hk_{n_2}\right) \\
 k_{n_4} &= f(t_n + h, y_n + hk_{n_3})
 \end{aligned} \tag{2.16}$$

La suma  $\frac{(k_{n_1}+2k_{n_2}+2k_{n_3}+k_{n_4})}{6}$  puede interpretarse como una pendiente promedio. Notese que  $k_{n_1}$  es la pendiente en el extremo izquierdo del intervalo,  $k_{n_2}$  es la pendiente en el punto medio, si se aplica la fórmula de Euler para ir de  $t_n$  a  $t_n + \frac{h}{2}$ ,  $k_{n_3}$  es una segunda aproximación para la pendiente en el punto medio, y por último  $k_{n_4}$  es la pendiente en  $t_{n+h}$ , si se aplica la fórmula de Euler.

Aunque en principio no es difícil demostrar que la ecuación (2.15) difiere del desarrollo de Taylor de la solución (2.16) en términos que son proporcionales a  $h^5$ , el algebra es más bien extensa. Por tanto se acepta el hecho de que el error local por truncamiento al aplicar la ecuación (2.16), es proporcional a  $h^5$  y que, para un intervalo finito, el error global por truncamiento es cuando se tiene a lo más una constante multiplicada por  $h^4$ .

Notese que si  $f$  no depende de  $y$ , se tiene:

$$\begin{aligned}
 k_{n_1} &= f(t_n); k_{n_2} = k_{n_3} = f\left(t_n + \frac{h}{2}\right) \\
 k_{n_4} &= f(t_n + h)
 \end{aligned} \tag{2.17}$$

y la ecuación (2.15) se reduce a:

$$y_{n+1} - y_n = \frac{h}{6}[f(t_n) + 4f\left(t_n + \frac{h}{2}\right) + f(t_n + h)] \tag{2.18}$$

### 2.2.3. Método de diferencias finitas y aproximación por mallas

El Método consiste en una aproximación de derivadas parciales por expresiones algebraicas envolviendo los valores de la variable dependiente en un limitado número de puntos seleccionados. Como resultado de la aproximación, la ecuación diferencial parcial que describe el problema es reemplazada por un número finito de ecuaciones algebraicas, escritas en términos de los valores de la variable dependiente en puntos seleccionados.

Las ecuaciones son lineales si las ecuaciones diferenciales parciales son también lineales [8]. El valor de los puntos seleccionados se convierten en las incógnitas, en vez de la distribución



espacial continua de la variable dependiente. El sistema de ecuaciones algebraicas debe ser resuelto y puede envolver un número largo de operaciones aritméticas.

Si  $u$  es una función de  $x$  con derivadas finitas y continuas, entonces por el teorema de Taylor puede ser:

$$u(x+h, t) = u(x, t) + hu_x(x, t) + \frac{1}{2}h^2u_{xx}(x, t) + \frac{1}{6}h^3u_{xxx}(x, t) + l_1(x, t) \quad (2.19)$$

$$u(x-h, t) = u(x, t) - hu_x(x, t) + \frac{1}{2}h^2u_{xx}(x, t) - \frac{1}{6}h^3u_{xxx}(x, t) + l_2(x, t) \quad (2.20)$$

$l_1(x, t)$  y  $l_2(x, t)$  representan los términos que contienen el grado de  $h$  y/o mayores.

sumando las expansiones (2.19 y 2.20) se tiene:

$$u(x+h, t) - u(x-h, t) = 2u_x(x, t) + h^2u_{xx}(x, t) + l(h)^4 \quad (2.21)$$

donde  $l(h^4)$  denota el nuevo conjunto de términos que contienen el grado de  $h$  de orden 4 o superior. Asumiendo que estos términos son pequeños en relación al grado de  $h$ , se tiene lo siguiente:

$$u_{xx}(x, t)_{tconst} \simeq \frac{1}{h^2} \{u(x+h, t) - 2u(x, t) + u(x-h, t)\} \quad (2.22)$$

con un orden de error  $h^2$  sustrayendo la ecuación (2.20) de la ecuación (2.21) y despreciando los términos de  $h^3$  se obtiene:

$$u_x(x, t) \simeq \frac{1}{2h} \{u(x+h, t) - u(x-h, t)\} \quad (2.23)$$

con un orden de error  $h^2$ .

La ecuación (2.23) se aproxima a la pendiente de la tangente en el punto  $(x, u(x))$ . Estos resultados se obtuvieron con la suposición de que  $t$  se mantiene sin cambios, esta aproximación se define como aproximación diferencial central. Este método no es utilizable para la aproximación basada en RND.

También se puede aproximar la pendiente de la tangente en  $(x, t, u(x, t))_{t, const}$  por la pendiente de la recta que pasa por los puntos  $(x-h, u(x-h))$  y  $(x+h, u(x+h))$ , la obtención de la aproximación por diferencias hacia atrás se tiene como:

$$u_x(x, t)_{t, const} \simeq \frac{1}{h} \{u(x, t) - u(x-h, t)\} \quad (2.24)$$

o la pendiente de la recta que pasa por los puntos  $(x-h, t, u(x-h, t))$  y  $(x, t, u(x, t))$ , obteniendo la aproximación por diferencias hacia adelante:

$$u_x(x, t)_{t, const} \simeq \frac{1}{h} \{u(x-h, t) - u(x, t)\} \quad (2.25)$$

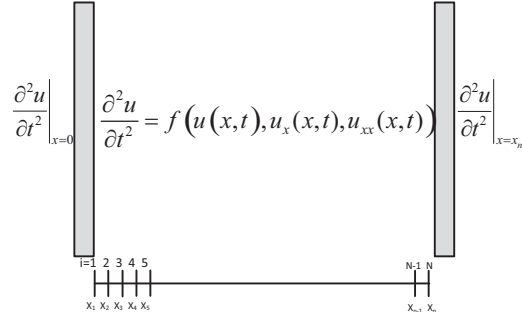


Fig.3 Representación Diferencial Parcial.

En este trabajo, la diferencia hacia atras será utilizada, porque la disponibilidad de medición  $u(x, t)$  se supone válida.

Por lo tanto, es necesario construir un conjunto denominado rejilla o malla que divide el subdominio  $x \in [0, 1]$  en las secciones  $N$  equidistantes (Fig. 3), se define como  $x_i$  de tal manera que  $x_0 = 0$  y  $x_N = 1$ . Usando esta descripción de malla, se puede utilizar la siguiente definición:

$$\begin{aligned}
 u_i(t) & : = u(x_i, t), u_{i,t}(x, t) = \frac{du(x, t)}{dt} \Big|_{x=x_i} \\
 u_{i,tt}(x, t) & = \frac{d^2u(x, t)}{dt^2} \Big|_{x=x_i} \\
 u_{i,x}(x, t) & = \frac{\partial u(x, t)}{\partial t} \Big|_{x=x_i}, u_{i,tt}(x, t) = \frac{\partial^2 u(x, t)}{\partial t^2} \Big|_{x=x_i}
 \end{aligned} \tag{2.26}$$

Utilizando la descripción y aplicación de malla y representación de diferencias finitas, se tiene:

$$u_{i,x}(x, t) \approx \frac{u(x_i, t) - u(x_{i-1}, t)}{\Delta x}, u_{i,xx}(x, t) \approx \frac{u_{i,x}(x, t) - u_{i-1}(x, t)}{\Delta x} \tag{2.27}$$

La aproximación basada en la malla de la EDP no lineal (2.27) puede ser representada de la siguiente manera:

$$u_{i,t}(x, t) = \dot{u}_i(t) \approx \Theta_i(u_i(t), u_{i-1}(t), u_{i-2}(t)) + \tilde{f}(u_i(t), u_{i-1}(t), u_{i-2}(t), v_i(x, t)) \tag{2.28}$$

donde:

$$\Theta_i(u_i(t), u_{i-1}(t), u_{i-2}(t)) := f_i \left( u(x_i, t), \frac{u(x_i, t) - u(x_{i-1}, t)}{\Delta x}, \frac{u_{i,x}(x, t) - u_{i-1,x}(x, t)}{\Delta x}, v_i(x, t) \right) \tag{2.29}$$

Evidentemente, este es un método numérico para aproximar la solución de la EDP. Sin embargo, esta aproximación puede ser formulada si  $f$  es perfectamente conocida. Este no será el caso en este trabajo.

#### 2.2.4. Método de elemento finito

El método de los elementos finitos (MEF) es un método numérico general para la aproximación de soluciones de ecuaciones diferenciales parciales muy utilizado en diversos problemas de ingeniería y física [34]. El MEF está pensado para ser usado en computadoras y permite resolver ecuaciones diferenciales asociadas a un problema físico sobre geometrías complicadas, además se utiliza en el diseño y mejora de productos y aplicaciones industriales, así como en la simulación de sistemas físicos y biológicos complejos. La variedad de problemas a los que puede aplicarse ha crecido enormemente, siendo el requisito básico que las ecuaciones constitutivas y ecuaciones de evolución temporal del problema a considerar sean conocidas de antemano.

El MEF permite obtener una solución numérica aproximada sobre un cuerpo, estructura o dominio (medio continuo) sobre el que están definidas ciertas ecuaciones diferenciales en forma débil o integral que caracterizan el comportamiento físico del problema, dividiéndolo en un número elevado de subdominios no-intersectantes entre sí denominados elementos finitos. El conjunto de elementos finitos forma una partición del dominio también denominada discretización. Dentro de cada elemento se distinguen una serie de puntos representativos llamados "nodos". Dos nodos son adyacentes si pertenecen al mismo elemento finito; además, un nodo sobre la frontera de un elemento finito puede pertenecer a varios elementos. El conjunto de nodos considerando sus relaciones de adyacencia se llama malla.

Los cálculos se realizan sobre una malla de puntos (llamados nodos), que sirven a su vez de base para discretización del dominio en elementos finitos. La generación de la malla se realiza usualmente con programas especiales llamados generadores de mallas, en una etapa previa a los cálculos que se denomina pre-proceso. De acuerdo con estas relaciones de adyacencia o conectividad se relaciona el valor de un conjunto de variables incógnitas definidas en cada nodo y denominadas grados de libertad. El conjunto de relaciones entre el valor de una determinada variable entre los nodos se puede escribir en forma de sistema de ecuaciones lineales (o linealizadas). La matriz de dicho sistema de ecuaciones se llama matriz de rigidez del sistema. El número de ecuaciones de dicho sistema es proporcional al número de nodos.

Típicamente el método de los elementos finitos se programa computacionalmente para

calcular el campo de desplazamientos y, posteriormente, a través de relaciones cinemáticas y constitutivas las deformaciones y tensiones respectivamente, cuando se trata de un problema de mecánica de sólidos deformables o más generalmente un problema de mecánica de medios continuos. El método de los elementos finitos es muy usado debido a su generalidad y a la facilidad de introducir dominios de cálculo complejos (en dos o tres dimensiones). Además el método es fácilmente adaptable a problemas de transmisión de calor, de mecánica de fluidos para calcular campos de velocidades y presiones o de campo electromagnético. Dada la imposibilidad práctica de encontrar la solución analítica de estos problemas [13], con frecuencia en la práctica ingenieril los métodos numéricos y, en particular, los elementos finitos, se convierten en la única alternativa práctica de cálculo.

Una importante propiedad del método es la convergencia; si se consideran particiones de elementos finitos sucesivamente más finas, la solución numérica calculada converge rápidamente hacia la solución exacta del sistema de ecuaciones.

El MEF, si bien en una dimensión espacial y en problemas con coeficientes constantes como los que se utilizan en este trabajo, debido a que proporciona esquemas numéricos de aproximación muy semejantes a los que se han comentado mediante diferencias finitas, conceptualmente es distinto, el MEF consiste en buscar soluciones en un espacio de dimensión finita. No se adopta el punto "de vista nodal" sino el del método de Galerkin: habiendo elegido una base de funciones y definido subespacios de dimensión finita, se elige en él la función que es solución de la EDP en el sentido más preciso posible. Para ello comprobamos si el producto escalar del operador diferencial con todas las funciones de dicho espacio es nulo (para lo cual es preciso adoptar una formulación variacional del problema que se obtiene integrando por partes). De este modo obtenemos una proyección de la solución real de la EDP sobre el espacio finito-dimensional considerado.

Esta metodología puede también ser aplicada en el marco de las ecuaciones de evolución y, en particular, en la ecuación del calor analizada en esta sección. Se tiene la función:

$$v_t = f(v_{xx}, v_{yy}, v_x, v_{xy}, v) \quad (2.30)$$

que es igual a:

$$v_t = \lambda^T \vec{v} \text{ en } \Omega \quad (2.31)$$

$$\vec{v} = \begin{bmatrix} v_{xx} \\ v_{yy} \\ v_x \\ v_{xy} \\ v \end{bmatrix}$$

$$\int \Omega w(v_t - \lambda^T \vec{v}) d\Omega - \int \Gamma w \frac{dv}{d\Omega} d\Gamma = 0$$

$$- \int y_1 y_2 \int x_1 x_2 (\lambda, v_{xx}) dx dy \text{ con } v \text{ suave}$$

$$ds_1 = v_{xx} dx \tag{2.32}$$

$$s_1 = \int_{x_1}^{x_2} v_x dx = v_x \Big|_{x_1}^{x_2}$$

$$\int_{y_1}^{y_2} \int_{x_1}^{x_2} w \lambda_1, v_{xx}, dx dy = \int_{y_1}^{y_2} w \lambda_1, v_x \Big|_{x_1}^{x_2} dy - \int_{x_1}^{x_2} \int_{y_1}^{y_2} w_x \lambda_1 v_x dx dy$$

$$\int_{y_1}^{y_2} \int_{x_1}^{x_2} w \lambda_2, v_{yy}, dx dy = \int_{y_1}^{y_2} w \lambda_2, v_y \Big|_{x_1}^{x_2} dy - \int_{x_1}^{x_2} \int_{y_1}^{y_2} w_y \lambda_2 v_x dx dy$$

sumandolos se obtiene:

$$\int_{y_1}^{y_2} w \lambda_1 v_x \Big|_{x_1}^{x_2} dy + \int_{x_1}^{x_2} w \lambda_2 v_y \Big|_{y_1}^{y_2} dx - \int \Omega (w_x \lambda_1 v_x + w_y \lambda_2 v_y) d\Omega$$

$$\int_{\Gamma_1} w \lambda_1 v_x \Big| n_x d\Gamma + \int_{\Gamma_2} w \lambda_2 v_y \Big| n_y d\Gamma - \int \Omega (w_x \lambda_1 v_x + w_y \lambda_2 v_y) d\Omega + \oint w (v_x n_x + \lambda v_y n_y) d\Gamma$$

$$\text{donde } (v_x n_x + \lambda v_y n_y) = \frac{d\Delta V}{dn} \tag{2.33}$$

y se necesita obtener:

$$V = \sum_{i=1}^n H_i(x, y) u_i(t) \tag{2.34}$$

## Método de Galerkin

El Método de Galerkin es una técnica que proporciona el marco general para la aproximación de algunos problemas variacionales en la práctica, se convierte en un algoritmo especial de discretización, que mediante la definición de una combinación lineal sobre un espacio de dimensión finita, se le denomina MEF. Su aplicación es muy importante, por ser muy usual en la resolución numérica de problemas de valores de frontera que con frecuencia se presentan en modelaciones fenomenológicas de la ingeniería. En el desarrollo de la formulación variacional, del problema diferencial, se introduce a medida que se requiere, la teoría funcional básica, de un modo que el lector tenga la facilidad de comprender los objetivos del trabajo. Luego se centrará en la utilización del método de elementos finitos para la búsqueda de una solución del problema variacional, el cual se formula a partir de un problema de contorno de tipo Dirichlet, así también en la comparación computacional, de la aplicación

del método con una base de funciones propuesta en el trabajo, para la generación del espacio de elementos finitos aproximados y la base usual en el espacio de igual dimensión.

Dado un conjunto abierto  $\Omega = (a, b) \subset \mathbb{R}$ , donde  $a, b$  son números reales sobre el que se ha planteado el siguiente problema diferencial:

$$-\frac{d}{dx} \left( p(x) \frac{du}{dx} \right) + q(x)u = f(x) \quad (2.35)$$

La variable desconocida  $y$  satisface las condiciones frontera:

$$u(a) = u_a, u(b) = u_b$$

Al que se le denomina también problema de contorno clásico. Se asumen las siguientes afirmaciones:

- Sean  $p, q$  y  $f$  funciones continuas sobre  $\Omega$ ,  $R, V \subset L^2(\Omega)$  es un espacio de Hilbert, de dimensión infinita y se denota por  $V'$  al espacio dual de  $V$ , así también denotemos por  $a(\cdot, \cdot) : V \times V \rightarrow R$  una forma bilineal de  $V$  y  $l \in V'$ , una forma lineal.
- Si  $f \in L^2(\Omega)$ , haciendo homogéneas a las condiciones de frontera, es decir  $y_a = y_b = 0$  y multiplicando por una función  $v$  del espacio de funciones infinitamente diferenciables y de soporte compacto, entonces la ecuación (2.35) se convierte en:

$$-\frac{d}{dx} \left( p(x) \frac{du}{dx} \right) v + q(x)uv = f(x)v, \forall u \in v \quad (2.36)$$

Aplicando una integración por partes en esta ecuación (2.36), y teniendo en cuenta las condiciones de frontera; resulta la siguiente ecuación integral:

$$\int_{\Omega} [p(x)u'(x)v'(x) + q(x)u(x)v(x)] dx = \int_{\Omega} f(x)v(x)dx \quad (2.37)$$

Si se asume que:

$$\begin{aligned} a(u, v) &= \int_{\Omega} [p(x)u'(x)v'(x) + q(x)u(x)v(x)] dx \\ l(v) &= \int_{\Omega} f(x)v(x)dx \end{aligned} \quad (2.38)$$

donde se puede ver  $V$  es un espacio de Hilbert,  $a(u, v)$  es una forma bilineal acotada y  $V$ -elíptica (o coerciva) de  $V$ , y  $l(v)$  una forma lineal acotada de  $V$ . Entonces el problema clásico

se reduce al siguiente problema denominado “problema variacional”, es decir: Hallar  $u \in V$ , que satisfaga la ecuación variacional

$$a(u, v) = l(v), \forall v \in V \quad (2.39)$$

donde  $a(\cdot, \cdot)$  es una forma bilineal acotada y  $V$ -elíptica; es decir:

$$\begin{aligned} \exists M \in \mathbb{R}, |a(u, v)| &\leq M \|u\|_V \|v\|_V, \forall u, v \in V \\ \exists C_0 \in \mathbb{R}, a(v, v) &\geq C_0 \|v\|_V^2, \forall v \in V \end{aligned}$$

Por el lema de Lax-Milgram [27] se tiene que el problema variacional dada en la ecuación (2.39) tiene solución única.

En general, es imposible encontrar una función de forma explícita que represente la solución exacta del problema ecuación (2.39), puesto que el espacio  $V$  es de dimensión infinita. Por tanto se plantea una forma de hallar una solución aproximada, es decir resolver un problema equivalente a la ecuación (2.39) pero en un espacio de dimensión finita. Para ello se considera el siguiente método de Galerkin, el cual dice que es posible suponer que la combinación lineal definida por  $u_N$ , donde:

$$u_N = \sum_{i=1}^N \xi_i \phi_i \quad (2.40)$$

es solución aproximada del problema de la ecuación (2.39), así también que el conjunto de funciones  $\phi_j, j = 1, \dots, N$  son linealmente independientes y genera el espacio de dimensión finita  $V_N$ , es decir,  $u_N \in V_N$ , es una sucesión de funciones que satisface la ecuación variacional:

$$a(u_N, v) = l(v), \forall v \in V_N \quad (2.41)$$

Por el principio del álgebra fundamental, se puede encontrar un subespacio de dimensión finita  $V_N \subseteq V$ . Entonces el problema (ecuación 2.39) se proyecta sobre  $V_N$  y extendiendo las propiedades mencionadas anteriormente para la forma bilineal  $a(\cdot, \cdot)$  y para la forma lineal  $l \in V'$ , se puede extender también el Lema de Lax-Milgram al problema discreto. Por tanto se puede predecir que existe solución única, que satisface la ecuación variacional dada en la expresión (2.39). Ahora para hallar esta solución, se asume que la combinación lineal presentada en la ecuación (2.40), esta determinada por el conjunto de funciones  $\{\phi_i\}_{i=1}^N$  la cual es una base de  $V_N$  y que  $\{\xi_i\}_{i=1}^N$ , es el conjunto de componentes que definen la solución

aproximada en cada nodo, ambos términos son desconocidos hasta ahora. Al reemplazar la ecuación (2.40) en la ecuación (2.41), se obtiene:

$$a\left(\sum_{i=1}^N \xi_i \varphi_i, v\right) = l(v), \forall v \in V_N \quad (2.42)$$

Según el método de Galerkin si se toma  $\varphi_j$  de la misma base de  $V_N$ , se genera un problema discreto equivalente a la ecuación (2.39). Es decir, hallar el conjunto de funciones  $\{\xi_i\}_{i=1}^N$ , tal que satisface el siguiente sistema de ecuaciones:

$$\sum_{i=1}^N \xi_i a(\phi_i, \phi_j) = l(\phi_j); j = 1, \dots, n \quad (2.43)$$

Significa que mediante el método de Galerkin se convierte la ecuación variacional dada en (2.39) en la forma de un sistema lineal de ecuaciones algebraicas (ecuación 2.43) que podemos expresarlo en forma matricial como:

$$A\xi = b$$

donde  $\xi = (\xi_i), i = 1, \dots, N$  es el vector desconocido y la matriz  $A = (a(\phi_i, \phi_j)) \in R^{N \times N}$ , cuyos coeficientes resultan de operar la forma bilineal sobre las funciones de base y en cada elemento de la malla. En la teoría de elasticidad la matriz  $A$  recibe el nombre de matriz de rigidez y es el vector de cargas y tiene la forma:

$$b = (l(\phi_j)) \in R^N \quad (2.44)$$

$$\begin{bmatrix} a(\phi_i, \phi_j) & \cdots & a(\phi_i, \phi_j) \\ \vdots & \ddots & \vdots \\ a(\phi_i, \phi_j) & \cdots & a(\phi_i, \phi_j) \end{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_N \end{bmatrix} = \begin{bmatrix} l(\phi_j) \\ \vdots \\ l(\phi_j) \end{bmatrix}$$

Luego al resolver este sistema lineal de ecuaciones algebraicas se encuentra la solución aproximada de la ecuación (2.36).

La solución  $u_N$ , es una aproximación de la solución exacta  $u$ . Con la finalidad de precisar mejor la solución, es lógico buscar una mayor aproximación, una alternativa es elegir  $N$  muy grande, lo que equivale a hacer un refinamiento sobre  $\Omega$ . Es decir, buscar una sucesión de subespacios  $V_{N_1} \subseteq V_{N_2} \subseteq \dots \subseteq V_{N_n}$ , y obtener la correspondiente sucesión de soluciones aproximadas  $u_{N_i}$ ,  $i = 1, 2, \dots, n$ . Pero con este procedimiento también se incrementará el tamaño de la matriz de coeficientes del sistema, lo que conduce a un cálculo más costoso



computacionalmente en el sistema. Otra alternativa es elegir óptimamente las funciones de base para el subespacio  $V_N$ , de dimensión finita. Este espacio es construido por medio de polinomios a trozos. Generalmente con el método de Galerkin, se utilizan los polinomios de interpolación lineal de tipo Lagrange, lo que lleva a la obtención de subespacios lineales, asociados a dos nodos por elemento en cada elemento finito de la malla definida sobre  $\Omega$ . A este esquema se le conoce como método de elementos finitos de Galerkin lineal. Numéricamente el método de elementos finitos es el más eficiente para resolver problemas de valores de frontera, la ventaja es que permite elegir diferentes bases de aproximación, bajo la condición de que sus elementos deben satisfacer las propiedades de ser linealmente independientes y generar el espacio solución. Anteriormente vimos que el problema de valores de frontera se reduce a resolver un sistema lineal de ecuaciones algebraicas, cuya matriz de coeficientes es la matriz  $A$ , la cual es simétrica y definida positiva. Los coeficientes de  $A$  juegan un rol esencial, tal es así que, si el número de condición de  $A$ , es demasiado grande, entonces es imposible encontrar directamente una solución directa del sistema, la alternativa sería utilizar un método iterativo, lo que es otro inconveniente sumado a la dispersión de la matriz  $A$ . Se dice que es dispersa, si la mayoría de sus entradas son ceros, en caso contrario, se dice que la matriz es densa.

La propiedad de que la matriz sea dispersa, hace mas costoso el cálculo computacional del sistema, si se puede observar el cálculo de cada entrada de la matriz, involucra un dominio de integración y a veces puede resultar que esta es no acotada. Otra forma de mejorar este inconveniente es, si previamente se realiza un tratamiento de la matriz, mediante un método de almacenamiento de tipo perfil o almacenar los coeficientes no nulos en un tipo vector. Generalmente si la matriz es muy dispersa el sistema es resuelto de manera más eficiente utilizando un método iterativo.

Para obtener una matriz de rigidez óptima, tenemos que ser muy cuidadosos con la elección del espacio de sus funciones de la base, para que pueda generar una matriz de tipo banda con un ancho óptimo. Un espacio aproximador de elementos finitos  $V_h$ , donde  $h$  representa el espaciado de cada elemento, es óptimo si se optimiza el grado de aproximación de las funciones de su base.

## **Elemento finito en la ecuación de onda**

No es difícil de adaptar el método de elementos finitos la ecuación de onda. Para lograrlo, conviene en primer lugar dar una formulación variacional de la ecuación de onda:

$$\begin{cases} u_{tt} - u_{xx} = 0, & 0 < x < \pi, t > 0 \\ u(0, t) = u(\pi, t) = 0, & t > 0 \\ u(x, 0) = \phi(x), u_t(x, 0) = \psi(x), & 0 < x < \pi \end{cases}$$

En este caso, el espacio de energía natural es:

$$u \in C([0, \infty); H_0^1(0, \pi)) \cap C^1([0, \infty); L^2(0, \pi)) \quad (2.45)$$

y la formulación variacional correspondiente es:

$$\frac{d^2}{dt^2} \int_0^\pi u(x, t) \phi(x) dx + \int_0^\pi u_x(x, t) \phi_x(x) dx = 0, \forall \phi \in H_0^1(0, \pi) \quad (2.46)$$

Conviene interpretar la ecuación (2.45) con cautela, puesto que la regularidad indicada en la expresión (2.46) no garantiza que el primer término obtenga un sentido clásico puesto que de la ecuación (2.45) no se deduce que la función  $\int_0^\pi u(x, t) \phi(x) dx$  sea de clase  $C^2$  con respecto al tiempo. Sin embargo lo es, puesto que las soluciones de la ecuación de ondas con la propiedad de regularidad la ecuación (2.45) verifican también que:

$$u \in C^2([0, \infty); H^{-1}(0, \pi)) \quad (2.47)$$

donde  $H^{-1}(0, \pi)$  es el espacio de Sobolev dual de  $H_0^1(0, \pi)$ . Esto permite interpretar la ecuación (2.46) como una ecuación diferencial clásica que se verifica para cada tiempo  $t \geq 0$  y cada función es  $\phi \in H_0^1(0, \pi)$ .

Obviamente las ecuaciones (2.45 y 2.46) han de ser completadas con las condiciones iniciales:

$$u(x, 0) = \varphi(x), u_t(x, 0) = \psi(x), 0 < x < \pi \quad (2.48)$$

De esta formulación débil de la ecuación de ondas es fácil obtener las ecuaciones de la aproximación por elementos finitos. Para ello se introduce el espacio  $V_h$  de funciones lineales a trozos y continuas de  $H_0^1(0, \pi)$ , de dimensión  $M$ , asociado al mallado de paso  $h$ .

La formulación variacional más natural es hallar  $u_h \in C^2([0, \infty); V_h)$  tal que:

$$\frac{d^2}{dt^2} \int_0^\pi u_h(x, t) \phi_j(x) dx + \int_0^\pi u_{h,x}(x, t) \phi_{j,x}(x) dx = 0, \forall t > 0, \forall j = 1, \dots, M \quad (2.49)$$

junto con las condiciones iniciales:

$$u_h(x, 0) = \phi_h(x), u_h'(x, 0) = \psi_h(x), 0 < x < \pi \quad (2.50)$$

Aquí y en lo sucesivo,  $\phi_j = \phi_j(x)$  denota la  $j$ -ésima función de base del espacio  $V_h$ , con centro en el punto  $x_j = j_h$  del mallado. Las funciones  $\varphi_h$  y  $\psi_h$  son por otra parte aproximaciones del dato inicial continuo  $(\varphi, \psi)$  en  $V_h$ .

Utilizando la notación vectorial habitual y las matrices de masa y rigidez, es fácil reescribir el problema como un sistema de  $M$  ecuaciones diferenciales lineales de orden 2, acopladas, con  $M$  incógnitas:

$$\begin{cases} M_h'' \vec{u}(t) + R_h \vec{u}(t) = 0, t > 0 \\ \vec{u}(0) = \vec{\phi}_h, \vec{u}'(0) = \vec{\psi}_h \end{cases}$$

El método desarrollado anteriormente, basado en las descomposiciones espectrales, permiten comprobar que se trata de un método convergente de orden 2.

Como es habitual en el método de elementos finitos, en una dimensión espacial, proporciona un esquema discreto que también puede interpretarse como una variación del método semidiscreto en diferencias finitas:

$$\begin{cases} u_j'' + \frac{[2u_j - u_{j+1} - u_{j-1}]}{h^2} & j = 1, \dots, M, t > 0 \\ u_0 = u_{M+1} = 0 & t > 0 \\ u_j(0) = \phi_j, u_j'(0) = \psi_j & j = 1, \dots, M \end{cases} \quad (2.51)$$

En efecto, la ecuación (2.51) puede reescribirse como:

$$h \left[ \frac{2}{3} u_j''(t) + \frac{1}{6} u_{j+1}''(t) + \frac{1}{6} u_{j-1}''(t) \right] = \frac{1}{h} [2u_j(t) - u_{j+1}(t) - u_{j-1}(t)] \quad (2.52)$$

Pero, la ventaja del método de elementos finitos reside precisamente no en este hecho, sino en su versatilidad para adaptarse a situaciones más complejas, donde el método de diferencias finitas es de difícil utilización (problemas no-lineales, geometrías complejas de varias variables espaciales, coeficientes variables, etc.).

La identidad de energía es también fácil de obtener. De la ecuación (2.51) y teniendo en cuenta que tanto la matriz de masa como de rigidez, son simétricas y definidas positivas, la energía correspondiente viene dada por:

$$E_h(t) = \frac{1}{2} \langle \mathcal{M}_h' \vec{u}(t), \vec{u}'(t) \rangle + \frac{1}{2} \langle \mathcal{R}_h \vec{u}_h(t), \vec{u}_h(t) \rangle \quad (2.53)$$

donde  $\langle \cdot, \cdot \rangle$  denota el producto escalar euclídeo.

La energía puede también escribirse componente a componente y se obtendrá:

$$E_h(t) = \frac{h^2}{2} \sum_{j=1}^M \left[ \frac{2}{3} |u_j'(t)|^2 + \frac{1}{3} u_j' + 1(t) u_j'(t) + \left| \frac{u_{j+1}(t) - u_j(t)}{h} \right|^2 \right] \quad (2.54)$$

Cuya conservación es también fácil deducirla de la escritura del sistema componente a componente como en la ecuación (2.52). En efecto, multiplicando en la expresión (2.52) por  $u'_j(t)$  y sumando en  $j$  no es difícil verificar que la energía se conserva. En la expresión (2.54) de la energía y con el objeto de comprobar su carácter definido-positivo es conveniente observar que los dos primeros términos de la suma pueden agruparse dando lugar al cuadrado perfecto:

$$\frac{1}{3} |u'_{j+1}(t) + u'_j(t)|^2$$

Una vez que se ha obtenido el esquema de aproximación semi-discreto mostrada en la expresión (2.51) y probando su convergencia, no es difícil introducir esquemas completamente discretos basados en el método de elementos finitos. Bastaría por ejemplo el utilizar diferencias finitas centradas en la discretización de la segunda derivada temporal del término dado en la ecuación (2.51).

## 2.3. Redes Neuronales

En 1949 Hebb propuso una ley de aprendizaje que puede considerarse como el punto de partida de los algoritmos de entrenamiento de las redes neuronales, mientras que la primera estructura de red neuronal fue propuesta por Frank Rosenblatt en 1958. Desde entonces el interés suscitado por las redes neuronales ha crecido enormemente, proponiéndose nuevos modelos que han venido a superar las limitaciones de las redes anteriores, dándoles mayores capacidades.

Las redes neuronales tienen su origen en la observación de los sistemas neuronales de los seres vivos, sin embargo, éstas no son un fiel reflejo de dichas redes biológicas. Como características más relevantes de las redes neuronales artificiales se citan las siguientes:

1. La capacidad de aprendizaje. La naturaleza no lineal de las mismas les permite exhibir comportamientos verdaderamente complejos. Esta característica junto con algoritmos que extraen un mayor rendimiento de tales sistemas les confieren esa capacidad.
2. La posibilidad de procesamiento paralelo. Aspecto que es aprovechado en la implementación y le lleva a aumentar considerablemente la velocidad de procesamiento.
3. La significativa tolerancia a fallos, dado que un fallo en unas pocas conexiones locales, rara vez contribuye significativamente al funcionamiento global de la red. Esta

propiedad las hace convenientes en ciertas aplicaciones, donde, una avería puede constituir un serio peligro.

Estas características, entre otras, hacen de las redes neuronales una herramienta fundamental en varios campos del conocimiento. Las redes neuronales generalmente han sido implementadas mediante la programación sobre computadoras digitales, sin embargo ya en nuestros días empiezan a proliferar implementaciones en hardware, ya sea mediante montajes electrónicos o incorporando dispositivos ópticos. Tales ingenios, presentan un paso importante hacia la consecución de mayores velocidades de procesamiento, así como un camino hacia la popularización de las mismas, siendo ésta un área de investigación muy activa.

En estos días la ingeniería de control se enfrenta a sistemas cada vez más complejos, a requisitos de diseño más exigentes y al desconocimiento de las plantas y sus entornos, revitalizando este hecho la investigación en el campo de las redes neuronales. En cuanto a la clasificación de las redes neuronales se puede establecer una división de las mismas en función de su arquitectura y forma de procesar las señales en tres clases. En general, existen las siguientes clases de RNA:

1. Las redes neuronales estáticas, que utilizan algoritmos de optimización local, como el método de “back-propagation”, para ajustar los parámetros de la red con la finalidad de aproximar una función no lineal desconocida.

2. Las redes neuronales recurrentes, que emplean ecuaciones en diferencias para resolver el problema de aproximación de funciones [1].

3. Las redes neuronales diferenciales, que explotan las propiedades de los sistemas retroalimentados, permitiendo evitar la mayoría de los problemas relacionados con la búsqueda de un extremo global, mediante el cambio de la ley de aprendizaje como un algoritmo estático a un modelo diferencial de ajuste de los pesos de la RNA [1].

El enfoque de Redes Neuronales Diferenciales que es presentado en éste trabajo, aprovecha la información de las propiedades aplicadas a la red neuronal recurrente, y si el modelo matemático de un proceso se considera incompleto o parcialmente conocido, ofrece una alternativa eficaz para atacar el problema de la identificación no paramétrica y otros, [17]. El problema que se discute en este estudio es la aplicación de una solución numérica basada en RND para resolver el problema de identificación no paramétrica de sistemas no lineales inciertos descritos en EDP.

A continuación se mencionan las memorias Hopfield, las cuales son la base de las redes neuronales utilizadas en este trabajo.

### 2.3.1. Memorias Hopfield

En 1982 las redes neuronales y memoria asociativas nace como consecuencia del artículo publicado por el físico John J. Hopfield (Anderson & Rosenfeld, Eds. 1990). En esa publicación, Hopfield presenta su memoria asociativa, que impactó positivamente a la comunidad científica. En ese trabajo se parte de un sistema físico descrito por un vector de estado  $x$  cuyas coordenadas son  $(x_1, x_2, \dots, x_n)$ ; se considera que el sistema tiene puntos límites localmente estables  $x^a, x^b, \dots$ . Entonces, si el sistema es activado en un estado suficientemente cercano a cualquier punto límite localmente estable, digamos en  $x = x^a + \Delta$ , al transcurrir el tiempo, el estado del sistema cambiará hasta llegar a  $x \approx x^a$ . El punto de arranque  $x^a + \Delta$  representa un conocimiento parcial del estado estable  $x^a$ , y a partir de ahí el sistema genera la información correcta  $x^a$ . Hopfield declara categóricamente que un sistema físico que se comporta de esta manera y que además, sea susceptible de que cualquier conjunto predeterminado de estados, pueda ser asignado como el conjunto de estados localmente estables, es útil como memoria asociativa.

En el modelo que originalmente propuso Hopfield, cada neurona tiene dos posibles estados:  $x_i = 0$  y  $x_i = 1$ ; sin embargo en el mismo artículo presentado por Hopfield, hace la relevante observación que la capacidad de almacenamiento de información en la memoria puede incrementarse por un factor de 2, si se escogen como posibles estados de las neuronas los valores  $x_i = -1$  y  $x_i = 1$  en lugar de los valores originales.

La intensidad de la fuerza de conexión de la neurona  $x_i$  a la neurona  $x_j$  se representa por el valor de  $m_{ij}$ , y se considera que hay simetría, es decir,  $m_{ij} = m_{ji}$ . Si  $x_i$  no está conectada con  $x_j$ , por lo tanto  $m_{ij} = 0$ ; en particular, no hay conexiones recurrentes de una neurona a si misma, lo cual significa que  $m_{ii} = 0, \forall i$ . El estado instantáneo del sistema está completamente especificado por el vector columna de dimensión  $n$ , cuyas coordenadas son los valores de las  $n$  neuronas.

La memoria es autoasociativa, simétrica, con ceros en la diagonal principal. En virtud de que la memoria es autoasociativa, el conjunto fundamental para la memoria Hopfield es  $\{(x^\mu, x^\mu) \mid \mu = 1, 2, \dots, p\}$  con

$$\begin{aligned} x^\mu &= \begin{pmatrix} x_1^\mu \\ x_2^\mu \\ \vdots \\ x_n^\mu \end{pmatrix} \in A^n \\ A &= \{-1, 1\} \end{aligned} \tag{2.55}$$

## Fase de aprendizaje

Esta fase de aprendizaje, es decir la creación de la memoria  $M$ , se realiza de la siguiente manera:

$$m_{ij} = \begin{cases} \sum_{\mu=1}^p x_i^\mu x_j^\mu & \text{si } i \neq j \\ 0 & \text{si } i = j \end{cases} \quad (2.56)$$

esta operación se obtiene en tres etapas:

Primera etapa. Para cada una de las  $p$  asociaciones  $(x^\mu, x^\mu)$ , se encuentra la matriz  $x^\mu \cdot (x^\mu)^T$  de dimensión  $n \times n$ :

$$\begin{aligned} x^\mu \cdot (x^\mu)^T &= \begin{pmatrix} x_1^\mu \\ x_2^\mu \\ \vdots \\ x_n^\mu \end{pmatrix} \cdot \begin{pmatrix} x_1^\mu & x_2^\mu & \cdots & x_n^\mu \end{pmatrix} \\ &= \begin{pmatrix} x_1^\mu x_1^\mu & x_1^\mu x_2^\mu & \cdots & x_1^\mu x_i^\mu & \cdots & x_1^\mu x_n^\mu \\ x_2^\mu x_1^\mu & x_2^\mu x_2^\mu & \cdots & x_2^\mu x_i^\mu & \cdots & x_2^\mu x_n^\mu \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_i^\mu x_1^\mu & x_i^\mu x_2^\mu & \cdots & x_i^\mu x_i^\mu & \cdots & x_i^\mu x_n^\mu \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_n^\mu x_1^\mu & x_n^\mu x_2^\mu & \cdots & x_n^\mu x_i^\mu & \cdots & x_n^\mu x_n^\mu \end{pmatrix} \end{aligned} \quad (2.57)$$

la matriz (2.57) debe tener valores 1 en su diagonal principal, porque los componentes de los patrones pertenecen al conjunto  $A = \{-1, 1\}$ , según la expresión 2.56, y para cualquiera de los valores, 1 o  $-1$ , se cumple que

$$\begin{aligned} x_i^\mu x_i^\mu &= 1, \forall i = 1, \dots, n : x^\mu \cdot (x^\mu)^T \\ &= \begin{pmatrix} 1 & x_1^\mu x_2^\mu & \cdots & x_1^\mu x_i^\mu & \cdots & x_1^\mu x_n^\mu \\ x_2^\mu x_1^\mu & 1 & \cdots & x_2^\mu x_i^\mu & \cdots & x_2^\mu x_n^\mu \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_i^\mu x_1^\mu & x_i^\mu x_2^\mu & \cdots & 1 & \cdots & x_i^\mu x_n^\mu \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ x_n^\mu x_1^\mu & x_n^\mu x_2^\mu & \cdots & x_n^\mu x_i^\mu & \cdots & 1 \end{pmatrix} \end{aligned} \quad (2.58)$$

Segunda etapa. A cada una de las  $p$  matrices  $x^\mu \cdot (x^\mu)^T$ , expresada en 2.58 se resta la

matriz identidad  $I$  de dimensión  $n \times n$ , con la finalidad de hacer ceros la diagonal principal:

$$= \begin{pmatrix} & & & [x^\mu \cdot (x^\mu)^T] - I & & \\ & 0 & x_1^\mu x_2^\mu & \cdots & x_1^\mu x_i^\mu & \cdots & x_1^\mu x_n^\mu \\ x_2^\mu x_1^\mu & & 0 & \cdots & x_2^\mu x_i^\mu & \cdots & x_2^\mu x_n^\mu \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \\ x_i^\mu x_1^\mu & x_i^\mu x_2^\mu & \cdots & 0 & \cdots & x_i^\mu x_n^\mu \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \\ x_n^\mu x_1^\mu & x_n^\mu x_2^\mu & \cdots & x_n^\mu x_i^\mu & \cdots & 0 & \end{pmatrix} \quad (2.59)$$

Tercera etapa. Se suman las  $p$  matrices  $x^\mu \cdot (x^\mu)^T - I$  para que finalmente se obtenga la memoria asociativa Hopfield  $M$ :

$$M = \sum_{\mu=1}^p [x^\mu \cdot (x^\mu)^T - I] = [m_{ij}]_{n \times n}$$

### Fase de recuperación

Esta recuperación consiste en tres etapas, se debe tener presente que un patrón  $\tilde{x}$  a la memoria autoasociativa Hopfield cambiará su estado con el tiempo, de modo que cada neurona  $x_i$ , ajuste su valor de acuerdo con el resultado que arroje la comparación de la cantidad  $\sum_{j=1}^n m_{ij}x_j$  con un valor umbral, el cual normalmente se coloca en cero. Se representa el estado de la memoria de Hopfield en el tiempo  $t$  por  $x(t)$ , entonces  $x_i(t)$  representa el valor de la neurona  $x_i$  en el tiempo  $t$  y  $x_i(t+1)$  el valor de  $x_i$  en el tiempo siguiente ( $t+1$ ). Dado un vector columna de entrada  $\tilde{x}$ .

Primera etapa. Para  $t=0$ , se hace  $x(t) = \tilde{x}$ ; es decir,  $x_i(0) = \tilde{x}_i, \forall i \in \{1, 2, 3, \dots, n\}$

Segunda etapa.  $\forall i \in \{1, 2, 3, \dots, n\}$  se calcula  $x_i(t+1)$  en conformidad con la siguiente condición:

$$x_i(t+1) = \begin{cases} 1 & \text{si } \sum_{j=1}^n m_{ij}x_j(t) > 0 \\ x_i(t) & \text{si } \sum_{j=1}^n m_{ij}x_j = 0 \\ -1 & \text{si } \sum_{j=1}^n m_{ij}x_j < 0 \end{cases} \quad (2.60)$$

Tercera etapa. Se compara  $x_i(t+1)$  con  $x_i(t) \forall i \in \{1, 2, 3, \dots, n\}$ . Si  $x(t+1) = x(t)$  el proceso termina y el vector recuperado es  $x(0) = \tilde{x}$ . De otra manera, se realizan tantas



iteraciones como sea necesario hasta llegar a un valor  $t = \tau$  para el cual  $x_i(\tau + 1) = x_i(\tau) \forall i \in \{1, 2, 3, \dots, n\}$ , el proceso termina y el patrón recuperado es  $x(\tau)$ .

Esta etapa indica que el sistema llega a un punto límite localmente estable en el tiempo  $\tau$ . La existencia de  $\tau$ , esta garantizada a través de la demostración que hace Hopfield, esto es debido a que existen puntos límite localmente estables en su modelo de memoria asociativa; para ello, se define  $E$  de la siguiente manera, tomando en cuenta la condición de que  $m_{ij} = 0, \forall i :$

$$E = -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n m_{ij} x_i x_j$$

### 2.3.2. Redes Neuronales Diferenciales

Los métodos de aproximación analítica a la solución de una EDPH, proporcionan frecuentemente información útil acerca del comportamiento de la solución en valores críticos de la variable dependiente, pero tienden a ser más difíciles de aplicar que los métodos numéricos. Entre las consideraciones que justifican el uso de métodos numéricos, para solucionar ciertos tipos de ecuaciones diferenciales ordinarias y en derivadas parciales se encuentran las siguientes:

- Los datos de los problemas reales presentan siempre errores de medición, y el trabajo aritmético para la solución está limitado a un número finito de cifras significativas que resultan en errores de redondeo. Por lo tanto, incluso los métodos analíticos proporcionan resultados que son aproximaciones numéricas.
- La evaluación numérica de las soluciones analíticas es a menudo una tarea laboriosa y computacionalmente ineficiente, pues requieren de un gran número de iteraciones en el cálculo y el tratamiento de los datos en serie, mientras que los métodos numéricos generalmente proporcionan soluciones numéricas adecuadas, de manera simple y eficiente.

Otro inconveniente es que para alcanzar una solución numérica, estos métodos están bien definidos sólo si la estructura de la EDP es perfectamente conocida. Es así que, si el modelo matemático de los procesos considerados está incompleto o es parcialmente conocido, la metodología de redes neuronales es la herramienta sugerida [1], en éste trabajo, para resolver este problema, se considera al sistema como una caja gris (ver Fig. 4) donde el conocimiento

del mismo es limitado y sólo se dispone de la entrada y/o la salida del sistema.

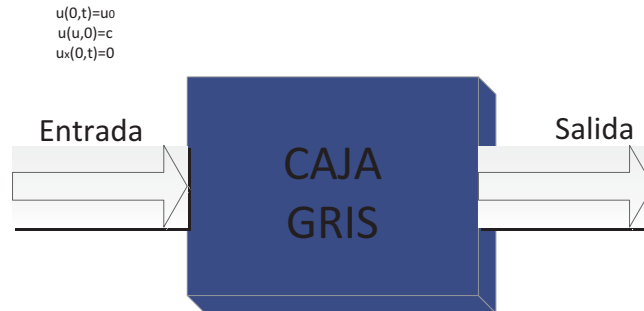


Fig. 4 Caja Gris

A continuación, se presentará una estrategia basada en redes neuronales diferenciales para resolver el problema de identificación no paramétrica de sistemas no lineales inciertos (caja gris) descritos en EDP.

### 2.3.3. Aplicación de la Técnica de RND en la aproximación de EDP

Por el buen comportamiento de las redes neuronales recurrentes dinámicas [19] como aproximadores universales de funciones complejas no lineales, es posible la representación y modelado de un sistema no lineal utilizando RNA. La representación real de muchos sistemas es obtenida mediante ecuaciones diferenciales de alto orden, de la misma manera, una red neuronal recurrente dinámica, puede ser modelada por medio de ecuaciones diferenciales de cierta complejidad, esta propiedad es la que las diferencia de otras y lo que les da el nombre de redes neuronales diferenciales.

Las redes neuronales han sido empleadas en la estimación de estos sistemas desconocidos debido a su masivo paralelismo, rápida adaptación y capacidad de aprendizaje con bastante éxito. Las RNA han demostrado buenos desempeños al identificar estados no lineales afectados por algunas incertidumbres o perturbaciones externas.

El enfoque de RND que es presentado en éste trabajo, aprovecha la información de las propiedades aplicadas a la RND [19], y si el modelo matemático de un proceso se considera incompleto o parcialmente conocido, ofrece una alternativa eficaz para atacar el problema de la identificación no paramétrica y otros [33]. El problema que se discute en este estudio es la aplicación de una solución numérica basada en RND para resolver el problema

de identificación no paramétrica de sistemas no lineales inciertos descritos en EDP. En la siguiente sección se hablará del método de control a través de modos deslizantes, el cual es un tema clave para lograr el propósito de este trabajo, debido a que se necesitará el uso de un observador por la teoría de modos deslizantes para conformar este trabajo.

## 2.4. Modos deslizantes

El control por MD desde hace mucho tiempo ha demostrado su interés en sistemas no lineales. Entre ellas, por la relativa simplicidad de diseño, control de movimiento independiente (siempre y cuando se mantengan las condiciones de deslizamiento), invarianza a la dinámica del proceso de las características y las perturbaciones externas, gran variedad de modos de operación tales como la regulación, control de trayectoria y la observación. Aunque el tema ya ha sido tratado en muchos trabajos [37][20][38], sigue siendo objeto de numerosos estudios teóricos y también relacionados con diversas aplicaciones [3].

El control por MD es fundamentalmente una consecuencia del control discontinuo. En los años sesenta, el control discontinuo era un tema de estudio para los ingenieros mecánicos e ingenieros de control, y se resolvía de forma rigurosa el problema de las oscilaciones que aparecen en los sistemas de control bang bang. Estos primeros estudios, estaban enfocados en el análisis y donde el fenómeno parece más bien como una molestia que debe evitarse, se volvió rápidamente a los problemas de la síntesis de diversas maneras. Uno de ellos estaba relacionado con el tiempo, y a este se le denominó control óptimo, hubo algunos otros que fueron aplicados para la linealización y robustez. En el primer caso, las discontinuidades en el control se producen en momentos determinados, como resultado de la solución de un problema variacional. En la segunda, que es de interés en este trabajo en el uso de un control discontinuo que fue una elección a priori. De hecho, era la hora de buscar maneras de diseñar lo que hoy se denomina: leyes de control robusto, el MD que fue descubierto a principios de los años sesenta, y fue enfocado para las necesidades de la aeronáutica militar, e incluso se utilizó antes del término de robustez, los ingenieros de control fueron en busca de leyes de control insensible a las variaciones del sistema a controlar. Las redes lineales utilizadas en estos, no dieron lugar a una indemnización suficiente para usar las altas ganancias necesarias para obtener la insensibilidad paramétrica. A principios de 1962, se inicio con la idea de B. Halmel [32] los estudios de compensadores no lineales, cuyo objetivo era superar las limitaciones anteriores. Típicamente, estas redes, que actúan sobre la señal de error  $x$  del sistema de retroalimentación.

Los sistemas con MD tienen la capacidad para ser una herramienta eficiente para el control complejo de alto orden en plantas dinámicas no lineales que operan bajo condiciones inciertas [24], este es un problema común para muchos procesos de tecnología moderna. Esto explica el alto nivel de la actividad de investigación y publicación en el área y el interés de prácticas de ingenieros en el control de MD durante las últimas dos décadas.

Las secciones siguientes mostrarán las inclusiones diferenciales, las cuales son necesarias, debido a que se trabaja con sistemas que son discontinuos o multivaluados y así abordar el tema del control por modos deslizantes.

### 2.4.1. Inclusiones diferenciales

Para sistemas dinámicos de la forma:

$$\frac{dx}{dt} = f(x, t) \quad (2.61)$$

no es posible asegurar la existencia y unicidad de soluciones cuando  $f$  no es localmente Lipschitz, discontinua o multivaluada en  $x$  [18]. Cuando  $f$  es multivaluada,  $f(x, t)$  es un subconjunto no vacío de  $\mathbb{R}^n$  para cada  $t$  en  $\mathbb{R}^+$  y  $x$  en  $\mathbb{R}^n$ , la expresión (2.61) se convierte en una inclusión diferencial en el sentido de Filippov:

$$\frac{dx}{dt} \in f(x, t) \quad (2.62)$$

Las funciones multivaluadas sirven en muchos casos para representar el comportamiento de sistemas físicos. Por ejemplo en un sistema en movimiento en el que se considera el efecto de fricción estática. Cuando la masa está en reposo, la fricción opone una fuerza igual a la aplicada, dentro de cierto rango, y la condición de movimiento no cambia. Funciones de este tipo no son continuas ni localmente Lipschitz, así que la existencia y unicidad de las soluciones no se asegura en el sentido clásico. La existencia (aunque no la unicidad) de soluciones para sistemas como (2.62) puede ser asegurada cuando  $f(x, t)$  satisface las condiciones, no muy restrictivas, que se establecen a continuación.

1.  $f(x, t)$  es tal que  $f(x, t)$  es un subconjunto de  $\mathbb{R}^n$ , no vacío, compacto y convexo para cada  $t \in \mathbb{R}^+$  y cada  $x \in \mathbb{R}^n$ .
2.  $f(x, t)$ , como función multivaluada de  $x$ , es semicontinua por arriba para todo  $t$ . Una función  $f(x, t)$  es semicontinua por arriba en un punto  $(x_0, t_0)$  si para cualquier  $\epsilon > 0$  existe  $\delta(\epsilon, x_0, t_0)$  tal que el conjunto  $f(x_1, t_1)$  está contenido en la  $\epsilon$ -vecindad del

conjunto  $f(x_0, t_0)$ , dado que un punto  $(x_1, t_1)$  pertenece a la  $\delta$ -vecindad del punto  $(x_0, t_0)$ .

3.  $f(x, t)$  es medible en  $x$  en el sentido de Lebesgue.
4.  $f(x, t)$  es localmente acotada.

Conceptualmente hablando, una función multivaluada  $f(x, t)$  es semicontinua por arriba en un punto  $(x_0, t_0)$  si, a medida que  $(x, t)$  tiende a  $(x_0, t_0)$ , el límite del conjunto  $f(x, t)$  está contenido en la cerradura del conjunto  $f(x_0, t_0)$ . En (Fig. 5) se ilustran las propiedades de convexidad y semicontinuidad por arriba de algunas funciones multivaluadas. Todas las funciones mostradas son semicontinuas por arriba con excepción de la señalada b). Ésta es, no convexa. De entre las funciones semicontinuas por arriba, la c) Es no convexa, y a) y d) cumplen las condiciones de semicontinuidad por arriba y convexidad requeridas para la existencia de soluciones de una inclusión diferencial.

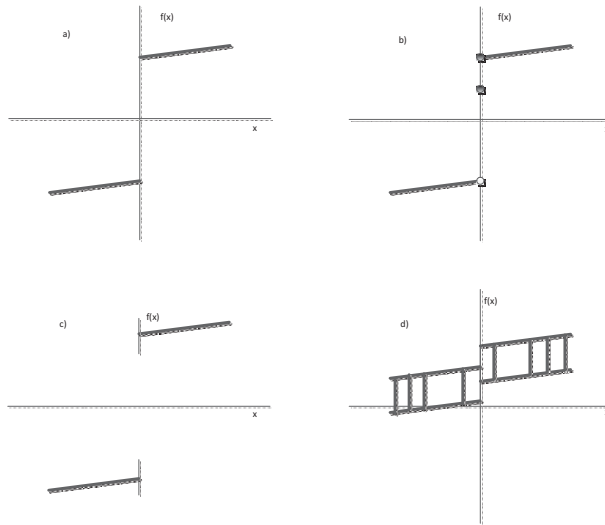


Fig. 5 Inclusiones Diferenciales.

Si  $f$  es discontinua, es posible modificarla de manera que en los puntos de discontinuidad la función sea determinada por medio de un conjunto, definido mediante los valores de la función en el punto de la discontinuidad. Este proceso se denomina convexificación y no afecta las propiedades globales de la función, pero permite asegurar la existencia de soluciones de ecuaciones (inclusiones) diferenciales que la incluyan.

Para ilustrar el procedimiento, puede considerarse la conocida función signo (Fig. 6). Su versión univaluada tradicional se define así:

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Convexificándola, se obtiene una función signo multivaluada:

$$F(x) = \begin{cases} 0 & x < 0 \\ [0, 1] & x = 0 \\ 1 & x > 0 \end{cases}$$

En la (Fig. 6) se puede observar la diferencia entre las funciones.

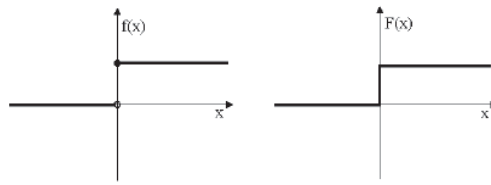


Figura 6. Función Signo.

El proceso de convexificación asegura la semicontinuidad por arriba de la función convexificada, y por ende, la existencia de soluciones. Se hace notar que la unicidad de soluciones de una inclusión diferencial ésta sujeta a condiciones bastantes restrictivas que la existencia, como por ejemplo la máxima monotonicidad de  $f(x, t)$ .

### 2.4.2. Modos deslizantes convencionales (MD)

El control por MD es una herramienta eficiente para controlar plantas de dinámica complejas de orden superior, que operan bajo condiciones de incertidumbre las cuales son comunes en muchos procesos de la tecnología moderna [26]. Los sistemas de control de estructura variable son una clase de sistemas, donde la ley de control es cambiada deliberadamente durante el proceso de control de acuerdo a algunas reglas predefinidas las cuales dependen del estado del sistema. Este tipo de control es un caso particular del control de estructura variable. En los sistemas de control de estructura variable (SCEV) se encuentra la combinación de leyes de control de realimentación y una regla de decisión. Esta regla tiene como

entrada alguna medida del comportamiento actual del sistema, y produce como salida el controlador de realimentación particular que debiera ser empleado en ese instante de tiempo. La combinación de estas leyes de control, y de la regla de decisión resulta de un sistema de estructura variable, el cual puede considerarse como una combinación de subsistemas donde cada uno tiene una estructura fija de control, y es solamente válido en regiones especificadas en el comportamiento del sistema. La ventaja primordial al introducir esta nueva complejidad al sistema, es la capacidad de disponer de propiedades útiles de cada una de las estructuras compuestas del sistema. Más aún, el sistema puede diseñarse para obtener nuevas propiedades que no están presentes en ninguna de las estructuras compuestas por sí solas.

Muchos sistemas físicos requieren naturalmente del uso de términos discontinuos en su dinámica o en acciones de su control. La piedra angular de este enfoque es la teoría de ecuaciones diferenciales discontinuos en el lado derecho [25][32]. Con base a ello, la lógica de conmutación está diseñada de tal manera, que una propiedad es que, el ajuste domine la dinámica controlada, lo que conduce a la estabilización múltiple deseada, lo que induce trayectorias deseables. Estos principios constituyen la idea principal en el desarrollo de uno de los métodos más eficaces para identificar y controlar una amplia clase de sistemas no lineales con incertidumbres. control por modos deslizantes (CMD) [21]. La característica esencial de ésta técnica, es la aplicación de las leyes retroalimentadas discontinuas para alcanzar y mantener la dinámica de ciclo cerrado en una variedad determinada en el espacio de estados (la superficie de conmutación) con algunas características deseadas para las trayectorias del sistema. Este control ofrece muchas ventajas en comparación con otro tipo de técnicas de control e identificación como son: comportamiento transitorio, la necesidad de una cantidad reducida de información en comparación con las técnicas de control clásico, la capacidad de rechazo a la perturbación no modelada, la insensibilidad a las no-linealidades de la planta o variaciones de los parámetros, notable estabilidad y robustez en el rendimiento.

**Definición 1** [26] *Se dice que existe un MD de primer o segundo orden en una variedad  $S$  dentro de una vecindad de un punto  $x$  de primer (o segundo) orden de MD, si en esta vecindad del punto  $x$  de primer (o segundo) orden de deslizamiento es un conjunto integrable, esto es, que son consistentes las trayectorias en el sentido de Filippov.*

Sea  $S_1 = S$ . Denotado por  $S_2$  el conjunto de puntos 2-deslizamiento con respecto a múltiples  $S$ , Se supone que  $S_2$  en sí puede ser considerado como una variedad suficientemente

suave. A continuación, la misma construcción se puede considerar con respecto a  $S_2$ . Denotado por  $S_3$  el correspondiente conjunto 2-deslizamiento se encuentra en lo que respecta a  $S_2$ .  $S_3$  es llamado conjunto de 3-deslizamiento con respecto a la variedad  $S$ . Se tienen dos casos para los sistemas de control utilizando modos deslizantes de alto orden los cuales son: primero; el deslizamiento ideal se produce cuando la restricción es idealmente mantenida, y segundo; el deslizamiento real tiene lugar cuando las imperfecciones de conmutación son tomadas en cuenta, y la restricción sólo se conserva aproximadamente.

### Modos deslizantes respecto a funciones con restricciones

Se tiene una restricción dada por una ecuación de  $s(x) = 0$ , donde  $s : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función con restricción suficientemente suave. También se supone que las derivadas del tiempo total a lo largo de la trayectoria de  $s, \dot{s}, \ddot{s}, \dots, s^{(r-1)}$  existen y son sólo las funciones con valores de  $x$ , lo cual no es trivial para sistemas dinámicos discontinuos. En otras palabras, esto significa que la discontinuidad no aparece en la primera  $r-1$  derivadas totales del tiempo de la función con restricción  $s$ . Entonces, el orden  $r$ -de deslizamiento está determinada por las igualdades:

$$s = \dot{s} = \ddot{s} = \dots = s^{(r-1)} = 0 \quad (2.63)$$

En esta ecuación (2.63) es una condición  $r$ -dimensional en el estado de la dinámica del sistema.

**Definición 2** [38] *Sea el conjunto de  $r$ - de deslizamiento dada en la ecuación (2.63) no es vacío y se supone que es localmente integrable en el sentido de Filippov (esto es, que consiste en las trayectorias de Filippov del sistema dinámico discontinuo). Entonces el correspondiente movimiento satisface la expresión (2.63) es llamado un  $r - MD$  con respecto a la función de restricción  $s$ .*

Para exhibir la relación con las definiciones anteriores, se considera a una variedad  $S$  dada por la ecuación  $s(x) = 0$ . Suponiendo que  $s, \dot{s}, \ddot{s}, \dots, s^{(r-2)}$  son funciones diferenciables de  $x$  y se tiene:

$$\text{rank} \{ \nabla s, \nabla \dot{s}, \dots, \nabla s^{(r-2)} \} = r - 1 \quad (2.64)$$

dominando localmente (aquí el rango  $V$  es una notación para el rango del conjunto de vectores  $V$ ). Entonces  $S_r$  esta determinado por (2.63) y todo  $S_i, i = 1, \dots, r - 1$  son variables suaves.



Si a su vez  $S_r$  se requiere para una variedad diferenciable, entonces esta última condición se extiende a:

$$\text{rank} \{ \nabla s, \nabla \dot{s}, \dots, \nabla s^{(r-1)} \} = r \quad (2.65)$$

La igualdad dada en la expresión (2.65) junto con la exigencia de las correspondientes derivadas de  $s$  que son funciones diferenciables de  $x$  se conoce como la condición de regularidad del deslizamiento, mientras que la condición (2.64) se llamará a las condiciones débiles de regularidad de deslizamiento.

Con la condición de regularidad débil satisfecha y  $S$  dada por la ecuación  $s = 0$ . La definición 3 es equivalente a la definición 2. Si la condición de regularidad (2.65) se mantiene, entonces las nuevas coordenadas locales pueden ser tomadas. En estas coordenadas, el sistema tiene la forma:

$$\begin{aligned} y_1 &= s, \dot{y}_1 = y_2; \dots; \dot{y}_{r-1} = y_r \\ s^{(r)} &= \dot{y}_r = \Phi(y, \varepsilon) \\ \varepsilon &= \Psi(y, \varepsilon), \varepsilon \in \mathbb{R}^{n-r} \end{aligned}$$

**Proposición 1** *Sea una condición de regularidad en la ecuación (2.65) se cumplan y  $r$ -de desplazamiento (2.72) que no está vacío. Entonces un  $r$ -MD con respecto a la función con restricción  $s$  existe si y sólo si la intersección con el conjunto de vectores de Filippov con el espacio tangencial para la variedad (2.72) es no vacío para cualquier punto  $r$ -de deslizamiento.*

## Deslizamiento ideal

Todas las consideraciones previas son trasladadas literalmente para el caso de un proceso controlado

$$\dot{x} = f(t, x, u), s = s(t, x) \in \mathbb{R}, u = U(t, x) \in \mathbb{R}$$

donde  $x \in \mathbb{R}^n$ ,  $t$  es el tiempo,  $u$  es el control,  $f$  y  $s$  son funciones suaves [38].

El control  $u$  esta determinado aquí por una retroalimentación  $u = U(t, x)$ , donde  $U$  es una función discontinua. Para simplificar la restricción se tiene el caso cuando  $s$  y  $u$  son escalares. Sin embargo, todos las siguientes afirmaciones podrían ser formuladas por el caso vectorial de MD.

Los MD estándar satisfacen la condición de que el conjunto de las velocidades posibles  $V$  no están en el espacio vectorial tangencial  $T$  para la variedad  $s = 0$ , pero se cruza con él, y

por lo tanto existe una trayectoria en la variedad  $S$  con el vector de velocidad que se extiende en  $T$ , tales MD son el principal modo de operación en los sistemas de estructura variable y de acuerdo con las definiciones anteriores son de primer orden. Cuando un error de conmutación está presente, los niveles de trayectoria están en un cierto ángulo. Por otra parte, en el caso del deslizamiento de segundo orden todas las velocidades posibles se encuentran en el espacio tangencial, e incluso cuando un error de conmutación está presente, la trayectoria de estado es tangencial a esa variedad en el momento de irse.

Para ver las conexiones con algunos resultados conocidos de la teoría de control, se examina en primer lugar el caso cuando:

$$\dot{x} = a(x) + b(x)u, s = s(x) \in \mathbb{R}, u \in \mathbb{R}$$

donde  $a, b, s$  es un vector de funciones suaves. El sistema tiene un grado relativo  $r$  respecto a la variable de salida  $s$  lo cual significa que las derivadas de posición  $L_b s, L_b L_a s, \dots, L_b L_a^{\tau-2} s$  son idénticamente igual a cero en una vecindad de un punto dado y  $L_b L_a^{\tau-1} s$  no es cero en ese punto. La igualdad del grado relativo para  $r$  significa, en una manera simplificada, que  $u$  aparece primero explícitamente sólo en el tiempo total derivable  $r^i$  de  $s$ . La igualdad del grado relativo de  $r$  significa, de una manera simplificada, que  $u$  aparece por primera vez explícitamente sólo en la derivada  $r^i$  del tiempo total de  $s$ . Se sabe que en ese caso  $s^{(i)} = L_a^i s$  para  $i = 1, \dots, r-1$ , la condición de regularidad dada en la ecuación (2.65) se cumple de forma automática y también  $\frac{\partial}{\partial u} s^{(r)} = L_b L_a^{\tau-1} s \neq 0$ . Hay una analogía directa entre la noción entre el grado relativo y la condición de regularidad de deslizamiento. En términos generales, puede decirse que la condición de regularidad de deslizamiento vista en la expresión (2.65) significa que el grado relativo con respecto a la discontinuidad no es menor que  $r$ . Del mismo modo, la idea del orden  $r^i$ -ésimo de MD es análogo a la idea de la dinámica cero [16].

La idea del grado relativo fue originalmente introducido para el caso autónomo únicamente. Sin embargo, se aplica ésta noción para el caso no-autónomo también. Como se mostró anteriormente, se introduce para el propósito de una variable ficticia  $x_{n+1} = t, x_{n+1} = 1$ .

Considerando un sistema dinámico de la forma:

$$\dot{x} = a(t, x) + b(t, x)u, s = s(t, x), u = U(t, x) \in \mathbb{R}$$

**Teorema 1** [38] *Sea el sistema que tiene grado relativo  $r$  con respecto a una función de salida  $s$  en algún punto  $r$  de deslizamiento  $(t_0, x_0)$ . Sea, también, la función de discontinuidad  $U$ , toma los valores de los conjuntos  $[K, \infty)$  y  $(-\infty, -K]$  en algunos conjuntos de medida no*

cero en cualquier vecindad de cualquier punto  $r$  de deslizamiento cerca del punto  $(t_0, x_0)$ . Entonces es desarrollado, con un  $K$  suficientemente grande, para la existencia de MD en alguna vecindad del punto  $(t_0, x_0)$ , el movimiento de deslizamiento satisface la ecuación dinámica-cero.

**Prueba** Este teorema es una consecuencia inmediata de la proposición, sin embargo, se puede detallar la prueba. Considerando algunas nuevas coordenadas locales  $y = (y_1, \dots, y_n)$  donde  $y_1 = s, y_2 = \dot{s}, \dots, y_r = s^{(r-1)}$ . En estas coordenadas de variedad  $L_r$  esta dada por la igualdad  $y_1 = y_2 = \dots = y_r = 0$  y la dinámica del sistema es el siguiente:

$$\begin{aligned} \dot{y}_1 &= y_2, \dots, \dot{y}_{r-1} = y_r \\ \dot{y}_r &= h(t, y) + g(t, y)u, g(t, y) \neq 0 \\ \dot{\xi} &= \Psi_1(t, y) + \Psi_2(t, y)u, \xi = (y_{r+1}, \dots, y_n) \end{aligned} \quad (2.66)$$

■

La noción de modo deslizante parece ser entendido en un sentido muy cerrado. Una serie de enfoques en algunos documentos aplican las técnicas de modos deslizantes de alto orden de una manera muy general desde el punto de vista diferencial-algebraico. Deje que la siguiente igualdad se cumpla de manera idéntica, como consecuencia de las ecuaciones del sistema dinámico:

$$P(s^{(r)}, \dots, \dot{s}, s, x, u^{(k)}, \dots, \dot{u}, u) = 0 \quad (2.67)$$

La ecuación (2.67) se supone que es solución con respecto a  $s^{(r)}$  y  $u^{(k)}$ . La función  $s$  puede depender de  $u$ . El orden  $r^i$  del MD se considera como un estado de equilibrio  $s = 0$  a ser alcanzado por un controlador que satisfaga la ecuación (2.67). Con el fin de lograr para algunas dinámicas  $s$  estable:

$$\sigma = s^{(r-1)} + a_1 s^{(r-2)} + \dots + a_{r-1} s = 0$$

la dinámica discontinua:

$$\dot{\sigma} = -\text{sign}\sigma \quad (2.68)$$

se proporciona. Para ello, el valor correspondiente de  $s^{(r)}$  se evalúa a partir de la expresión (2.68) y sustituir en (2.67). La ecuación obtenida se resuelve para  $u^{(k)}$ . Entonces el control dinámico está constituido por una ecuación diferencial obtenida para  $u$  la cual tiene una discontinuidad del lado derecho. Con este control de derivación sucesiva  $s, \dots, s^{(r-1)}$  sera una función suave de un sistema cerrado de variables de espacio.

## Deslizamiento Real y Tiempo Finito de Convergencia

Se recuerda que el objetivo es la síntesis de una  $u$  de control que tiene la restricción de  $s(t, x) = 0$ . La calidad del diseño de control está estrechamente relacionado con la precisión del deslizamiento. En realidad, ningún enfoque para el problema de diseño se prevé para mantener idealmente la restricción prescrita. Por lo que existe una necesidad para introducir algunos medios con el fin de proporcionar una capacidad de comparación de los diferentes controladores. Cualquier idea de MD puede ser comprendida como un límite de movimientos, cuando la imperfección de la conmutación desaparece y la frecuencia de conmutación tiende a infinito. Sea  $\varepsilon$  es alguna medida de las imperfecciones de conmutación. Entonces la precisión de deslizamiento de cualquier técnica de MD puede ser caracterizada por precisión asintótica de deslizamiento con  $\varepsilon \rightarrow 0$ .

**Definición 3** [32] Sea  $(t, x(t, \varepsilon))$  es una familia de trayectorias, indexada por  $\varepsilon \in \mathbb{R}^\mu$ , con condición común inicial  $(t_0, x(t_0))$ , y sea  $t \geq t_0$  (o  $t \in [t_0, T]$ ). Asumiendo que existe  $t_1 \geq t_0$  (o  $t_1 \in [t_0, T]$ ) de tal manera que en todos los segmentos  $[t', t'']$ , donde  $t' \geq t_1$ , la función  $s(t, x(t, \varepsilon))$  tiende uniformemente a cero con  $\varepsilon$  tendiendo a cero. En ese caso, se llama a una familia la familia real de deslizamiento con la restricción  $s = 0$ . Se le denomina al movimiento en el intervalo  $[t_0, t_1]$  un proceso transitorio, y el movimiento en el intervalo  $[t_1, \infty)$  un proceso de estado estacionario.

**Definición 4** [38] Un algoritmo de control, dependiente de un parámetro  $\varepsilon \in \mathbb{R}^\mu$ , se llama algoritmo real de deslizamiento en la restricción  $s = 0$  si, con  $\varepsilon \rightarrow 0$ , se forma una verdadera familia real de deslizamiento para cualquier condición inicial.

**Definición 5** Sea  $\gamma(\varepsilon)$  es una función real valuada tal que  $\gamma(\varepsilon) \rightarrow 0$  como  $\varepsilon \rightarrow 0$ . Un algoritmo real de deslizamiento en la restricción  $s = 0$ , se dice que es de orden  $r$  ( $r > 0$ ) con respecto a  $\gamma(\varepsilon)$  si para cualquier conjunto compacto con condiciones iniciales y para cualquier intervalo  $[T_1, T_2]$  existe una constante  $C$ , tal que el proceso de estado estacionario para  $t \in [T_1, T_2]$  satisface

$$|s(t, x(t, \varepsilon))| \leq C |\gamma(\varepsilon)|^r$$

En el caso particular cuando  $\gamma(\varepsilon)$  es un pequeño intervalo de tiempo de la suavidad de control, las palabras con respecto a  $\gamma$  pueden ser omitidas. Este es el caso cuando el deslizamiento real aparece como resultado de la discreción en la conmutación.

Este tipo de control tienen las siguientes características:

1. Insensible respecto a entradas desconocidas (mayor robustes).
2. Posibilidades de utilizar los valores de la aportación de la producción equivalente para la identificación de entrada desconocida.
3. Converge en un tiempo finito por los valores de los vectores de estado.

En efecto, el diseño se podrá resumir en dos pasos: primero se elige una variedad  $s = 0$ , tal que el sistema alcanza la función especificada mientras se mantenga en dicha variedad. Segundo, se diseña un control discontinuo que lleve al sistema a la variedad y lo mantenga para todo tiempo futuro [20][26]. La idea intuitiva es reaccionar inmediatamente a cualquier desviación que tenga el sistema de la variedad  $s = 0$ , empujándolo nuevamente hacia ella.

A continuación, se muestra un observador utilizando la teoría de control de modos deslizantes de orden superior que generaliza la idea del control por modos deslizantes de primer orden, actuando en las derivadas temporales superiores de la función de restricción  $s(x)$ , en lugar de hacerlo sólo sobre la primera derivada como ocurre en el MD de primer orden o estándar. Manteniendo las principales ventajas del enfoque original, el control por MDAO permite además reducir el efecto de chattering o ripple de alta frecuencia, que es considerablemente perjudicial en muchas aplicaciones prácticas.

En términos generales, la expresión de función de restricción  $s(x)$  se diseña según las especificaciones de control, siendo necesario que  $s$  se anule para que las mismas se verifiquen. En este contexto, el orden de deslizamiento está definido por el número de derivadas totales continuas de  $s$ , que se anulan en la vecindad del punto de deslizamiento (incluyendo la de orden cero). Por lo tanto, el  $r$ -ésimo orden de modo deslizante será determinado por las igualdades:

$$s = \dot{s} = \ddot{s} = \dots = \overset{n-1}{s} = 0$$

Formando así una condición  $r$  dimensional en el estado dinámico del sistema. Si bien los resultados teóricos para modos deslizantes de alto orden son prometedores, en la actualidad los algoritmos de modos deslizantes de segundo orden son los que poseen mayor aplicabilidad en la implementación, debido a su mayor simplicidad de diseño y menor requerimiento de información.

El principal problema de un sistema de modos deslizantes de orden  $n$ , es que por lo general es necesaria la información de  $s, \dot{s}, \ddot{s}, \dots, \overset{n-1}{s}$ . El algoritmo conocido como el super twisting, es un

controlador de sistemas de grado dos, implica 2 de deslizamiento, el cual necesita mediciones de  $s$  únicamente.

El problema consiste en determinar la velocidad de un sistema mecánico de segundo orden, cuando sólo se tiene disponible la posición. El modelo de un sistema mecánico de segundo orden es el siguiente:

$$M(q)\ddot{q} + C(q, \dot{q})\dot{q} + P(\dot{q}) + G(q) + \Delta(t, q, \dot{q}) = u \quad (2.69)$$

donde  $q \in \mathbb{R}^n$  es el vector de coordenadas generalizadas,  $M(q)$  es la matriz de inercia,  $C(q, \dot{q})$  es la matriz que incluye las fuerzas de coriolis y centrifugas,  $P(\dot{q})$  es la fricción de Coulomb,  $G(q)$  es el vector de fuerzas gravitacionales,  $\Delta(t, q, \dot{q})$  es un término de incertidumbre y  $u$  es el torque ejercido por el actuador.

La matriz  $M(q)$  es simétrica y definida positiva, por lo tanto tiene sentido su inversa, y podemos expresar el sistema con las variables de estado:

$$\begin{aligned} x_1 &= q \\ x_2 &= \dot{q} \end{aligned}$$

por lo que se tiene:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= f(t, x_1, x_2, u) + \xi(t, x_1, x_2, u) \\ y &= x_1 \end{aligned} \quad (2.70)$$

donde la parte nominal del sistema dinámico esta representado por la función:

$$f(t, x_1, x_2, u) = -M^{-1}(x_1)[C(x_1, x_2)x_2 + P(x_2) + G(x_1) - u] \quad (2.71)$$

y

$$\xi(t, x_1, x_2, u)$$

es el que incluye los términos de incertidumbre. Supondremos que  $f$  y  $\xi$  son medibles en el sentido de Lebesgue y uniformemente acotadas en cualquier conjunto compacto de los estados  $x_1$  y  $x_2$ . El objetivo es determinar  $x_2$  cuando sólo está disponible  $x_1$  y no conocemos  $\xi$ . Los observadores lineales no alcanzan un desarrollo adecuado para estos sistemas. Los observadores basados en el modelo y los diferenciadores de alta ganancia también tienen desventajas de exactitud y efectos no deseados.

### 2.4.3. Modos deslizantes respecto a funciones con restricciones

Se tiene una restricción dada por una ecuación de  $s(x) = 0$ , donde  $s : \mathbb{R}^n \rightarrow \mathbb{R}$  es una función con restricción suficientemente suave. También se supone que las derivadas del tiempo total a lo largo de la trayectoria de  $s, \dot{s}, \ddot{s}, \dots, s^{(r-1)}$  existen y son solo las funciones con valores de  $x$ , lo cual no es trivial para sistemas dinámicos discontinuos. En otras palabras, esto significa que la discontinuidad no aparece en la primera  $r - 1$  derivadas del tiempo total de la función con restricción  $s$ . Entonces, el orden  $r$ -deslizamiento está determinada por las igualdades:

$$s = \dot{s} = \ddot{s} = \dots = s^{(r-1)} = 0 \quad (2.72)$$

En esta ecuación (2.72) es una condición en el estado de la dinámica del sistema. Un MD se dice que es estable si el correspondiente conjunto integral deslizante es estable. A continuación se describen brevemente los algoritmos de los modos deslizantes, empezando por el twisting, hasta llegar al super twisting.

#### Algoritmo Twisting

Se debe tener un sistema de grado relativo 1. Considerando coordenadas locales  $y_1 = s$  y  $y_2 = \dot{s}$ , entonces después de la inicialización adecuada de fase, el problema de control por MD de segundo orden es equivalente al problema de estabilización en tiempo finito para sistemas de segundo orden inciertos con:

$$\begin{aligned} |\varphi| \leq \Phi, 0 < \Gamma_m \leq \gamma \leq \Gamma_M, \Phi > 0 \\ \begin{cases} \dot{y}_1 = y_2 \\ \dot{y}_2 = \varphi(t, x) + \gamma(t, x)\dot{u} \end{cases} \end{aligned} \quad (2.73)$$

con  $y_2(t)$  no medible pero con una función *sign* posiblemente conocida, con  $\varphi(t, x)$  y  $\gamma(t, x)$  son funciones inciertas con:

$$\Phi > 0, |\varphi| \leq \Phi, 0 \ll \Gamma_m \leq \gamma \leq \Gamma_M \quad (2.74)$$

Es históricamente el primer controlador conocido de 2 de deslizamiento, y que las características del algoritmo twisting entorno al origen del plano 2 de deslizamiento  $y_1 - y_2$  (Fig. 7). Las trayectorias realizan un número infinito de rotaciones, mientras convergen en un tiempo finito en el origen. La vibración de las magnitudes a lo largo de los ejes, así como la rotación decrece en una progresión geométrica. Los valores del control derivativo conmutan

en cada cruce con el eje, lo cual requiere disponibilidad de la función *sign* de la variable de deslizamiento en el tiempo de derivación  $y_2$ .

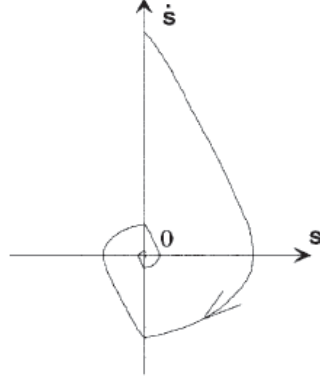


Fig. 7 Twisting SM

El algoritmo de control está definido por la siguiente ley de control, en la cual la condición de  $|u|$  prevé  $|u| \leq 1$  :

$$\dot{u}(t) = \begin{cases} -u & \text{si } |u| > 1 \\ -V_m \text{sign}(y_1) & \text{si } y_1 y_2 \leq 0; |u| \leq 1 \\ -V_M \text{sign}(y_1) & \text{si } y_1 y_2 > 0; |u| \leq 1 \end{cases} \quad (2.75)$$

Las correspondientes condiciones suficientes para la convergencia en tiempo finito en la variedad del deslizamiento son:

$$\begin{aligned} V_M &> V_m \\ V_m &> \frac{4\Gamma_M}{S_0} \\ V_m &> \frac{\Phi}{\Gamma_m} \\ \Gamma_m V_M - \Phi &> \Gamma_M V_m + \Phi \end{aligned} \quad (2.76)$$

El control similar esta dado como:

$$u(t) = \begin{cases} -V_m \text{sign}(y_1) & \text{si } y_1 y_2 \leq 0 \\ -V_M \text{sign}(y_1) & \text{si } y_1 y_2 > 0 \end{cases}$$

se va a utilizar para el control del sistema:

$$f(t, x, u) = a(t, x) + b(t, x)u(t) \quad (2.77)$$



cuando el grado relativo es 2.

Teniendo en cuenta el límite de las diferentes trayectorias de la dinámica incierta dada:

$$\ddot{s} = \varphi(t, x) + \gamma(t, x)u \quad (2.78)$$

y la evaluación de los intervalos de tiempo entre los pasos sucesivos del eje de las abscisas, es posible definir lo siguiente. El límite superior para el tiempo de convergencia es:

$$t_{tw\infty} \leq t_{M_1} + \Theta_{tw} \frac{1}{1 - \theta_{tw}} \sqrt{|y_{1M_1}|}$$

Aquí  $y_{1M_1}$  es el valor de la variable  $y_1$ , en el primer cruce de las abscisas en el plano  $y_1 - y_2$ ,  $t_{M_1}$  es el correspondiente instante de tiempo:

$$\begin{aligned} \Theta_{tw} &= \sqrt{2} \frac{\Gamma_m V_M + \Gamma_M V_m}{(\Gamma_m V_M - \Phi) \sqrt{\Gamma_M V_m + \Phi}} \\ \theta_{tw} &= \sqrt{\frac{\Gamma_m V_M + \Phi}{\Gamma_m V_M - \Phi}} \end{aligned}$$

En la práctica cuando  $y_2$  es no medible, su signo puede ser estimado por el signo de la primera derivada de la variable disponible de deslizamiento  $y_1$  en un intervalo de tiempo  $\tau$ , es decir,  $sign(y_2(t))$  se estima por  $sign(y_1(t) - y_1(t - \tau))$ . En ese caso, la precisión de 2 de deslizamiento con respecto al intervalo de tiempo medible se proporciona, y el tamaño de la frontera de la variedad de deslizamiento  $\Delta \sim O(\tau^2)$ . Hay que recordar que es la mejor precisión asintótica posible con discontinuidades  $\dot{y}_2 = \ddot{s}$ .

### Algoritmo Sub-optimal

En este caso se tiene el controlador 2 de deslizamiento y fue desarrollado como una implementación de retroalimentación sub-optimal de un control clásico de tiempo óptimo para un integrador doble. Dejando el grado relativo 2. El sistema auxiliar es:

$$\begin{cases} \dot{y}_1 = y_2 \\ \dot{y}_2 = \varphi(t, x) + \gamma(t, x)u \end{cases} \quad (2.79)$$

Las trayectorias en el plano  $y_1 y_2$  están confinadas dentro de los límites arcos parabólicos, que incluyen el origen, por lo que ambos comportamientos son posibles, esto es; saltan y se deslizan (Fig. 8) (cuando  $y_1$  y  $y_2$  no cambian de signo). También aquí las coordenadas con

las inserciones de la trayectoria con el eje  $y_1$  decrece geométricamente. Después de una fase de inicialización del algoritmo se define por la siguiente ley de control:

$$v(t) = -\alpha(t)V_M \text{sign}(y_1(t) - \frac{1}{2}y_{1M})$$

$$\alpha(t) = \begin{cases} \alpha^* & \text{si } [y_1(t) - \frac{1}{2}y_{1M}] [y_{1M} - y_1(t)] > 0 \\ 1 & \text{si } [y_1(t) - \frac{1}{2}y_{1M}] [y_{1M} - y_1(t)] \leq 0 \end{cases} \quad (2.80)$$

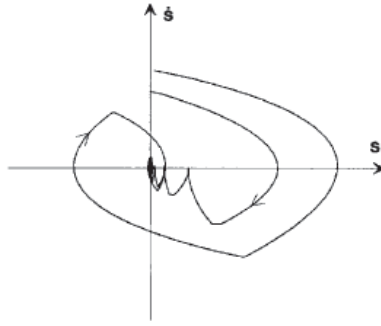


Fig. 8 Algoritmo Sub-optimal  
Trayectoria de Fase

donde  $y_{1M}$  es el último valor singular de la función  $y_1(t)$ , es decir, el último valor que corresponde al valor cero de  $\dot{y}_1 = y_2$ . Las condiciones correspondientes son suficientes para la convergencia en tiempo finito para la variedad deslizamiento y son las siguientes:

$$\alpha^* \in (0, 1] \cap (0, \frac{3\Gamma_m}{\Gamma_M})$$

$$V_M > \text{máx}(\frac{\Phi}{\alpha^*\Gamma_m}, \frac{4\Phi}{3\Gamma_m - \alpha^*\Gamma_M}) \quad (2.81)$$

También en ese caso un límite superior para el tiempo de convergencia se puede determinar:

$$t_{opt\infty} \leq t_{M_1} + \Theta_{opt} \frac{1}{1 - \theta_{opt}} \sqrt{|y_{1M_1}|}$$

$y_{1M_1}$  y  $t_{M_1}$  se definen para el algoritmo twisting, donde  $\Theta_{opt}$  y  $\theta_{opt}$  :

$$\Theta_{opt} = \frac{(\Gamma_m + \alpha^*\Gamma_M) V_M}{(\Gamma_m V_M - \Phi) \sqrt{\alpha^*\Gamma_m V_M + \Phi}}$$

$$\theta_{opt} = \sqrt{\frac{(\alpha^*\Gamma_M - \Gamma_m) V_M + 2\Phi}{2(\Gamma_m V_M - \Phi)}}$$

La eficacia del algoritmo anterior se extendió a una clase mucho mayor de sistemas inciertos. Se demostró que, en caso de una función de ganancia unitaria, la ley de control presentada en la expresión (2.80) se puede simplificar mediante  $\alpha = 1$  y la elección de  $V_M > 2\Phi$ .

El algoritmo de sub optimal requiere algún mecanismo con el fin de detectar los valores singulares de la variable disponible de deslizamiento  $y_1 = s$ . En el caso más práctico  $y_{1_M}$  puede ser estimado por el control para checar el signo de la cantidad  $D(t) = [y_1(t - \tau) - y_1(t)]$ , en el cual  $\frac{\tau}{2}$  es la estimación del retardo. En este caso, el control de la amplitud  $V_M$  debe pertenecer a un instante en el intervalo de la línea media:

$$V_M \in \left( \text{máx} \left( \frac{\Phi}{\alpha^* \Gamma_m}, V_{M_1}(\tau, y_{1_M}) \right), V_{M_2}(\tau; y_{1_M}) \right)$$

$V_{M_1} < V_{M_2}$  son soluciones de la ecuación algebraica de segundo orden dado:

$$\left[ \frac{(3\Gamma_m - \alpha^* \Gamma_M) V_{M_i}}{\Phi} - 4 \right] \frac{y_{1_M}}{\Phi \delta^2} - \frac{V_{M_i}}{8\Phi} [\Gamma_m + \Gamma_M(2 - \alpha^*)] \left( \frac{\Gamma_M V_{M_i}}{\Phi} + 1 \right) = 0$$

En el caso de la evaluación aproximada de  $y_{1_M}$  el MD real de segundo orden real del modo deslizante se consigue, y el tamaño de la frontera de la variedad del deslizamiento es  $\Delta \sim O(\tau^2)$ . Se puede minimizar eligiendo  $V_M$  de la siguiente manera:

$$V_M = \frac{4\Phi}{3\Gamma_m - \alpha^* \Gamma_M} \left[ 1 + \sqrt{1 + \frac{3\Gamma_m - \alpha^* \Gamma_M}{4\Gamma_M}} \right]$$

Una extensión del controlador sub-optimal 2 de deslizamiento para una clase de sistemas de datos muestreados, de manera que la función de ganancia de la ecuación (2.78) es constante, es decir:  $\gamma(\cdot) = 1$ .

### Algoritmo Super-twisting

Este algoritmo ha sido desarrollado para el control de sistemas con un grado relativo de orden 1, con el fin de evitar el chattering en control de sistemas variables. También en este caso las trayectorias en el plano 2-deslizamiento se caracterizan por el twisting alrededor al origen (Fig. 9), pero la ley de control continua  $u(t)$  está constituida por dos términos. El primero se define por medio de derivadas discontinuas en el tiempo, mientras que el otro es

una función continua de la variable de deslizamiento disponible.

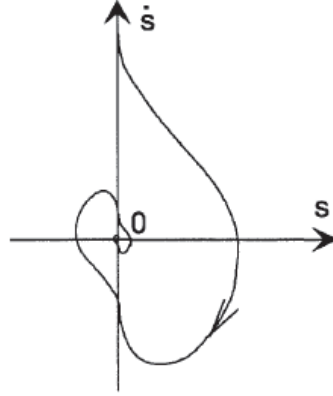


Fig.9 Algoritmo  
Super-Twisting Trayectoria  
de Fase.

El algoritmo de control está definido por la siguiente ley de control:

$$\begin{aligned}
 u(t) &= u_1(t) + u_2(t) & (2.82) \\
 \dot{u}_1(t) &= \begin{cases} -u & \text{si } |u| > 1 \\ -W \text{sign}(y_1) & \text{si } |u| \leq 1 \end{cases} \\
 u_2(t) &= \begin{cases} -\lambda |s_0|^\rho \text{sign}(y_1) & \text{si } |y_1| > s_0 \\ -\lambda |y_1|^\rho \text{sign}(y_1) & \text{si } |y_1| \leq s_0 \end{cases}
 \end{aligned}$$

y las correspondientes condiciones suficientes para la convergencia en tiempo finito para la variedad de deslizamiento son:

$$\begin{aligned}
 W &> \frac{\Phi}{\Gamma_m} \\
 \lambda^2 &\geq \frac{4\Phi(W+\Phi)}{\Gamma_m^2(W-\Phi)} \\
 0 < \rho &\leq 0,5
 \end{aligned} \tag{2.83}$$

El controlador puede simplificarse cuando los sistemas controlados (2.77) son linealmente dependientes sobre el control,  $u$  no necesariamente tiene que ser acotada y  $s_0 = \infty$ , entonces se tiene que el control se puede definir de la siguiente forma:

$$\begin{aligned}
 u &= -\lambda |s|^\rho \text{sign}(y_1) + u_1 \\
 \dot{u}_1 &= -W \text{sign}(y_1)
 \end{aligned}$$

El algoritmo de super twisting no necesita ninguna información sobre la derivada respecto al tiempo de la variable de deslizamiento. Un 2-MD es exponencialmente estable si la ley de control (2.82) con  $\rho = 1$  se utiliza. La elección de  $\rho = 0,5$  garantiza que la realización 2-deslizamiento se logre. Este tipo de control es extremadamente robusto, este controlador se utiliza con éxito para diferenciación exacta en tiempo real.

En la siguiente subsección se describe un enfoque que muestra la aplicación de los MD utilizando el algoritmo super twisting y proporcionando una función de Lyapunov fuerte, que cumple con los requisitos para implementarla en MD.

#### 2.4.4. Un enfoque de control de MDAO a través de Lyapunov

En este subtema, se presenta, una función de Lyapunov que se obtiene, para el algoritmo super twisting, el cual, es una clase importante de modo deslizante de segundo orden (MDSO). Este algoritmo se utiliza ampliamente en la literatura de modos deslizantes, para el diseño de controladores, observadores y diferenciadores exactos. La introducción de una función de Lyapunov no solo permite estudiar profundamente las propiedades conocidas de la convergencia en tiempo finito y la robustez a fuertes perturbaciones, sino también para mejorar el rendimiento mediante la adición de términos lineales de corrección para el algoritmo. Estas modificaciones permiten al sistema hacer frente a las perturbaciones de crecimiento lineal, que no son soportados por el algoritmo básico de super twisting. Por otra parte, la introducción de funciones de Lyapunov abre muchos nuevos análisis y herramientas de diseño para el área de modos deslizantes de orden superior.

Un algoritmo para un controlador y observador utilizando MDAO se llama super twisting, que esta descrito por la inclusión diferencial:

$$\begin{aligned}\dot{x}_1 &= -k_1 |x_1|^{1/2} \text{sign}(x_1) + x_2 + \varrho_1(x, t) \\ \dot{x}_2 &= -k_3 \text{sign}(x_1) + \varrho_2(x, t)\end{aligned}\tag{2.84}$$

donde  $x_i$  son las variables de estado,  $k_i$  son las ganancias diseñadas, y  $\varrho_i$  son las perturbaciones.

#### Dinámica sin perturbación

Para el sistema mostrada en la ecuación (2.84) se tiene la función:

$$V(x) = 2k_3 |x_1| + \frac{1}{2}x_2^2 + \frac{1}{2}(k_1 |x_1|^{1/2} \text{sign}(x_1) - x_2)^2\tag{2.85}$$

Es continua en todas partes pero no diferenciable en  $x_1 = 0$ , es una función de Lyapunov fuerte para el sistema (2.84) sin perturbaciones.

Se tiene que  $V(x)$  es continua pero no diferenciable, por lo que se requiere una versión no suave. Se hace notar que la versión usual para las funciones de Lyapunov para que sea localmente Lipschitz, en este caso no se cumple, donde  $V(x)$  no es localmente Lipschitz.

**Teorema 2** [21] *Se supone que  $k_1 > 0$  y  $k_3 > 0$ . Entonces todas las trayectorias del sistema sin perturbaciones (2.84), con  $\varrho_1 = \varrho_2 = 0$ , converge en tiempo finito al origen  $x = 0$ , en un tiempo pequeño  $t(x_0)$  que  $T = 2V^{1/2}(x_0)/\gamma$ , donde  $x_0$  es el estado inicial  $\gamma$  es una constante que depende de las ganancias  $k_1$  y  $k_3$ . Más aún,  $V(x)$ , definida en (2.85) es una función fuerte de Lyapunov que asegura estas propiedades.*

### Dinámica perturbada

El siguiente teorema asegura robustez de estabilidad asintóticamente estable de el equilibrio (2.84) en tiempo finito, cuando las ganancias son seleccionadas suficientemente grandes.

**Teorema 3** [21] *Supóngase que el término de la perturbación del sistema (2.84) esta globalmente acotada por:*

$$|\rho_1| \leq \delta_1 |x_1|^{1/2}, |\rho_2| \leq \delta_2$$

*para algunas constantes  $\delta_1, \delta_2 \geq 0$ . Entonces el origen  $x = 0$  es un punto de equilibrio que es fuertemente globalmente asintóticamente estable si las ganancias satisfacen:*

$$\begin{aligned} k_1 &> 2\delta_1 \\ k_3 &> k_1 \frac{5\delta_1 k_1 + 6\delta_2 + 4(\delta_1 + \delta_2/k_1)^2}{2(k_1 - 2\delta_1)} \end{aligned}$$

*sin embargo, todas las trayectorias convergen en tiempo finito a el origen, el límite superior es  $\tilde{T} = \frac{2V^{1/2}(x_0)}{\bar{\gamma}}$ , donde  $x_0$  es el estado inicial y  $\bar{\gamma}$  es una constante dependiente de las ganancias  $k_1, k_3$  y los coeficientes de perturbación  $\delta_1, \delta_2$ .*

### Algoritmo mejorado (MDAOM)

El algoritmo super twisting estudiado anteriormente es una versión no lineal del siguiente algoritmo básico lineal que tiene la siguiente estructura:

$$\begin{aligned} \dot{x}_1 &= -k_2 x_1 + x_2 + \varrho_1(x, t) \\ \dot{x}_2 &= -k_4 x_1 + \varrho_2(x, t) \end{aligned} \tag{2.86}$$

donde a pesar de sus propiedades ambos son muy diferentes. Por medio de la función candidata de Lyapunov dada como:

$$V(x) = k_4 x_1^2 + \frac{1}{2} x_2^2 + \frac{1}{2} (k_2 x_1 - x_2)^2 \quad (2.87)$$

es fácil ver que  $V(x)$  es definida positiva y radialmente no acotada si  $k_4 > 0$ , su derivada es:

$$\dot{V} = -x^T P_0 x + \varrho_1 q_0^T + \varrho_2 q_1^T x \quad (2.88)$$

donde:

$$\begin{aligned} P_0 &= k_2 \begin{vmatrix} (k_2^2 + k_4) & -k_2 \\ -k_2 & 1 \end{vmatrix} \\ q_0^T &= \begin{vmatrix} k_2^2 + 2k_4 & -k_2 \end{vmatrix} \\ q_1^T &= \begin{vmatrix} -k_2 & 2 \end{vmatrix} \end{aligned}$$

para el caso en donde las perturbaciones desaparecen i.e.  $\varrho_1 = \varrho_2 = 0$ , el estado  $x$  converge exponencialmente rápido a cero si  $P_0 > 0$ , que es, si  $k_2 > 0$  y  $k_4 > 0$ . Supóngase que las perturbaciones son globalmente acotadas por:

$$\begin{aligned} |\varrho_1| &\leq \delta_3 |x_1| \\ |\varrho_2| &\leq \delta_4 |x_1| \end{aligned}$$

para algunas constantes  $\delta_3, \delta_4 \geq 0$ , entonces:

$$\begin{aligned} \dot{V} &\leq -x^T (P_0 - \tilde{Q}) x \\ \tilde{Q} &= \begin{vmatrix} (k_2^2 + 2k_4)\delta_3 + k_2\delta_4 & \star \\ \frac{1}{2}(k_2\delta_3 + 2\delta_4) & 0 \end{vmatrix} \end{aligned}$$

existe convergencia exponencial si  $(P_0 - \tilde{Q}) > 0$  i.e. si:

$$\begin{aligned} k_2 &> 2\delta_3 \\ k_4 &> \frac{2\delta_3 k_2^3 + (\frac{1}{2}\delta_3^2 + 3\delta_4)k_2^2 + \delta_3\delta_4 k_2 + \delta_4^2}{k_2(k_2 - 2\delta_3)} \end{aligned}$$

La modificación de MDAO (MDAOM). Para un caso sin perturbaciones. Este nuevo algoritmo combina los terminos lineales con los no lineales, descrita por la siguiente inclusión diferencial:

$$\begin{aligned} \dot{x}_1 &= -k_1 |x_1|^{1/2} \text{sign}(x_1) - k_2 x_1 + x_2 \\ \dot{x}_2 &= -k_3 \text{sign}(x_1) - k_4 x_1 \end{aligned} \quad (2.89)$$

donde  $x_i$  son variables de estado y  $k_i$  son constantes de las ganancias.

**Teorema 4** [21] El origen  $x = 0$  es un punto de equilibrio del sistema (2.89) que es muy fuerte global y asintóticamente estable si  $k_i > 0, i = 1, \dots, 4$ , y  $4k_3k_4 > (8k_3 + 9k_1^2)k_2^2$ . En las mismas condiciones

$$V(x) = 2k_3 |x_1| + k_4 x_1^2 + \frac{1}{2} x_2^2 + \frac{1}{2} (k_1 |x_1|^{1/2} \text{sign}(x_1) + k_2 x_1 - x_2)^2$$

### Comportamiento con perturbaciones (MDAOM)

Ahora se presentan los términos de las perturbaciones en el sistema dado como:

$$\begin{aligned} \dot{x}_1 &= -k_1 |x_1|^{1/2} \text{sign}(x_1) - k_2 x_1 + x_2 + \varrho_1(x, t) \\ \dot{x}_2 &= -k_3 \text{sign}(x_1) - k_4 x_1 + \varrho_2(x, t) \end{aligned} \quad (2.90)$$

**Teorema 5** [21] Se supone que los términos de las perturbaciones del sistema (ecuación 2.90) son globalmente acotadas por:

$$\begin{aligned} |\varrho_1| &\leq \delta_1 |x_1|^{1/2} + \delta_3 |x_1| \\ |\varrho_2| &\leq \delta_2 + \delta_4 |x_1| \end{aligned}$$

para algunas constantes  $\delta_1, \delta_2, \delta_3, \delta_4 \geq 0$ . Entonces las ganancias se pueden seleccionar muy altas tal que el origen  $x = 0$  es un punto de equilibrio que es muy fuerte global y asintóticamente estable, y todas las trayectorias convergen en tiempo finito al origen. Para visualizar lo anterior se muestra el siguiente ejemplo.

**Ejemplo 1** [21] Considerando el modelo del péndulo el cual es

$$\begin{aligned} \ddot{\theta} &= \frac{1}{J} u - \frac{MgL}{2J} \sin(\theta) - \frac{V}{J} \dot{\theta} + v(t) \\ \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{1}{J} u - \frac{MgL}{2J} \sin(x_1) - \frac{V}{J} x_2 + v(t); y = x_1 \end{aligned}$$

donde  $x_1 = \theta$  (ángulo de oscilación) y  $x_2 = \dot{\theta}$  (velocidad angular),  $M$  es la masa del péndulo,  $g$  es la fuerza gravitacional,  $L$  es el largo del péndulo,  $J = ML^2$  polar de inercia,  $V$  es el coeficiente de fricción viscosa del péndulo, y  $v$  es la perturbación dada:

$$v(t) = 0,5 \sin(2t) + 0,5 \cos(5t)$$



Tenemos un observador MDAOM:

$$\begin{aligned}
 \dot{\hat{x}}_1 &= \hat{x}_2 + k_1 |e_1|^{1/2} \text{sign}(e_1) + k_2 e_1 \\
 \dot{\hat{x}}_2 &= \frac{1}{J} u - \frac{MgL}{2J} \sin(\hat{x}_1) - \frac{V}{J} \hat{x}_2 + k_3 \text{sign}(e_1) + k_4 e_1 \\
 e_1 &= x_1 - \hat{x}_1
 \end{aligned} \tag{2.91}$$

el error dinámico esta dado por:

$$\begin{aligned}
 \dot{e}_1 &= e_2 - k_1 |e_1|^{1/2} \text{sign}(e_1) - k_2(e_1) \\
 \dot{e}_2 &= \frac{MgL}{2J} (\sin(\hat{x}_1) - \sin(x_1)) - \frac{V}{J} e_2 - k_3 \text{sign}(e_1) - k_4 e_1 + v \\
 &= -k_3 \text{sign}(e_1) - k_4 e_1 + \varrho_2(e, t)
 \end{aligned}$$

donde  $\varrho_2(e, t) = \rho - \frac{MgL}{2J} \cos(z(t))e_1 - \frac{Vg}{J}e_2$  es el término de perturbación. Se observa que la perturbación contiene un término lineal en  $e_1$  (obtenido utilizando el teorema de valor medio), una perturbación acotada  $\rho$  y un término lineal en  $e_2$ . Se observa que la dinámica de este error es similar a (2.90), y por lo tanto, se usa esa relación mostrada para obtener  $k_i$ .

# Capítulo 3

## Desarrollo del Trabajo

### 3.1. Identificación por redes neuronales diferenciales hiperbólicas utilizando un pseudo observador de modos deslizantes.

Este capítulo expone la identificación de las EDP definidas en un dominio que tiene condiciones a la frontera Neuman y Dirichlet, la cual está definida en el espacio de Sobolev y es desarrollada en el sentido de Lebesgue, para que posteriormente se utilice el método de diferencias finitas y su aproximación basada en mallas, que servirá para aproximar la solución de las EDPH, más adelante se muestra la teoría general utilizada para aproximar las funciones inciertas, y se realiza la aproximación de RND basado en diferencias finitas, después se introduce el identificador no paramétrico basado en diferencias finitas. Para finalizar este capítulo se muestra el pseudo observador con RND, y las leyes de aprendizaje para los pesos del identificador donde se muestran los teoremas para probar la convergencia del identificador con ganancias constantes y variables.

#### 3.1.1. Identificación de Ecuaciones Hiperbólicas

Sea considerada una ecuación diferencial parcial hiperbólica incierta:

$$u_{tt}(x, t) = f(u(x, t), u_x(x, t), u_{xx}(x, t), v(x, t)) \quad (3.1)$$

donde  $u(x, t)$  está definida en el dominio  $x \in [0, 1], t \geq 0$  con condiciones en la frontera Neuman y Dirichlet. Las condiciones iniciales:

$$\begin{aligned} u_x(0, t) &= 0 \\ u(0, t) &= u_0, u(x, 0) = c \end{aligned} \quad (3.2)$$

La función  $v(x, t) \in \mathbb{R}$  puede considerarse como una perturbación externa medible o un diseño de acción de control distribuido. De hecho el identificador basado en RND únicamente necesita medir cada punto de  $(x)$  para todo tiempo  $(t)$  en (ecuación 3.1) y se tiene que su primera y segunda derivada respecto a  $t$  y a  $x$  se definen:

$$\begin{aligned} u_x(x, t) &= \frac{\partial u(x, t)}{\partial x} & u_{xx}(x, t) &= \frac{\partial^2 u(x, t)}{\partial x^2} \\ u_t(x, t) &= \frac{\partial u(x, t)}{\partial t} & u_{tt}(x, t) &= \frac{\partial^2 u(x, t)}{\partial t^2} \end{aligned} \quad (3.3)$$

el sistema presentado en la ecuación (3.1) considerado con condiciones iniciales y frontera mostrados en (3.2) es manejado como un espacio de  $H$  equipado con un producto interno  $\langle \cdot, \cdot \rangle$ .

En este trabajo se utiliza una función  $g(t)$  continua por partes en  $t$ . Por  $L(a, b, H)$  denota el conjunto de funciones  $g$  en valores de  $H$ , tal que  $(g(\cdot), x)$  es medible en el sentido de Lebesgue para toda  $x \in H$  y es  $\max_{t \in [a, b]} \|g(x, t)\| < \infty$ . Suponiendo que la función no lineal  $g(x, t)$  satisface la condición de Lipschitz:

$$\|g(x, t) - g(y, t)\| \leq L \|x - y\|, \forall z, y \in B_r(z_0) := \{z \in \mathbb{R}^n \mid \|z - z_0\| \leq r\}, \forall t \in [t_0, t_1]$$

Donde  $L$  es una constante y  $\|g\|^2 = \langle g, g \rangle$  es usado para asegurar que existe alguna  $\delta > 0$  tal que la ecuación de estado  $\dot{z} = g(z, t)$  con  $z(t_0) = z_0$  tiene una solución única sobre  $[t_0, t_0 + \delta]$  [7]. La norma está definida por el espacio de Sobolev y a continuación se define.

**Definición 6** [2] Sea  $\Omega$  un conjunto abierto en  $\mathbb{R}^n$  y sea  $v \in C^m(\Omega)$ . Definida la norma de  $v(x)$  como

$$\|v\|_{m,p} := \sum_{0 \leq \alpha \leq m} \left( \int_{\Omega} |D^\alpha v(x)|^p dx \right)^{\frac{1}{p}} \quad (3.4)$$

$(1 \leq p < \infty, D^\alpha v(x)) := \frac{\partial^\alpha}{\partial x^\alpha} v(x)$ . Esta es la norma de Sobolev en la cual la integración esta desarrollada en el sentido de Lebesgue. El complemento del espacio de la función  $v(x) \in C^m(\Omega) : \|v\|_{m,p} < \infty$  con respecto a  $\|\cdot\|_{m,p}$  es el espacio de Sobolev  $H^{m,p}(\Omega)$  para  $p = 2$ , el espacio de Sobolev forma un espacio  $H$ . A continuación se utilizará la norma (3.4) para las funciones  $u(x, t)$  para cada  $t$  fijo.

Se tiene que la aproximación de la solución de la EDP a través del método de diferencias finitas y aproximación por mallas mostrado en el capítulo anterior, se retomará para la solución de las EDP, mostradas en las ecuaciones (2.27,2.28 y 2.29).

### 3.1.2. Identificación con RND hiperbólica utilizando un pseudo observador de modos deslizantes

El sistema presentado en la expresión (3.1) es una clase de un sistema generalizado de segundo orden definido en el espacio de  $H$ . Por la representación de diferencias finitas, cada subsistema en el punto  $i$  puede ser representado como:

$$u = u_1; \dot{u} = u_2 \quad (3.5)$$

$$\dot{u}_{1,i}(t) = u_{2,i}(x, t)_{x=x_i} \quad (3.6)$$

$$\dot{u}_{2,i}(t) = f_i(x, v, t)_{x=x_i}$$

que puede ser representado de la siguiente forma:

$$\begin{aligned} \dot{\mu}_i(t) &= F_i(x, v, t)_{x=x_i} \\ F_i(x, v, t) &: = \begin{bmatrix} u_{2,i}(x, t)_{x=x_i} \\ f_i(x, v, t)_{x=x_i} \end{bmatrix} \end{aligned} \quad (3.7)$$

donde  $\mu_i(t) = [u_{1,i}(t), u_{2,i}(t)]^T \in \mathbb{R}^2$  es el vector de estado y  $v_i(t) \in \mathbb{R}^1$  es la acción de control aplicado dentro del sistema en tiempo  $t$ .

La solución de las anteriores ecuaciones diferenciales sin perturbación son comprendidas en el sentido de Filippov [18], esto es, la segunda ecuación en (3.7) donde aparece  $F_i(x, t)$  es remplazada por una inclusión diferencial equivalente:

$$\frac{d}{dt}\mu_i(x, v, t) \in \bar{F}_i(x, t, v) \quad (3.8)$$

Esta suposición está dada para considerar la posible aplicación de las perturbaciones discontinuas o controladores  $v_i(t)$ . Incluso cuando estas señales no sean discontinuas, la clase de soluciones son comprendidas en el sentido generalizado de Filippov para el caso continuo.

En vista de la continuidad casi en todos lados de  $F_i(x, v, t)$ , el conjunto valuado  $\bar{F}_i(x_0, v_0) = [\bar{F}_i(\cdot)]_{i=1,n}$ , es cerrado convexo de  $F_i(x, v, t)$  del conjunto de todos los límites de  $F_i(x_a, v_a)$  como  $[x_a, v_a] \rightarrow [x_0, v_0]$  donde  $[x_0, v_0]$  es el conjunto de todos los puntos continuos de  $F_i(x, v, t)$

para cualquier  $x_a \in X \subset \mathbb{R}^2$  y  $v_a \in V^{adm}$  donde  $V^{adm}$  es el conjunto de todos los controladores admisibles no lineales dado por:

$$V^{adm} := \left\{ v : \|v_t\|^2 \leq r_1 \|\mu(t)\|_{\Lambda_\mu}^2 < \infty \right\} \quad (3.9)$$

Este conjunto  $V^{adm}$  puede incluir diferentes diseños de control tales como la retroalimentación lineal, modos deslizantes convencionales y de alto orden, control integral, etc. Aquí  $r_1 \in \mathbb{R}^+$  y  $\Lambda_\mu$  es una matriz definida positiva ( $\Lambda_\mu = \Lambda_\mu^T, \Lambda_\mu > \infty$ ) con una adecuada dimensión ( $\Lambda_\mu \in \mathbb{R}^{2 \times 2}$ ).

### 3.1.3. Diferencias finitas y aproximación de RND para EDP inciertas

Utilizando una red neuronal para la aproximación de funciones no lineales desconocidas  $f_i(x, v, t)_{x=x_i}$ , han sido consideradas como una herramienta muy importante para resolver muchos problemas comunes dentro de la teoría de sistemas. Sin embargo, la presencia del error de modelado  $\tilde{f}_i$  han sido considerados para suavizar las condiciones de diseño. Este último término se asocia a la información no disponible para la construcción de la reproducción numérica de  $f_i(x, v, t)_{x=x_i}$ .

Por lo tanto, la siguiente ecuación es válida:

$$f_i(x, v, t)_{x=x_i} = f_{0,i}(x, t) + \tilde{f}_i(x, t) \quad (3.10)$$

La descomposición anterior esta basada en la capacidad de aproximación de las redes neuronales. Aquí se hace notar que  $f(\cdot, \cdot) \in \mathbb{R}^n$  y siempre puede ser representado (por los teoremas de Kolmogorot y Stone-Weisstrass [14]) como la composición nominal de  $f_0(x_t, u_t | \Omega) : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$  y el término del error de modelado  $\tilde{f}_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$  (como es usual cuando una aproximación de un modelo se aplica). La parte nominal  $f_0$  puede ser aproximada usando una descripción no lineal basada en la teoría de redes neuronales [33][31], utilizando una base general para reproducir la supuesta función no lineal desconocida. Muchas posibles sugerencias sofisticadas se pueden hacer para diseñar cualquier base adecuada para una aproximación numérica.

La idea principal empieza con la aplicación de RND para aproximar la solución de EDP's para utilizar una clase de método de diferencias finitas pero por funciones no lineales inciertas. En la sección de introducción, se estableció la contribución de la teoría de aproximación para construir modelos numéricos (usualmente llamados no-paramétricos) basados en una

aproximación general. En los párrafos siguientes se ofrece una descripción general sobre cómo producir una aproximación adecuada para cada  $f_i$ .

### 3.1.4. Teoría general utilizada para aproximar funciones inciertas

Para dar un enfoque suave a la propuesta presentada en este trabajo, vamos a considerar una función continua  $h_0(\cdot)$  definida en un espacio de  $H$  por [5],  $h_0(\cdot)$  puede ser reescrito en términos de la base  $H$  ( $\Psi_{ij}$ ) como:

$$\begin{aligned} h_0(z, \theta^*) &= \sum_i \sum_j \theta_{ij}^* \Psi_{ij}(z) \\ \theta_{ij}^* &= \int_{-\infty}^{+\infty} h_0(z) \Psi_{ij}(z) dz, \forall i, j \in \mathbb{Z} \end{aligned} \quad (3.11)$$

donde  $\{\Psi_{ij}(z)\}$  son funciones que contituyen una base en  $H$ . La última expresión se conoce como una función de expansión en series de  $h_0(z, \theta^*)$ . La red neuronal propuesta corresponde a  $h_0$ . La idea principal detrás del diseño de la red neuronal consiste en utilizar una construcción similar, pero con una clase especial de funciones base. Sin embargo, no se puede asumir que  $\theta_{ij}^*$  sea conocida. Por lo tanto, se debe proponer la función de approximation  $\hat{h}_0(z, \theta)$  diseñado para ser ajustada utilizando una estructura que varíe con el tiempo, la RN ajustable puede tener la siguiente estructura matemática:

$$\hat{h}_0(z, \theta(t)) := \sum_{i=M_1j=N_1}^{M_2} \sum_{j=N_1}^{N_2} \theta_{ij}(t) \Psi_{ij}(z) = \Theta(t) \Pi(z) \quad (3.12)$$

$$\begin{aligned} \Theta(t) &= [\theta_{M_1N_1}(t), \dots, \theta_{M_1N_2}, \dots, \theta_{M_2N_1}, \dots, \theta_{M_2N_2}]^T \\ \Pi(z) &= [\Psi_{M_1N_1}(z), \dots, \Psi_{M_1N_2}, \dots, \Psi_{M_2N_1}, \dots, \Psi_{M_2N_2}]^T \end{aligned} \quad (3.13)$$

que puede ser utilizado para aproximar una función no lineal  $h_0(z, \theta^*) \in H$  con una adecuada selección de enteros  $M_1, M_2, N_1, N_2 \in \mathbb{Z}^+$ . Siguiendo el teorema de Stone-Weistrass [14], si:

$$\varepsilon(M_1, M_2, N_1, N_2, t) = h_0(x, \theta^*) - \hat{h}_0(x, \theta(t)) \quad (3.14)$$

es el error de aproximación de la RN. Entonces para cualquier constante arbitraria definida positiva  $\varepsilon$  existen algunas constantes  $M_1, M_2, N_1, N_2 \in \mathbb{Z}$  tal que  $\forall z \in \mathbb{Z} \subset \mathbb{R}$ , se acota de la siguiente manera:

$$\sup_t \|\varepsilon(M_1, M_2, N_1, N_2, t)\|_2 \leq \varepsilon \quad (3.15)$$

**Observación 1** La selección apropiada de la función  $\Psi_{ij}(\cdot)$  es una importante tarea para construir una aproximación adecuada de funciones no lineales. Muchas funciones han sido reportadas en la literatura [22] que tienen resultados destacables para aproximar funciones no lineales desconocidas. La cual es la base más adecuada en la aplicación práctica que depende de cada especificación particular de diseño.

**Observación 2**  $M_1, M_2, N_1, N_2$  son los parámetros de diseño de las RN que están estrechamente relacionados con la aproximación de calidad  $\varepsilon(M_1, M_2, N_1, N_2, t)$ . La RN ha demostrado ser eficaz para reproducir funciones no lineales inciertas, las cuales satisfacen la condición de Lipschitz.

### 3.1.5. Aproximación con RND para EDPH

Siguiendo la idea presentada anteriormente, por una simple suma y sustracción de los términos necesarios, se puede representar la ecuación (3.1) como:

$$\begin{aligned} u_{tt}(x, t) = & A\mu(x, t) + \mathring{V}_1(x)\bar{\sigma}(x)\mu(x, t) + \mathring{V}_2(x)\bar{\varphi}(x)\mu_x(x, t) \\ & + \mathring{V}_3(x)\bar{\gamma}(x)\mu_{xx}(x, t) + \mathring{V}_4(x)\bar{\eta}(x)v(x, t) + \tilde{f}(x, t) \end{aligned} \quad (3.16)$$

Aquí  $\tilde{f}(x, t) \in \mathbb{R}$  representa el error del modelado y se define explícitamente por:

$$\tilde{f}(x, t) := f(u(x, t), u_x(x, t), u_{xx}(x, t)) - u_{tt}(x, t) \quad (3.17)$$

Los vectores  $A \in \mathbb{R}^{1 \times 2}$ ,  $\mathring{V}_1(\cdot) \in \mathbb{R}^{n_1}$ ,  $\mathring{V}_2(\cdot) \in \mathbb{R}^{n_2}$ ,  $\mathring{V}_3(\cdot) \in \mathbb{R}^{n_3}$ ,  $\mathring{V}_4(\cdot) \in \mathbb{R}^{n_4}$  son constantes y el conjunto de funciones  $\bar{\sigma}(x) \in \mathbb{R}^{n_1}$ ,  $\bar{\varphi}(x) \in \mathbb{R}^{n_2}$ ,  $\bar{\gamma}(x) \in \mathbb{R}^{n_3}$ ,  $\bar{\eta}(x) \in \mathbb{R}^{n_4}$  forman una base obedeciendo las siguientes condiciones:

$$\begin{aligned} \|\bar{\sigma}(x) - \bar{\sigma}(x')\| &\leq L_{\bar{\sigma}} \|x - x'\| & \|\bar{\varphi}(x) - \bar{\varphi}(x')\| &\leq L_{\bar{\varphi}} \|x - x'\| \\ \|\bar{\gamma}(x) - \bar{\gamma}(x')\| &\leq L_{\bar{\gamma}} \|x - x'\| & \|\bar{\eta}(x) - \bar{\eta}(x')\| &\leq L_{\bar{\eta}} \|x - x'\| \end{aligned} \quad (3.18)$$

También se puede mostrar que están acotados en  $x$ , esto es:

$$\|\bar{\sigma}(\cdot)\| \leq \sigma^+, \|\bar{\varphi}(\cdot)\| \leq \varphi^+, \|\bar{\gamma}(\cdot)\| \leq \gamma^+, \|\bar{\eta}(\cdot)\| \leq \eta^+ \quad (3.19)$$

note que  $\bar{\sigma}(x_i), \bar{\varphi}(x_i), \bar{\gamma}(x_i), \bar{\eta}(x_i)$  son los vectores de activación, no es un conjunto de funciones propias.

### 3.1.6. Aproximación de RND basado en diferencias finitas

La construcción de RND y aplicando la misma representación para la ecuación (3.16), se tiene que para cada  $i \in [1, N]$  :

$$u_{i,tt}(x, t) = f_{0,i}(x, t) + \tilde{f}_i(x, t) \quad (3.20)$$

Aquí el término  $f_{0,i}(x, t)$  se refiere usualmente a la dinámica nominal o aproximación de RND. Esta estructura obedece la forma básica de regresión descrita en [28]. Por lo tanto, el  $f_{0,i}(x, t)$  tiene la forma:

$$\begin{aligned} f_{0,i}(x, t) := & A\mu(x, t) \Big|_{x=x_i} + V_1(x)\bar{\sigma}(x)\mu(x, t) \Big|_{x=x_i} \\ & + V_2(x)\bar{\varphi}(x)\mu_x(x, t) \Big|_{x=x_i} + V_3(x)\bar{\gamma}(x)\mu_{xx}(x, t) \Big|_{x=x_i} \\ & + V_4(x)\bar{\eta}(x)v(x, t) \Big|_{x=x_i} \end{aligned} \quad (3.21)$$

por otra parte, el término:

$$\tilde{f}_i(x, t) := \tilde{f}(x, t) \Big|_{x=x_i} \quad (3.22)$$

es llamado el error del modelado, que representa la distancia entre la aproximación producida por la RND y la trayectoria real de la EDP en cada punto dentro del dominio espacial  $[0, 1]$ . Por un simple algoritmo matemático, se tiene:

$$\tilde{f}_i(x, t) = R_i(u_i(t), u_{i-1}(t) - u_{i-2}(t)) - f_{0,i}(x, t) \quad (3.23)$$

En esta ecuación  $V_{1,i} \in \mathbb{R}^{n_1}$ ,  $V_{2,i} \in \mathbb{R}^{n_2}$ ,  $V_{3,i} \in \mathbb{R}^{n_3}$ ,  $V_{4,i} \in \mathbb{R}^{n_4}$ . A partir de ahora, se supone que los términos el error del modelado satisfacen las siguientes suposiciones:

**Proposición 2** *El error del modelado esta absolutamente acotado en  $\Omega$  :*

$$\left\| \tilde{f}_i(x, t) \right\|^2 \leq \tilde{f}_{1,i} \|\mu_i(t)\|^2 \quad (3.24)$$

*La aplicación directa del método de diferencias finitas para la ecuación (3.10) conduce a:*

$$\begin{aligned} f_{0,i}(x, t) := & A_i\mu_i(x, t) + [\mathring{V}_{1,i}(x)\bar{\sigma}_i(x) + (\Delta x)^{-1}\mathring{V}_2(x_i)\bar{\varphi}(x_i) + (\Delta x)^{-2}\mathring{V}_3(x_i)\bar{\gamma}(x_i)]\mu_i(x, t) \\ & - [(\Delta x)^{-1}\mathring{V}_2(x_i)\bar{\varphi}(x_i) + 2(\Delta x)^{-2}\mathring{V}_3(x_i)\bar{\gamma}(x_i)]\mu_{i-1}(x, t) \\ & + [(\Delta x)^{-2}\mathring{V}_3(x_i)\bar{\gamma}(x_i)]\mu_{i-2}(x, t) + \mathring{V}_4(x_i)\bar{\eta}(x_i)v(x_i, t) \end{aligned} \quad (3.25)$$

*De hecho por la teoría de la aproximación introducido anteriormente, se puede representar esta expresión como:*

$$f_{0,i}(x(t), t) := \Omega_{0,i}^T(t)\Pi_i(t) \quad (3.26)$$



estas nuevas variables estan definidas así:

$$\begin{aligned}\Omega_{0,i}^T(t)u &:= [A_i \mathring{W}_{i,1}(x) \mathring{W}_{i,2}(x) \mathring{W}_{i,3}(x) \mathring{W}_{i,4}(x)] \\ \Pi_i^T(t) &:= [\mu(x_i, t)\sigma(x_i) \ \mu(x_i, t)\varphi(x_i) \ \mu(x_{i-1}, t)\gamma(x_i) \ \mu(x_{i-2}, t)\eta(x_i) \ v(x_i, t)]\end{aligned}\quad (3.27)$$

Esto se define como una aproximación por diferencias finitas de las RND de ecuaciones diferenciales parciales hiperbólicas.

Considerando la estructura dada en (3.5) y utilizando la descripción de la RND, la EDP puede ser aproximada por el siguiente identificador de las RND de segundo orden:

$$\begin{aligned}\dot{u}_{1,i}(t) &= u_{2,i}(x, t) \\ \dot{u}_{2,i}(t) &= \Omega_{0,i}^T(t)\Xi_i(t) + \tilde{f}_i(x, t)\end{aligned}\quad (3.28)$$

esta forma puede ser utilizada para mostrar la convergencia del identificador por RND.

### 3.1.7. Identificador no paramétrico

En este caso, siguiendo la teoría de redes neuronales, la sección nominal es propuesta como:

$$\begin{aligned}\bar{f}_{0,i}(x(t), t) &:= \Omega_{0,i}^T(t)\bar{\Pi}_i(t) \\ \bar{\Theta}_i^T(t) &:= [A_i V_1(x, t) |_{x=x_i} \ V_2(x, t) |_{x=x_i} \ V_3(x, t) |_{x=x_i} \ V_4(x, t) |_{x=x_i}] \\ \bar{\Pi}_i^T(t) &:= [\mu(x, t) |_{x=x_i} \ \bar{\sigma}(x)\mu(x, t) |_{x=x_i} \ \bar{\varphi}(x)\mu_x(x, t) |_{x=x_i} \ \bar{\gamma}(x)\mu_{xx}(x, t) |_{x=x_i} \ \bar{\eta}(x)v(x, t) |_{x=x_i}]\end{aligned}\quad (3.29)$$

donde  $V_j(x, t) \in \mathbb{R}^{n_j}$  son matrices constantes que definen los pesos óptimos y se definen de la siguiente manera:

$$V_j(x, t) |_{x=x_i} := V_j(x_i, t) \quad (3.30)$$

La función de activación  $\bar{\Pi}_i(z)$  son constituidas por la función sigmoide definida como:

$$S_r(x_t) := a_r \left( 1 + b_r \exp \left( - \sum_{j=1}^n c_j x_{j,t} \right) \right)^{-1}, \quad r = [1, n_j] \quad (3.31)$$

Estas funciones satisfacen las siguientes condiciones:

$$\begin{aligned}|S_r(x_t^1) - S_r(x_t^2)|^2 &\leq l_{S_r} \|x_t^1 - x_t^2\|^2 \\ x_t^1, x_t^2 &\in \mathbb{R}^l, \quad l \geq l_{S_r} \in \mathbb{R}^+\end{aligned}\quad (3.32)$$

El límite superior (3.15) garantiza si el sistema no lineal con incertidumbre (3.12) puede ser aproximado por un posible algoritmo adaptable. Esta propiedad es usualmente referenciada cuando el sistema es identificable [28].

La aplicación directa del método de diferencias finitas para (ecuación 3.21) conduce a:

$$\begin{aligned} \bar{f}_{0,i}(x(t), t) := & A_i \mu_i(x, t) + [\mathring{V}_{1,i}(x) \bar{\sigma}_i(x) + (\Delta x)^{-1} \mathring{V}_2(x_i) \bar{\varphi}(x_i) + (\Delta x)^{-2} \mathring{V}_3(x_i) \bar{\gamma}(x_i)] \mu_i(x, t) \\ & - [(\Delta x)^{-1} \mathring{V}_2(x_i) \bar{\varphi}(x_i) + 2(\Delta x)^{-2} \mathring{V}_3(x_i) \bar{\gamma}(x_i)] \mu_{i-1}(x, t) + \\ & [(\Delta x)^{-2} \mathring{V}_3(x_i) \bar{\gamma}(x_i)] \mu_{i-2}(x, t) + \mathring{V}_4(x_i) \bar{\eta}(x_i) v(x_i, t) \end{aligned} \quad (3.33)$$

De hecho por la teoría de la aproximación introducida anteriormente, se puede representar esta expresión como:

$$\bar{f}_{0,i}(x(t), t) := \Omega_i^T(t) \bar{\Pi}_i(t) \quad (3.34)$$

Estas nuevas variables se definen como:

$$\begin{aligned} \Omega_i^T(t) &:= [A_i W_{i,1}(x, t) W_{i,2}(x, t) W_{i,3}(x, t) W_{i,4}(x, t)] \\ \bar{\Pi}_i^T(t) &:= [\mu(x_i, t) \sigma(x_i) \hat{\mu}(x_i, t) \varphi(x_i) \hat{\mu}(x_{i-1}, t) \gamma(x_i) \hat{\mu}(x_{i-2}, t) \eta(x_i) v(x_i, t)] \\ W_{i,j}(x, t) &:= W_j(x, t) \big|_{x=x_i} \end{aligned} \quad (3.35)$$

Esta estructura es al del identificador basado en RND para la EDP basado en el método de diferencias finitas con las RND.

Como se puede comprender, el identificador real basado en la EDP será constituida por los  $N$  identificadores, trabajando cada uno en un punto específico  $x_i$ . Evidentemente, la aproximación final será obtenida por un algoritmo usual de interpolación usado en la reconstrucción de la solución final. Utilizando un método similar que se describio en la sección previa, se tiene:

$$\begin{aligned} \frac{d}{dt} \bar{u}_{1,i}(t) &= \bar{u}_{2,i}(x, t) \\ \frac{d}{dt} \bar{u}_{2,i}(t) &= \bar{f}_{0,i}(x(t), t) \end{aligned} \quad (3.36)$$

Esta forma será utilizada para mostrar la convergencia del identificador basado en las RND.

### 3.1.8. Pseudo-observador adaptativo con la RND

Hay que señalar que las condiciones de medición sólo permiten obtener el valor de  $\bar{u}_{1,i}(t)$ . Por lo tanto, la propuesta introducida del identificador en (3.36) no puede ser desarrollado. Así, en este trabajo se propone la aplicación de un llamado pseudo observador basado en RND. Esta idea proviene de la similitud que la estructura mostrada en (3.36) que tiene con los sistemas no lineales de segundo orden. Entre otros, los sistemas mecánicos son buenos ejemplos de tales sistemas. En estos sistemas mecánicos, donde la dinámica no lineal  $f_i(x, v, t)_{x=x_i}$

que puede ser descrita explícitamente utilizando el método de Euler-Lagrange, como muchos dispositivos electromecánicos, tales como el motor de inducción (MI) o generadores de alta potencia, obedecen a una dinámica no lineal con una estructura similar.

La estimación de estado de este tipo de sistemas ha atraído una gran cantidad de trabajos de investigación [12][38], especialmente cuando no hay sensores mecánicos para la medición de la velocidad o la posición. Esta es una consecuencia de la dificultad para medir simultáneamente ambas variables  $\bar{u}_{1,i}(t)$  y  $\bar{u}_{2,i}(t)$ , resultando en la propuesta natural para construir observadores adaptables. Por otra parte, la estructura no lineal compleja hace de la estimación de estado un verdadero reto para los diseñadores. Estos inconvenientes se han resuelto por muchas propuestas de observadores adaptables como son: paso a paso, el método convencional de modo deslizante, observador adaptable, las técnicas óptimas, etc.

Los observadores son sistemas dinámicos obtenidos de la información de la planta (conociendo la estructura del modelo y las variables de entrada y salida disponibles a partir de mediciones), las variables estimadas (conocidos como estados) o parámetros que no son conocidos o directamente medibles.

Existe una amplia gama de sistemas para los cuales los observadores no pueden ser diseñados con la teoría estándar, ya que se requiere la condición de Lipschitz para la existencia y unicidad de soluciones, implica que las funciones que definen al sistema sean continuas en todos los valores del estado. Entre los sistemas que se definen por funciones discontinuas o multivaluadas, y por lo tanto no cumplen con la condición de Lipschitz, se incluyen algunos que son tan comunes como la restricción mecánica o histéresis, así como híbridos, que han sido ampliamente estudiados. Para estos sistemas el trabajo de diseño de los observadores ha sido baja, debido en parte a la complejidad matemática inherente a la imposibilidad del uso de la teoría tradicional.

En la literatura sobre los observadores no lineales existen varios usos del diseño no lineal discontinuo. Por ejemplo, un observador de modo deslizante mediante la inyección de términos discontinuos con el fin de mejorar el rendimiento. También hay algunos trabajos publicados en el diseño de observadores para los sistemas que incluyen explícitamente no linealidades discontinuas o multivaluadas. El enfoque mencionado en estas publicaciones restringe severamente el tipo de sistemas no lineales con discontinuidades o multivaluadas, ya que requiere la unicidad de las soluciones, para lo cual, las no linealidades deben ser monótonas. Este es un requisito muy fuerte, ya que la unicidad de soluciones en los sistemas descritos por inclusiones diferenciales es más bien la excepción en la regla.

La teoría de inclusiones diferenciales se desarrollo hace varias décadas, para sugerir un

sistema adecuado y unificado, que incluyen no linealidades discontinuas o multivaluadas. La inclusión diferencial dentro de un mapa que define la dinámica del sistema, asigna que a cada elemento del dominio no es un elemento del codominio, sino un subconjunto de este, esto define como una función multivaluada. Las ecuaciones diferenciales tradicionales se incorporan en esta formulación, teniendo en cuenta a cada imagen como un conjunto de un elemento. Dado que la existencia (aunque no la unicidad) de soluciones con inclusiones diferenciales requiere que la imagen de cada elemento del dominio sea convexa y compacta, a menudo se encuentran en la práctica como son la función discontinua o multivaluada, debe ser convexificada para conocer cada condición y puede ser tratado por la teoría de inclusiones diferenciales [23][18].

Los observadores neuronales diferenciales son estudiados en el enfoque de MD, y se utilizan para obtener el proceso de aprendizaje algebraico, para la identificación en línea de una planta no lineal (diseño del modelo) completamente disponible para los estados. La característica esencial de la técnica de MD es la aplicación de las leyes de retroalimentación que es discontinuo para así lograr, y mantener la dinámica en ciclo cerrado de una variedad determinada en el espacio de estados (para la conmutación de la superficie), con algunas características deseadas para las trayectorias del sistema. Este método ofrece muchas ventajas sobre otras técnicas de identificación y control con un buen comportamiento transitorio, la necesidad de una cantidad reducida de información en comparación con las técnicas tradicionales de control, no como un modelo con capacidad de rechazo de perturbaciones, insensible a la no linealidad de la planta o variación de parámetros, un notable rendimiento y robustes a la estabilidad.

En general, hay algunos resultados agradables para el diseño de varios posibles observadores de la clase de sistemas no lineales dada en (3.12), incluso utilizando la técnica de modos deslizantes de alto orden (MDAO) [32]. Sin embargo, muchos de ellos requieren el completo conocimiento de mi sistema  $f_i(x, v, t)_{x=x_i}$  o debe admitirse que  $\|f_i(x, v, t)_{x=x_i}\|$  está acotada (existe una constante  $f_i^+ \in \mathbb{R}^+$  tal que  $\|f_i(x, v, t)_{x=x_i}\| \leq f_i^+, \forall i \in [1, N]$ ). Algunos trabajos han demostrado la capacidad de MDAO para lograr una reproducción excelente del estado desconocido  $\bar{u}_{2,i}(t)$  empleando sólo la información  $f_i^+$  [23]. Por otra parte, las propiedades agradables del método de SM para ser invariantes a algún tipo de incertidumbre podría no ser deseable, porque si existe la posibilidad de diseñar cualquier función de control utilizando estimación de estados, el dominio donde la acción de control es válida puede alcanzar grandes valores que no pueden ser realizables por los actuadores y, además, provocar un sobrecalentamiento por ejemplo en los amplificadores de potencia, (que se ded-

ican a darles movilidad a motores DC, motores paso a paso y otros).

Algunas de las características atractivas de los MDAO en comparación con el modo deslizante clásico de primer orden son ampliamente reconocidas y estas son: mayor precisión de movimientos, reducción del chattering, la convergencia en tiempo finito para sistemas con grado relativo dos [23][15], etc. En la mayoría de los casos, los modos deslizantes se obtienen mediante la inyección de términos no lineales discontinuos, dependiendo del error a la salida. Este diseño puede ser utilizado para construir sólidos algoritmos de control o de observación. La inyección discontinua deberá estar diseñada de tal manera que las trayectorias del sistema están forzadas a permanecer en una subvariedad contenida en la estimación del espacio de error (llamada superficie de deslizamiento). Para ambos, el problema de control y de observación tiene un movimiento resultante que se conoce como MD [37][33]. Este término discontinuo permite el rechazo de perturbaciones externas [36].

Los observadores de MD son ampliamente utilizados porque pueden converger en tiempo finito, tienen robustez en la estimación respecto a las perturbaciones e incertidumbres [6][16]. Una nueva generación de observadores basado en el modo deslizante de segundo orden ha sido desarrollado recientemente en [15][30]. En él fueron desarrollados diferenciadores exactos y robustos. Ese observador basado en el algoritmo llamado super twisting asegura la convergencia en tiempo finito a las trayectorias reales sin filtración o derivación numérica.

En [20] se propone un observador de modo deslizante de segundo orden basandose en una modificación del algoritmo super twisting para observar una amplia clase de sistemas mecánicos. Una versión discreta de tal observador es también presentado [15], su tiempo de convergencia es finito esto se demuestra por medio de las curvas mejoradas. En el mismo sentido, en [27], se muestra que las diferencias finitas son aplicables para la estimación en línea de derivadas de orden arbitrario en control homogéneo discontinuo.

Otra importante contribución en el campo del análisis de estabilidad para observadores MDAO, fueron desarrollados por [16]. En ese artículo, se obtuvo una función fuerte de Lyapunov para una clase de algoritmos de MDAO. Adicionalmente, una versión modificada del algoritmo super twisting es implementado sumando un término proporcional en su estructura. Este término ayuda al proceso de estimación con MDAOM.

### **3.1.9. Estructura del Pseudo Observador**

Continuando con los métodos de identificación de estados no paramétricos, a través de la observación adaptativa está compuesto por una reproducción adaptable de la sección nom-

inal desconocida  $f_0(\cdot, \cdot | \cdot)$  y un conjunto de términos correctivos utilizando la información disponible para el sistema incierto, que es la señal de salida. La propuesta dada en este trabajo, se refiere a la adaptación de ambas secciones, mediante la teoría de RND para reproducir la descripción del vagaje matemático para el sistema no lineal y el método de MDAO para evitar la presencia del chattering sobre la trayectoria del observador. La descripción del estado estimado utiliza los enfoques desarrollados en [15][1]. Estas técnicas demostraron de forma independiente, una gran capacidad para reconstruir los estados desconocidos para un sistema no lineal incierto afectado por perturbaciones. Sin embargo, ambos tienen algunos inconvenientes asociados, primero, para las RND se tiene la dificultad asociada con una rápida convergencia entre los estados medibles, las estimaciones correspondientes y la segunda por la incapacidad para proporcionar una buena aproximación para la estructura de la sección incierta del sistema no lineal.

Una nueva estructura mezclando las habilidades de este par de métodos es la base del observador:

$$\begin{aligned}\frac{d}{dt}\hat{u}_{1,i}(t) &:= \hat{u}_{2,i}(t) + \beta_1\lambda(\tilde{u}_{1,i}(t))\text{sign}(\tilde{u}_{1,i}(t)) \\ \frac{d}{dt}\hat{u}_{2,i}(t) &:= \Omega_i^T(t)\bar{\Pi}_i(t) + \beta_2\text{sign}(\tilde{u}_{1,i}(t)) \\ \tilde{u}_{1,i}(t) &:= u_{1,i}(t) - \hat{u}_{1,i}(t)\end{aligned}\tag{3.37}$$

Los pesos de las matrices  $(\Omega_i)$  proporcionan el tiempo en el que varía el comportamiento de adaptación para esta clase de observadores. Este observador utiliza un método de entrenamiento que se ejecuta en línea para ajustar los pesos, para desarrollar la representación de la ecuación (3.28) por las RND. Esta mezcla alternativa da un segundo nivel de robustez sobre las perturbaciones externas y las incertidumbres de modelado. Como se dijo anteriormente, las matrices  $\Omega_i$  son responsables de reproducir la dinámica desconocida del sistema.

La solución de la ecuación (3.37) se comprende en el sentido de Filippov. Eso significa que el pseudo observador genera un conjunto de trayectorias que dependen de la definición aplicada para la función signo, en este trabajo se utiliza:

$$\text{sign}(z) := \begin{cases} 1 & \text{si } x > 0 \\ [-1, 1] & \text{si } x = 0 \\ -1 & \text{si } x < 0 \end{cases}\tag{3.38}$$

Por lo tanto, el concepto de estabilidad fuerte se utilizará para demostrar que todas las trayectorias descritas en la expresión (3.37) convergen en el conjunto de EDO con incertidumbres, presentado en la ecuación (3.28).

### 3.1.10. Leyes de aprendizaje para identificar los pesos

Para cada  $i = 0, \dots, N$  se define el vector de funciones definiendo el error entre las trayectorias producidas por el modelo y el identificador de las RND, así como sus derivadas con respecto a  $x$  para cada  $i$ .

$$\tilde{u}_i(t) := \hat{u}_i(t) - u_i(t) \quad (3.39)$$

Las matrices de pesos satisfacen las siguientes ecuaciones no lineales de la matriz de ecuaciones diferenciales:

$$\dot{W}_{i,k}(t) := \Phi_{i,k}(\tilde{u}_i(t), \tilde{W}_{i,k}(t)) \quad (3.40)$$

donde:

$$\Phi_{i,k}(\tilde{u}_i(t), \tilde{W}_{i,k}(t)) := -k_{i,k} M^T P_i \zeta_i(t) \Xi_k^T(x_i, t) \quad (3.41)$$

$$\begin{aligned} \Xi_1(x_i, t) &:= \sigma(x_i) \hat{\mu}_i(x_i, t) \\ \Xi_2(x_i, t) &:= \varphi(x_i) \hat{\mu}_i(x_{i-1}, t) \\ \Xi_3(x_i, t) &:= \gamma(x_i) \hat{\mu}_i(x_{i-2}, t) \\ \Xi_4(x_i, t) &:= \eta(x_i) v(x_i, t) \end{aligned} \quad (3.42)$$

con constantes positivas  $k_{i,k} > 0$  ( $k = 1, \dots, 4$ ) y  $P_i$  ( $i = 0, \dots, N$ ) la cual es una matriz definida positiva y simétrica ( $(P_i)^T = P_i > 0$ ) de la ecuación algebraica de Riccati se define lo siguiente:

$$Ric(P_i) := P_i M \tilde{A}_i + \left[ M \tilde{A}_i \right]^T P_i + P_i R_i P_i + Q_i^P = 0 \quad (3.43)$$

donde cada una de las variables se definen como:

$$\begin{aligned} R_i^P &:= M \dot{W}_{i,1}(x) \Lambda_1^{-1} \dot{W}_{i,1} M^T + M \dot{W}_{i,2}(x) \Lambda_2^{-1} \dot{W}_{i,2} M^T + \\ &M \dot{W}_{i,3}(x) \Lambda_3^{-1} \dot{W}_{i,3} M^T + M \dot{W}_{i,4}(x) \Lambda_4^{-1} \dot{W}_{i,4} M^T + \frac{1}{2} N \Lambda_5 N^T \\ &M \Lambda_f M^T + \lambda_{\max} \{ \Lambda_f^{-1} \} \tilde{f}_{1,i} \\ Q_i^P &:= \alpha_\sigma I_{n \times n} + \alpha_v I_{n \times n} + d^2 O^T I_{n \times n} O \\ M &:= [0 \ 0 \ 1]^T \end{aligned} \quad (3.44)$$

Las leyes de aprendizaje presentadas en (3.40) se han obtenido utilizando la metodología de Lyapunov como se explica en el teorema principal de este trabajo y desarrollado en el apéndice. Debe tenerse en cuenta que las leyes de aprendizaje obtenidos por la metodología propuesta exclusivamente dependen de  $t$ . Una vez que el método de entrenamiento ha sido completado, se puede desactivar la ley de aprendizaje, mientras que la estructura de identificación esta fija con los pesos de los valores generados después del entrenamiento.

**Observación 3** *La clase especial de la ecuación de Riccati*

$$PA + AP + PRP + Q = 0$$

tiene una única solución positiva  $P$  si y sólo si [1] las siguientes cuatro condiciones dadas a continuación se deben cumplir:

1. La matriz  $A$  es estable.
2. El par  $(A, R^{1/2})$  es controlable.
3. El par  $(Q^{1/2}, A)$  es observable.
4. Las matrices  $(A, Q, R)$  son seleccionadas de tal manera que satisfacen la siguiente desigualdad

$$\frac{1}{4} (AR^{-1} - R^{-1}A) R (AR^{-1} - R^{-1}A) + Q \leq AR^{-1}A$$

lo cual restringe el mayor eigenvalor de  $R$  garantizando la existencia de una única solución positiva.

**Teorema 6** *Considere el modelo no lineal presentado en (3.1) dado por el sistema de EDP's con incertidumbres (perturbaciones) en los estados y en las salidas, sobre las condiciones en la frontera dadas en la ecuación (3.2). Sea también suponiendo que los parámetros en la identificación de la RND dados por la EDPH (3.5) son ajustados por las leyes de aprendizaje de la ecuación (3.40). Si la matriz definida positiva  $Q_i^P$  proporciona la existencia de soluciones positivas para las ecuaciones de Riccati  $P^i (i = 0, \dots, N)$  dada en (3.43), y para cualquier escalar positivo  $d \in \mathbb{R}^+$  tal que la siguiente desigualdad lineal matricial (DLM) tiene la solución:*

$$\left( P_i \begin{bmatrix} -\beta_1 & 0 & 1 \\ 0 & -d & 0 \\ -r_1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -\beta_1 & 0 & 1 \\ 0 & -d & 0 \\ -r_1 & 0 & 0 \end{bmatrix} P_i \right) \leq H_i \quad (3.45)$$

entonces el origen

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

es un punto de equilibrio que es global y asintóticamente estable. La prueba se muestra en el apéndice.



### 3.1.11. Estrategia basada en Lyapunov para probar la convergencia del identificador

La función candidata de Lyapunov es:

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

el vector  $\varsigma_i(t)$  se define como sigue:

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

en [16][], se muestran las características de esta clase de funciones estrictas de Lyapunov. Esta función en realidad es diferenciable, por lo tanto una opción no suave es requerida por la teoría de Lyapunov. De hecho, hay que señalar que generalmente el gradiente acostumbrado no se puede utilizar aquí. Además, el primer elemento de  $\varsigma_i(t)$  no puede diferenciable mediante la regla del producto.

## 3.2. Identificación por redes neuronales diferenciales hiperbólicas utilizando un pseudo observador de modos deslizantes con ganancias variables

En esta sección se propone el diseño del pseudo observador, utilizando el algoritmo super twisting con ganancias variables con una estrategia basada en RND, para la identificación no paramétrica de un modelo matemático descrito por EDPH. El problema de identificación se reduce hallando una expresión exacta de la dinámica de pesos utilizando las propiedades de RND. La metodología propuesta hace que ésta estrategia de control global como es aplicando el control por MD utilizando ganancias variables proporcione un alivio al problema del chattering, aquí el problema del modelado no paramétrico de la ecuación se resuelve con éxito.

En este apartado, consideremos que las ganancias del observador serán variantes en el tiempo, a continuación se muestra el pseudo observador que ya se había mencionado anteri-

ormente que se vuelve a mostrar:

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

La inclusión de ganancias variables en el algoritmo permite considerar una mayor cantidad de perturbaciones externas, incluyendo perturbaciones que pueden crecer con respecto al estado. Dado que la mayor parte del desarrollo presentado para el identificador anterior es muy similar al que se va a proponer a continuación, se omitirá mucho del mismo. Por esta razón, únicamente se integrará directamente el teorema.

**Teorema 7** *Considere el modelo no lineal exhibidas en la ecuación (3.1) dado por el sistema de EDP's con incertidumbres (perturbaciones) en los estados y las salidas del tipo expuesto en (3.24) sobre las trayectorias del sistema descrito en EDPH. Suponiendo que los parámetros en la identificación RND dado por (3.5) son ajustados por las leyes de aprendizaje en la ecuación (3.40). Si la matriz definida positiva  $Q_i^P$  proporciona la existencia de soluciones positivas para las ecuaciones de Riccati  $P^i (i = 0, \dots, N)$  dada en (3.43), y para cualquier escalar positivo  $d \in \mathbb{R}^+$  tal que la siguiente DLM tiene la solución:*

$$\left( P_i \begin{bmatrix} -\beta 1 & 0 & 1 \\ 0 & -d & 0 \\ -r1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -\beta 1 & 0 & 1 \\ 0 & -d & 0 \\ -r1 & 0 & 0 \end{bmatrix} P_i \right) \leq H_i$$

entonces el origen

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

es un punto de equilibrio que es global y asintóticamente estable, y la superficie de deslizamiento se puede alcanzar en un tiempo finito si las ganancias variables son seleccionadas como:

$$\begin{aligned}\beta_1(t) &\leq 0 \\ \beta_2(t) &:= (\text{sign}(\tilde{u}_{1,i}(t)))_{eq}\end{aligned}$$

donde  $(\text{sign}(\tilde{u}_{1,i}(t)))_{eq}$  satisface la siguiente ecuación:

$$\frac{d}{dt} (\text{sign}(\tilde{u}_{1,i}(t)))_{eq} = \tau (\text{sign}(\tilde{u}_{1,i}(t)))_{eq} + \text{sign}(\tilde{u}_{1,i}(t))$$

donde  $\tau$  es una constante positiva pequeña que satisface las condiciones del prefiltro establecido en literatura clásica de modos deslizantes.

La prueba de este teorema es muy similar al anterior teorema por lo cual se obvia su exposición.

### 3.2.1. Estrategia basada en Lyapunov para probar la convergencia del identificador

La función candidata de Lyapunov es:

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

el vector  $\varsigma_i(t)$  se define como sigue:

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

en [16], se muestran las características de esta clase de funciones estrictas de Lyapunov. Esta función en realidad es diferenciable, por lo tanto una opción no suave es requerida por la teoría de Lyapunov. De hecho, hay que señalar que generalmente el gradiente acostumbrado no se puede utilizar aquí. Además, el primer elemento de  $\varsigma_i(t)$  no pueden diferenciarse mediante la regla del producto.

$$\begin{aligned} \dot{V}_i(t) &= 2\zeta_i^T(t) P_i \dot{\zeta}_i(t) + \sum_{j=1}^4 \left[ k_{i,j} \text{tr} \left\{ \dot{W}_{i,j}^T(t), \tilde{W}_{i,j}(t) \right\} \right] \\ \dot{\zeta}_i(t) &= \begin{bmatrix} \frac{1}{2} |\tilde{u}_{1,i}|^{-1/2} \frac{d}{dt} (\tilde{u}_{1,i}(t)) \\ \frac{d}{dt} \tilde{u}_{1,i}(t) \\ \frac{d}{dt} \tilde{u}_{2,i}(t) \end{bmatrix} \end{aligned}$$

Por la misma razón se ha utilizado el análisis de continuidad asociado a la monotonía decreciente de la estructura del algoritmo del super twisting. Este esquema de análisis se mantiene dado que las ganancias variables no afectan este comportamiento.

# Capítulo 4

## Resultados Numéricos

Esta sección muestra los resultados por el algoritmo de identificación no paramétrica y de estimación desarrollados en el capítulo precedente, esto a través del uso de un pseudo observador por medio de modos deslizantes de alto orden, este trabajo ofrece una ventaja real en el análisis de sistemas con parámetros distribuidos, esto sólo puede ser comprobado por los resultados teóricos presentados en cada uno de los teoremas correspondientes de cada sistema. Con la finalidad de demostrar que el esquema de identificación y de estimación obtenidos pueden ser realizables por la simulación.

Se resolvieron una serie de métodos numéricos para simular el comportamiento de los algoritmos basados en redes neuronales diferenciales, los cuales han sido probados mediante los modelos matemáticos simplificados de sistemas de EDPH. Hay que mencionar que una de las habilidades de la redes neuronales diferenciales es el desconocimiento de la descripción matemática del sistema, pero identifica las trayectorias del sistema y por ello son una gran herramienta que se utiliza en el presente trabajo.

### 4.1. Primer sistema

Se ha seleccionado un sistema representado por ecuaciones diferenciales parciales de tipo hiperbólico con incertidumbres. Este sistema cumple las condiciones necesarias requeridas para lograr los resultados para la identificación utilizando el diseño del pseudo observador por modos deslizantes basado en la metodología de con RND.

A continuación, las simulaciones numéricas muestran la ilustración cualitativa de un sistema de referencia. Consideremos el problema simplificado de una cuerda vibrante. Se

representa la posición de un punto, en un instante  $t$ , por una función real continua  $u : [0; L] \times [0; 1)$ , donde  $[0, L]$  representa la cuerda en el marco de referencia. Se considera el siguiente sistema con parámetros distribuidos que representa el modelo de una cuerda vibrante definida como:

$$\begin{aligned} u_{tt}(x, t) &= 0,01u_{xx}(x, t) + 0,01 \sin(x, t) \\ u_x(0, t) &= 0, u_x(1, t) = 0 \end{aligned}$$

Este modelo supone que tiene acceso a medidas discretas del estado  $\sin(x, t)$  a lo largo de su dominio. Este modelo se utilizará sólo para generar los datos necesarios para probar el identificador basado en RND. Las condiciones de frontera anteriores corresponden a la situación en el extremo de la cuerda en  $x = 1$  se fija en el extremo  $x = 0$  es libre. La condición de frontera pendiente cero en  $x = 0$  tiene el significado físico de que ninguna fuerza se aplica en ese extremo. Los parámetros utilizados en la identificación para las simulaciones fueron seleccionados de la siguiente manera:

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ -88,51 & -95,62 \end{bmatrix}, P = \begin{bmatrix} 0,004 & 1 \\ 0 & 0,0028 \end{bmatrix} \\ S &= \begin{bmatrix} 5^{-7} & 0 \\ 0 & 5^{-7} \end{bmatrix}, T = \begin{bmatrix} 2^{-7} & 0 \\ 0 & 2^{-7} \end{bmatrix} \\ W_{1,i}(0) &= W_{2,i}(0) = W_{3,i}(0) = \begin{bmatrix} 31,78 \\ 39,2 \\ 23,8 \end{bmatrix} \end{aligned}$$

El pseudo observador basado en las RND genera las trayectorias que están muy cerca de la trayectoria real del sistema mostrado en (Fig. 10). Las trayectorias representadas en la figura fueron generadas por el pseudo observador que se propone en este trabajo. Se puede ver la cercanía entre ambas figuras (Fig. 10 y Fig. 11) que es una consecuencia de la estructura variable propuesta. El error entre las trayectorias generadas por el modelo y la propuesta del identificador es cercana a cero casi para todo  $x$  y  $t$  lo que muestra la eficiencia del proceso de identificación proporcionada por el algoritmo sugerido por medio del uso de las RND (Fig. 12). Esta semejanza demuestra la posibilidad de aproximar con gran precisión la solución de una ecuación diferencial parcial hiperbólica con incertidumbre. Aquí se puede observar la convergencia a una pequeña zona cercana a cero que se define por las incertidumbres considerado en la ecuación diferencial parcial hiperbólica.

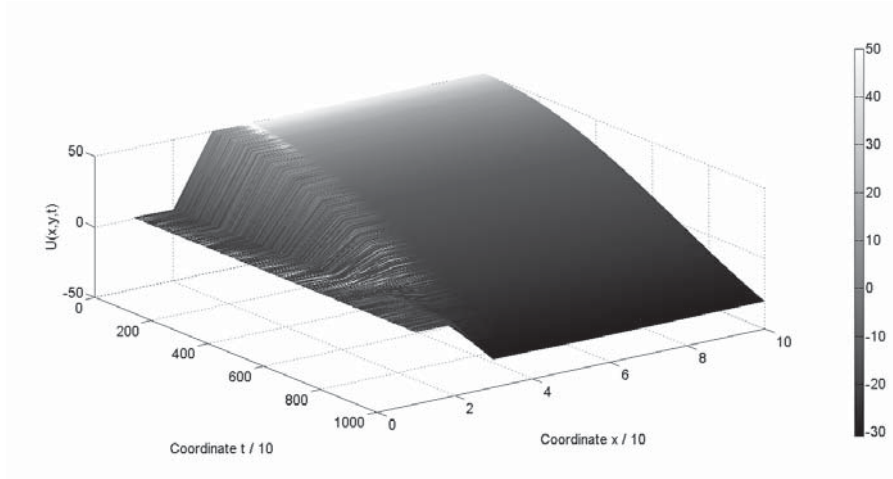


Fig. 10 Ecuación Diferencial Hiperbólica Original.

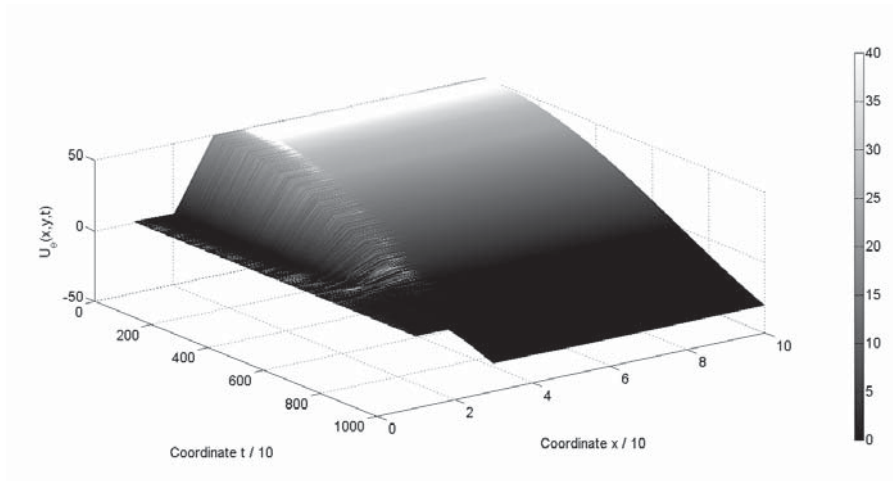


Fig. 11 Ecuación Diferencial Hiperbólica Estimada.

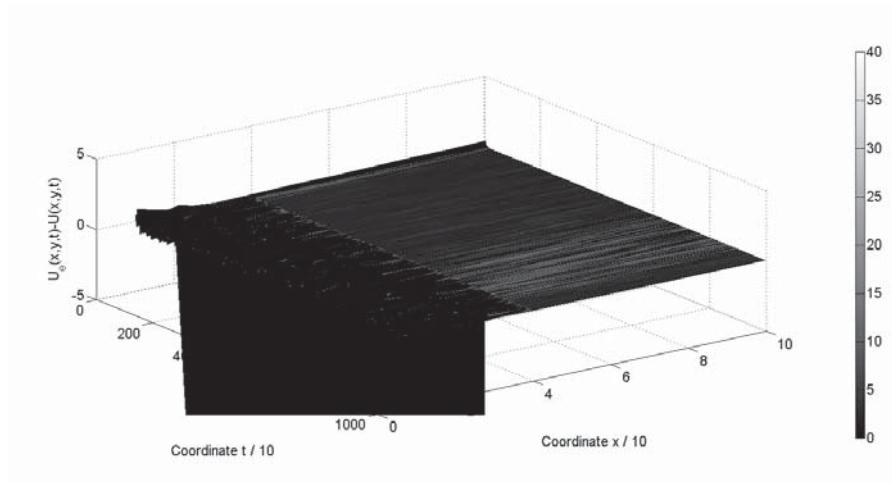


Fig. 12 Error del Sistema 1

Se puede visualizar que en (Fig. 12) este error del modelado se obtiene al restar las imágenes original-estimada y se observa que este error está cercano a cero durante casi todo  $x$  y todo  $t$ . Esto demuestra, al menos cualitativamente la eficiencia del proceso de identificación proporcionada por este algoritmo de RND.

## 4.2. Segundo Sistema

El segundo sistema es un poco más complejo que el anterior como se puede observar en la siguiente ecuación (4.1). Este sistema se utilizó para demostrar que es posible la estabilidad con el pseudo observador, para identificar cualquier otro sistema a través de las RND:

$$u_{tt}(x, t) = Du_{xx}(x, t) + Vu_x(x, t) + \Gamma u(x, t) + A \sin(x, t) \quad (4.1)$$

Se supone que tienen acceso a medidas discretas del estado  $\sin(x, t)$ ;  $u_{xx}$ ;  $u_x$ , a lo largo de todo su dominio y así lo hace  $u$ . Este modelo se utilizará sólo para generar los datos para probar el identificador basado en RND. Se asume que las condiciones iniciales han de ser

fijas como en el ejemplo anterior. Los parámetros utilizados en la simulación son:

$$D = \begin{bmatrix} 0 & 1 \\ -88,51 & -95,62 \end{bmatrix}, V = \begin{bmatrix} 0,004 & 1 \\ 0 & 0,0028 \end{bmatrix}$$

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

$$W_{1,i}(0) = W_{2,i}(0) = W_{3,i}(0) = \begin{bmatrix} 31,78 \\ 39,2 \\ 23,8 \end{bmatrix}$$

Los resultados numéricos siguientes se han obtenido con los parámetros anteriores. Estas trayectorias se utilizan sólo como generador de datos. No hay información con respecto al modelo y se utiliza en la definición del identificador. Considerando el efecto de la perturbación externa  $A \sin(x, t)$  y la trayectoria oscilante natural asociado a este sistema, la simulación numérica se muestra en (Fig. 13). La trayectoria producida por el identificador se muestra (Fig. 14). Se pueden ver las similitudes entre ambas trayectorias respecto al sistema original y el observador. Incluso cuando ambas son similares, definidas en la escala de grises las trayectorias muestran pequeñas variaciones que son consecuencias naturales del régimen de adaptación utilizadas por el Identificador para producir su trayectoria.

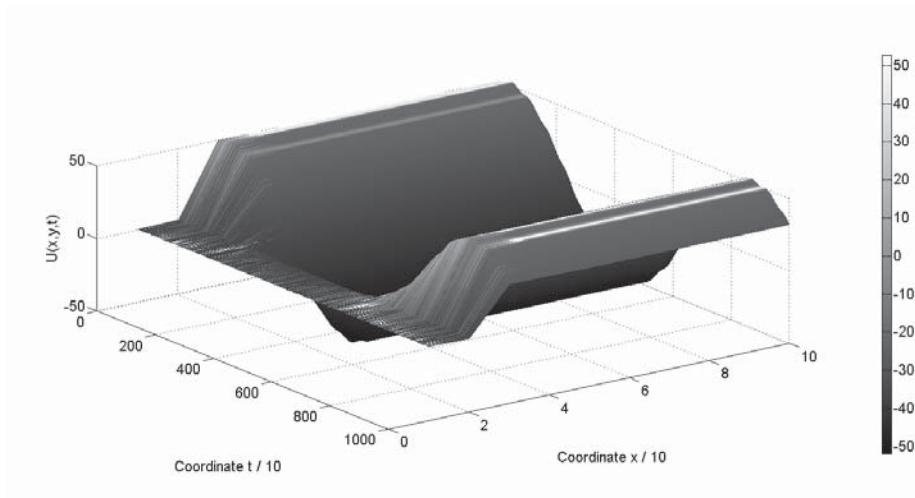


Fig. 12 Ecuación Diferencial Hiperbólica 2 Original.



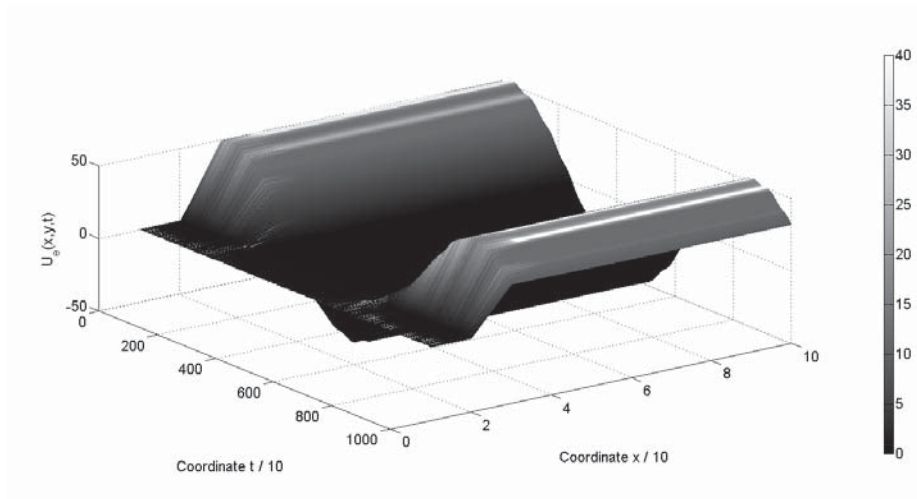


Fig. 13 Ecuación Diferencial Hiperbólica 2 Estimada.

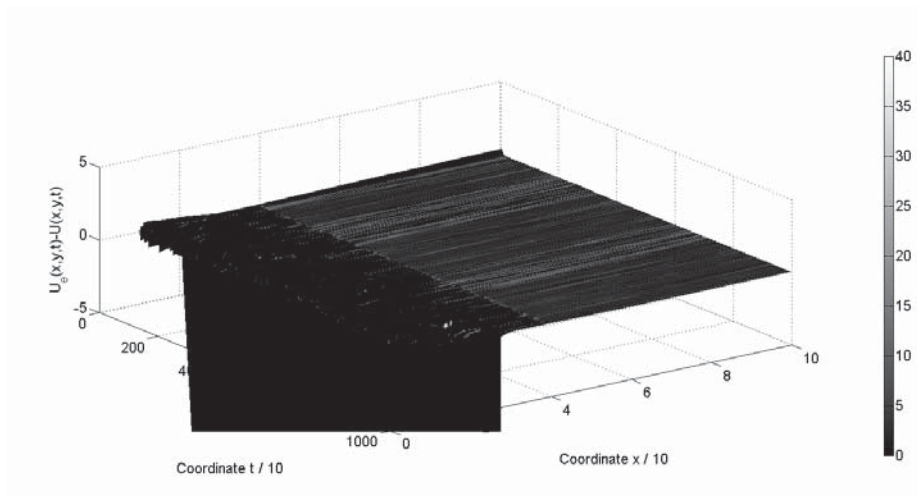


Fig.15 Error del Sistema 2

Se puede visualizar que en la figura mostrada en (Fig. 15), la cual muestra el error que al igual que en el sistema anterior está muy cercano a cero durante todo  $x$  y  $t$ . Esto demuestra de manera cualitativa la eficiencia del proceso de identificación proporcionada por el algoritmo de las RND.

### 4.3. Tercer sistema: Ganancias variables

Se ha seleccionado un sistema descrito mediante EDPH con incertidumbres. Este sistema cumple con las condiciones necesarias requeridas para lograr los resultados para la identificación de los parámetros con el uso del diseño del pseudo observador de modo deslizante basado en la metodología de RND, pero en este sistema se utilizan ganancias variables las cuales se mostraron anteriormente.

A continuación, las simulaciones numéricas muestran la ilustración cualitativa de un sistema de referencia. Considerando el problema simplificado de un sistema de alas de avión cuerda vibrante. Se representa la posición de un punto, en un instante  $t$ , por una función real continua  $u : [0; L] \times [0; 1)$ , donde  $[0, L]$  representa la cuerda en el marco de referencia. Se considera el siguiente sistema con parámetros distribuidos que representa el modelo de una cuerda vibrante

$$\begin{aligned} u_{tt}(x, t) &= 0,01u_{xx}(x, t) + 0,01 \sin(x, t) \\ u_x(0, t) &= 0, u_x(1, t) = 0 \end{aligned}$$

Este modelo supone que tiene acceso a medidas discretas del estado  $\sin(x, t)$  a lo largo de su dominio. Este modelo se utilizará sólo para generar los datos necesarios para probar el identificador basado en RND. Las condiciones de frontera anteriores corresponden a la situación en el extremo de la cuerda en  $x = 1$  se fija en el extremo  $x = 0$  es libre. La condición de frontera pendiente cero en  $x = 0$  tiene el significado físico de que ninguna fuerza se aplica en ese extremo. Los parámetros utilizados en la identificación para las simulaciones fueron seleccionados de la siguiente manera:

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ -88,51 & -95,76 \end{bmatrix}, P = \begin{bmatrix} 4^{-3} & 0 \\ 0 & 2,8^{-3} \end{bmatrix} \\ S &= \begin{bmatrix} 5^{-7} & 0 \\ 0 & 5^{-7} \end{bmatrix}, T = \begin{bmatrix} 2^{-7} & 0 \\ 0 & 2^{-7} \end{bmatrix} \\ W_{1,i}(0) &= W_{2,i}(0) = W_{3,i}(0) = \begin{bmatrix} 31,78 \\ 39,2 \\ 23,8 \end{bmatrix} \end{aligned}$$

El identificador RND genera las trayectorias que están muy cerca de la trayectoria real del sistema (Fig. 16). Las trayectorias representadas en la figura fueron generadas por el pseudo

observador que se propone en este trabajo. Se puede ver la cercanía entre ambas figuras (Fig. 17) que es una consecuencia de la estructura variable propuesta del identificador. El error entre las trayectorias generadas por el modelo y la propuesta del identificador es cercana a cero casi para todo  $x$  y  $t$  lo que muestra la eficiencia del proceso de identificación proporcionada por el algoritmo sugerido por la RND. Esta semejanza demuestra la posibilidad de aproximar con gran precisión la solución de una ecuación diferencial parcial hiperbólica con incertidumbre. Aquí se puede ver la convergencia a una pequeña zona cercana a cero que se define por las incertidumbres considerado en la ecuación diferencial parcial hiperbólica.

En las siguientes figuras se muestran los resultados gráficos de como el sistema converge en tiempo finito y se logra la estabilidad asintótica, es visible observar como el sistema decrece y se obtiene dicha estabilidad

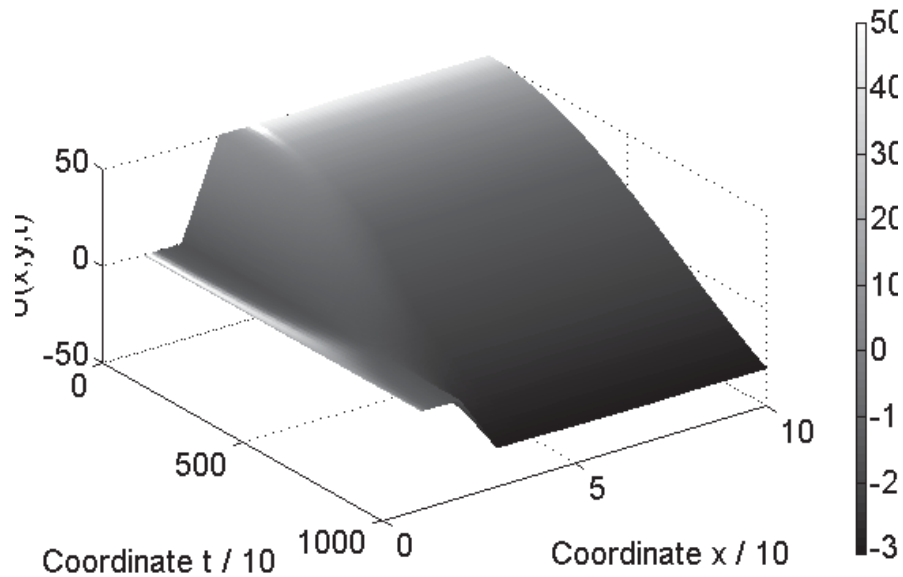


Fig.16 Sistema Original con GV

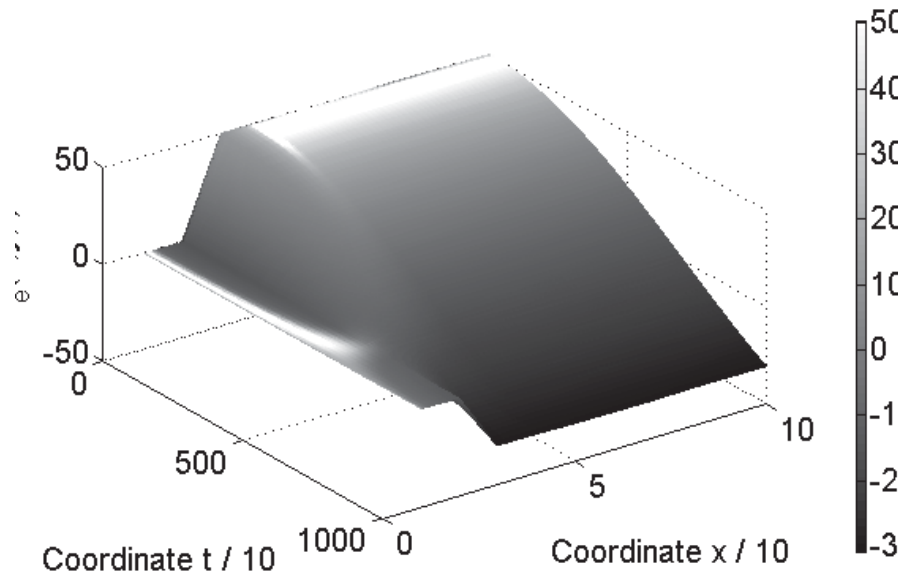


Fig.17 Sistema Identificado con GV.

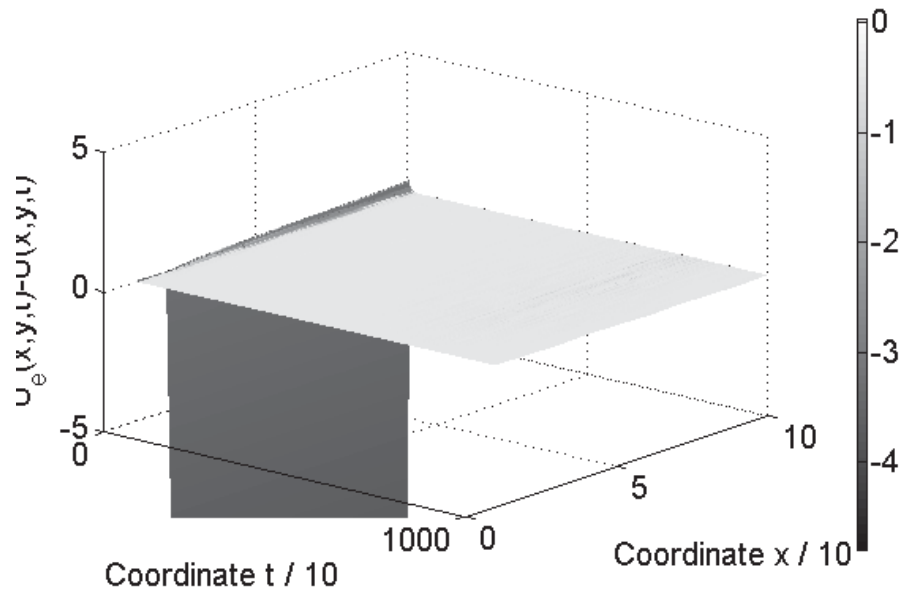


Fig.18 Error del Sistema con GV.

# Capítulo 5

## Conclusiones

En este trabajo se ha demostrado que existen algunos métodos, tales como la RND, para aproximar a un sistema incierto. Esta metodología es útil cuando sólo unos pocos parámetros están disponibles para su estudio. Los parámetros conocidos, como las condiciones de frontera, se utilizan para aproximar el valor del punto vecino. Este proceso se repite varias veces hasta que el centro del espacio de estudio se alcanza. De esta manera la naturaleza del sistema es conocido a pesar de que el modelo no se conoce. Se ha demostrado que las RND son útiles para conocer el comportamiento de un sistema incierto e incluso las anomalías podrían ser descubiertas a través de este método. Debido al hecho de que no todos los modelos matemáticos de enfermedades y patologías se conocen, una aplicación de esta metodología para modelar los nuevos sistemas es una forma útil en áreas de investigación de la medicina que no ha sido explorada. El enfoque propuesto resuelve el problema de la identificación no paramétrica incierta no lineal descrita por las ecuaciones diferenciales parciales hiperbólicas. La convergencia asintótica para la identificación del error se ha demostrado aplicando un análisis en Lyapunov utilizando una clase especial de una función de Lyapunov. Además, el mismo análisis conduce a la generación de las condiciones correspondientes para el límite superior de los pesos implicados en la estructura de la estructura del identificador. La estructura del identificador se basa en el pseudo observador construido mediante la estructura de observador lineal. El diseñar un pseudo observador basado en modos deslizantes de alto orden aplicando el algoritmo super twisting, donde este pseudo observador es realimentado por la identificación del sistema a través de las redes neuronales diferenciales es la aportación principal en este trabajo, y por medio de los ejemplos numéricos se demuestra la viabilidad de esta nueva metodología basada en redes neuronales continuas.

# Bibliografía

- [1] A., P., E., S., AND YU, W. *Differential Neural Networks for Robust Nonlinear Control (Identification, state Estimation an trajectory Tracking)*. Worl Scientific, 2001.
- [2] ADAMS, R., AND FOURNIER, J. *Sobolev Spaces*. 2003.
- [3] A.DERDIYOK, YAN, Z., M.GUVEN, AND UTKIN, V. Sliding mode speed and rotor observer for induction machines in: Industrial electronics society. In *Annual Conference of the IEEE* (2001).
- [4] ALEXANDER, P. *Advanced Mathematical Tools for Automatic Control Engineers. Vol. 1: Deterministic technique*. Elsevier, 2008.
- [5] B, D., A., J., AND A., B. Accuracy analisys for wavelet approximations. *IEEE Transaction on Neural Networks* 6 (1995), 332–348.
- [6] BARBOT, J., DJEMAI, M., AND BOUKHOBZA, T. *Sliding Mode Observers*. Sliding Mode Control in Engineering, 2002.
- [7] BELLMAN, R., AND ADOMIAN, G. *Partial differential equations: new methods for their treatment and solution*. D. Reidel Publishing Company, 1985.
- [8] BENTRCIA, T., AND CHAHDI, M. A low complexity finite-differences neural network combined approach for the modelling of p-n electrical properties. *Journal of Electron Devices* 7 (2010), 203–210.
- [9] BIEY, M., GILLI, M., AND CHECO, P. Complex dynamics phenomena in space-invariant cellular neural networks. *IEEE Transaction on Circuits and System I* 49-3 (2002), 340–345.
- [10] BORUNDIYA, A. *Implementation of Hopfield Neural Network*. VDM Verlag, 2009.

- [11] CHAIREZ, I., FUENTES, R., POZNYAK, A., POZNIAK, T., ESCUDERO, M., AND VIANA, L. Neural network identification of uncertain 2d partial differential equations. In *6th International Conference on Electrical Engineering Computing Science and Automatic Control, CCE* (2009).
- [12] CHEN, F., AND KHALIL, H. Adaptive control of a class of nonlinear discrete-time systems using neural networks. *IEEE Transaction on Automatic Control* *40*(5) (1995), 791–801.
- [13] COOK, R. D. *Finite Element Modeling For Stress Analysis*. 1994.
- [14] COTTERÑ., E. The stone-weierstrass theorem and its application to neural networks. *IEEE Transactions on Neural Network* *1* (1990), 290–295.
- [15] DAVILA, J., F., AND A., L. Second-order sliding-modes observer for mechanical systems. *IEEE Transactions on Automatic Control* *50* (2005), 1785–1789.
- [16] DAVILA, A., AND MORENO, J. Variable gains super-twisting algorithm: A lypaunov based desig. In *American Control Conference* (2010).
- [17] F., L., A., Y., AND K., L. Multilayer neural-net robot controllerwith guaranteed tracking performance. *IEEE Transactions Neural Networks* *7* (2001), 987–1000.
- [18] FILIPPOV, A. *Differential Equations with Discontinuos Right-Hand Side*. 1998.
- [19] FUENTES, R., POZNYAK, A., CHAIREZ, I., AND POZNYAK, T. Neural numerical modeling for uncertain distributed parameter systems. In *International Joint Conference on Neural Networks, pp 909-916* (2009).
- [20] J., D., L., F., AND A., P. Observation and identification of mechanical systems via second order sliding modes. *International Workshop on Variable Structure Systems 9* (June 5-7, 2006), 232–327.
- [21] J., M., AND M., O. Lyapunov approach to second-order sliding mode controllers and observers. In *47th IEEE Conference on Decision and Control* (2008).
- [22] KHALI, H. K. *NonLinear Systems*. Prentice Hall, Upper Saddle River, 1996.
- [23] LEVANT, A. Sliding order and sliding accuracy in sliding mode control. *International Journal of Control* *58*(6) (1993), 1247–1263.



- [24] LEVANT, A. *High Order Sliding Mode Control in Engineering*. 2002.
- [25] LEVANT, A. Finite differences in homogeneous discontinuous control. *IEEE Transaction on Automatic Control* 52(7) (2007), 1208–1217.
- [26] LEVANT, A., AND FRIDMAN, L. *Sliding Mode in Control in Engineering*. 2002.
- [27] LEVEQUE, R. *Finite Differences Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems (Classics in Applied Mathematics)*. 2007.
- [28] LJUNG, L. *System Identification: Theory for the User*. Englewood, 1987.
- [29] NARENDRA, K., AND MUKHOPADHYAY, S. Adaptive control using neural network and approximation models. *IEEE Transaction on Neural Networks* 8(3) (1997), 475–485.
- [30] NELLES, O. *Non linear System Identification*. Springer-Verlag, 2001.
- [31] O., O., AND L., E. D. *Neural systems for control*. Academic Press, 1997.
- [32] PERRUQUETTI, W., AND BARBOT, J. P. *Sliding Mode Control in Engineering*. New York Basel, 2002.
- [33] S., H. *Neural Networks. A comprehensive Foundation*. New York, 1994.
- [34] SAXENA, A., AND SAHAY, B. *Computed Aided Engineering Design*. 2005.
- [35] SMITH, G. D. *Numerical solution of partial differential equations: Infinite difference methods*. Oxford: Clarendon Press,, 1978.
- [36] TAN, C., AND EDWARDS, C. An imi approach for designing sliding mode observers. *International Journal of Control* 74(16) (2001), 1559–1568.
- [37] UTKIN, V. *Sliding Modes in Control and Optimization*. 1992.
- [38] WILFRID, P. *Sliding Mode Control in Engineering*. Marcel Dekker, Inc and New York-Besel, 2002.

# Capítulo 6

## Apéndice

Utilizando la estimación del error dado  $[u_{1,i}(t) - u_{1,i}(t), u_{2,i}(t) - u_{2,i}(t)]$ , que esta descrito por la siguiente dinámica:

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

La convergencia de la identificación del error podrá ser basada en una función especial de Lyapunov [1][22]. Ésta función esta definida como:

$$\begin{aligned} V(t) &= \sum_{i=1}^N V_i(t) \\ V_i(t) &:= \sum_{i=1}^N \|\zeta_i\|_{p_i}^2 + \sum_{j=1}^4 \left[ k_{i,j} \text{tr} \left\{ \tilde{W}_{i,j}^T(t) \tilde{W}_{i,j}(t) \right\} \right] \end{aligned}$$

El vector  $\zeta_i(t)$  se define de la siguiente forma:

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

Siguiendo la idea dada en [16], se puede mostrar facilmente como:

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

Por lo tanto, la derivada del tiempo de cada función específica de Lyapunov tiene la siguiente estructura:

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

Por sustitución directa en el primer término del lado derecho, se tiene:

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

donde:

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

Se puede ver que:

$$\begin{aligned} \zeta_i^T(t) P_i \begin{bmatrix} \frac{1}{2} |\tilde{u}_{1,i}(t)|^{-1/2} [\tilde{u}_{2,i}(t) - \beta_1 \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t))] \\ \tilde{u}_{2,i}(t) - \beta_1 \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) \\ G_i(t) - \beta_2 \text{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix} &= \\ \frac{1}{2} |\tilde{u}_{1,i}(t)|^{-1/2} \zeta_i^T(t) P_i \begin{bmatrix} -\beta_1 & 0 & 1 \\ 0 & 0 & 0 \\ -r & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) \\ \tilde{u}_{1,i}(t) \\ \tilde{u}_{2,i}(t) \end{bmatrix} &+ \\ \zeta_i^T(t) P_i \begin{bmatrix} 0 \\ \tilde{u}_{2,i}(t) - \beta_1 \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) \\ G_i(t) - \beta_2 \text{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix} & \end{aligned}$$

Los términos anteriores de la ecuación previa puede ser rescrita como:

$$\begin{aligned} \zeta_i^T(t) P_i \begin{bmatrix} 0 \\ 0 \\ G_i(t) - \beta_2 \text{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix} &:= \\ \zeta_i^T(t) P_i M [G_i(t) - \beta_2 \text{sign}(\tilde{u}_{1,i}(t))] & \end{aligned}$$

Para analizar el término anterior, utilizando la desigualdad matricial:

$$X^T Y + Y^T X \leq X^T \Lambda X + Y^T \Lambda Y$$

donde:

$$\begin{aligned} X, Y &\in \mathfrak{R}^{n \times m} \\ \Lambda &\in \mathfrak{R}^{n \times n}, \Lambda^T = \Lambda > 0 \end{aligned}$$

Aplicando de manera directa la desigualdad, se tiene:

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

Utilizando esta información, se consigue por la inclusión el término:

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

Por la suposición en la existencia de la solución definida positiva de (3.43), esta desigualdad es intercambiada por:

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

donde  $H_i$  satisface la igualdad dada en (3.45). Finalmente, si se utiliza las leyes de ajuste en (3.40) se tiene:

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

La desigualdad anterior puede demostrarse siguiendo el procedimiento dado en [16] y considerando dentro que  $|\tilde{u}_{1,i}(t)|^{1/2} \leq \lambda_{\max} \zeta_i^T(t) \{P_i\} \sqrt{V_t}$ . Donde la solución de la ecuación diferencial:

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

esta dada por:

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

partiendo del principio de comparación de que  $V_i(t) \leq v(t)$  por lo que  $V_i(0) \leq v(0)$ . La prueba se da por concluida.

# Capítulo 7

## Trabajo a Futuro

1. Implementar este método en otros tipos de redes como son las memorias asociativas.
2. Desarrollar este método en el área aeroespacial, y hacer comparaciones respecto a lo ya existente.
3. Implementar este método en la industria metal-mecánica a través de microcontroladores para lograr estabilidad en los sistemas.
4. Probar otras clases de estructuras por modos deslizantes.
5. Desarrollar algoritmos de observación plenos basados en mediciones puntuales para sistemas hiperbólicos.

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

SALVADOR JUÁREZ LOPEZ<sup>1</sup>, OSCAR CAMACHO NIETO<sup>1</sup>  
AND JORGE ISAAC CHAIREZ ORIA<sup>2</sup>

<sup>1</sup>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

<sup>2</sup>Bioprocess Department  
Professional Interdisciplinary Unit of Biotechnology  
National Polytechnic Institute  
Av. Acueducto s/n, Barrio La Laguna, Col. Ticomán, México, D.F.  
jchaires@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 phenomena 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

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

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:

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

The function  $v(x, t) \in \mathfrak{R}$  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

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

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

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

( $1 \leq p < \infty$ ,  $D^\alpha \nu(x) := \frac{\partial^\alpha}{\partial x^\alpha} \nu(x)$ ). 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

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

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

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

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:

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

with error de order  $h^2$ .

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

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

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

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

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

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

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

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

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

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

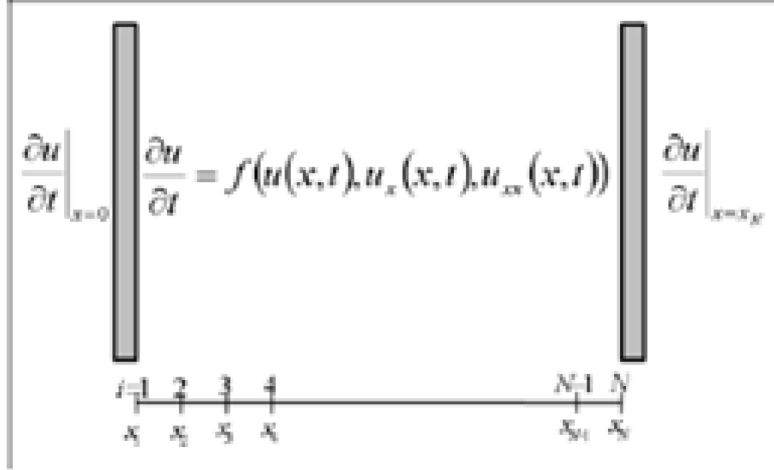


FIGURE 1. Discretization of the spacial domain. This scheme is used to solve the finite difference method to approximate the hyperbolic partial difference equation used in this paper.

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

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

where

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

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-differences representation, each subsystem at the point  $i$  may be represented as

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

That may be represented as follows

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

where  $\mu_i(t) = [u_{1,i}(t), u_{2,i}(t)]^\top \in \mathfrak{R}^2$  is the state vector and  $v_i(t) \in \mathfrak{R}^1$  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

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

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

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

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

**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

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

The last decomposition is based on the approximation capabilities of neural networks. Here it should be noticed that  $f(\cdot, \cdot) \in \mathfrak{R}^n$  always could be presented (by the Stone-Weisstrass and the Kolmogorov theorems [5]) as the composition of nominal  $f_0(x_t, u_t | \Omega) : \mathfrak{R}^{n+m} \rightarrow \mathfrak{R}^n$  and a modeling error  $\tilde{f}_t : \mathfrak{R}^n \rightarrow \mathfrak{R}^n$  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 ( $\Psi_{ij}$ ) as

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

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

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  $\theta_{ij}^*$  are known. So, one must propose the function approximation  $\hat{h}_0(z, \theta)$  designed to be adjusted using a time varying structure.

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

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

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 Weierstrass 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  $\epsilon$  there are some constants  $M_1, M_2, N_1, N_2 \in \mathbb{Z}$  such that for all  $z \in Z \subset \mathfrak{R}$ .

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

**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

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

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:

$$\begin{aligned}\|\bar{\sigma}(x) - \bar{\sigma}(x')\| &\leq L_{\bar{\sigma}} \|x - x'\| & \|\bar{\varphi}(x) - \bar{\varphi}(x')\| &\leq L_{\bar{\varphi}} \|x - x'\| \\ \|\bar{\gamma}(x) - \bar{\gamma}(x')\| &\leq L_{\bar{\gamma}} \|x - x'\| & \|\bar{\eta}(x) - \bar{\eta}(x')\| &\leq L_{\bar{\eta}} \|x - x'\|\end{aligned}$$

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

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

**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

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

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

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

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

**Assumption:** The modelling error is absolutely bounded in  $\Omega$  :

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

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

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

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

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

These new variables are defined as

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

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

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

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

$$\begin{aligned} \bar{f}_{0,i}(x(t), t) &:= \Omega_i^\Gamma(t) \bar{\Pi}_i(z) \\ \bar{\Theta}_i^\Gamma(t) &:= [ A_i \quad V_1(x, t)|_{x=x_i} \quad V_2(x, t)|_{x=x_i} \quad V_3(x, t)|_{x=x_i} \quad V_4(x, t)|_{x=x_i} ] \\ \bar{\Pi}_i^\Gamma(t) &:= \begin{bmatrix} \mu(x, t)|_{x=x_i} & \bar{\sigma}(x)\mu(x, t)|_{x=x_i} \\ \bar{\varphi}(x)\mu_x(x, t)|_{x=x_i} & \bar{\gamma}(x)\mu_{xx}(x, t)|_{x=x_i} & \bar{\eta}(x)v(x, t)|_{x=x_i} \end{bmatrix} \end{aligned} \quad (18)$$

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

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

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

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

These functions satisfy the following sector conditions

$$\begin{aligned} |S_r(x_t^1) - S_r(x_t^2)|^2 &\leq l_{S_r} \|x_t^1 - x_t^2\|^2 \\ x_t^1, x_t^2 &\in \mathbb{R}^l, \quad l \geq 1 \quad l_{S_r} \in \mathbb{R}^+ \end{aligned}$$

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

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

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

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

These new variables are defined as

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

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:

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

**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  $f_i^+ \in \mathfrak{R}^+$  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  $f_0(\cdot, \cdot | \cdot)$  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 inconveniencias 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:

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

Here, the weights matrices  $(\Omega_i)$  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  $\Omega_i$  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:

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

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  $i = 0, \dots, N$  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  $i$ .

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

Let the weights matrices satisfy the following nonlinear matrix differential equations

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

where

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

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

$$\text{Ric}(P_i) := P_i M A_i + [M A_i]^\top P_i + P_i R_i P_i + Q_i^P = 0 \quad (23)$$

where

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

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*

$$P A + A^\top P + P R P + Q = 0$$

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  $(A, R^{1/2})$  is controllable; 3) Pair  $(Q^{1/2}, A)$  is observable, and 4) Matrices  $(A, Q, R)$  should be selected in such a way to satisfy the following inequality

$$\frac{1}{4} (A^\top R^{-1} - R^{-1} A) R (A^\top R^{-1} - R^{-1} A)^\top + Q \leq A^\top R^{-1} A$$

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  $P^i$  ( $i = 0, \dots, N$ ) to the Riccati Equation (23), and for any positive scalar  $d \in \mathfrak{R}^+$  such that the following LMI has solution for some  $H_i > 0$ ,  $H_i^\top = H_i \in \mathbb{R}^{n \times n}$*

$$\left( P_i \begin{bmatrix} -\beta_1 & 0 & 1 \\ 0 & -d & 0 \\ -r_1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -\beta_1 & 0 & 1 \\ 0 & -d & 0 \\ -r_1 & 0 & 0 \end{bmatrix}^\top P_i \right) \leq H_i \quad (24)$$

then the origin

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

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

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

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

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

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:

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ -88.51 & -95.62 \end{bmatrix}, \quad P = \begin{bmatrix} 0.004 & 0 \\ 0 & 0.0028 \end{bmatrix} \\ S &= \begin{bmatrix} 5^{-7} & 0 \\ 0 & 5^{-7} \end{bmatrix}, \quad T = \begin{bmatrix} 2^{-7} & 0 \\ 0 & 2^{-7} \end{bmatrix} \\ W_{1,i}(0) &= W_{2,i}(0) = W_{3,i}(0) = \begin{bmatrix} 31.78 \\ 39.2 \\ 23.8 \end{bmatrix} \end{aligned}$$

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

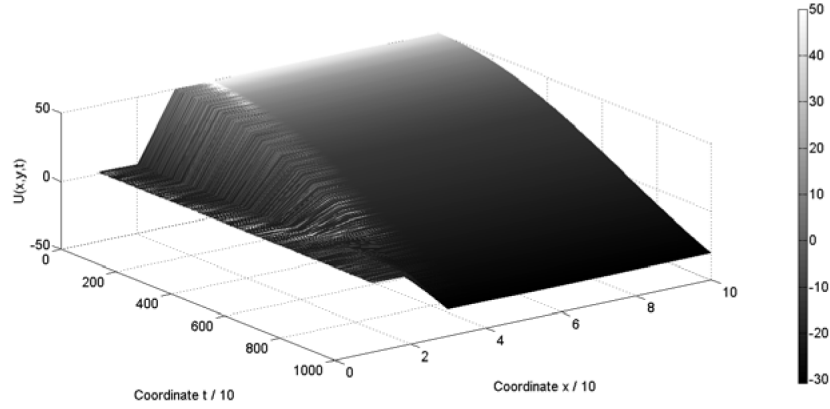


FIGURE 2. Trajectories of the hyperbolic partial differential equation obtained by the finite difference method. These trajectories were generated using the parameters defined in the first numerical problem without perturbations.

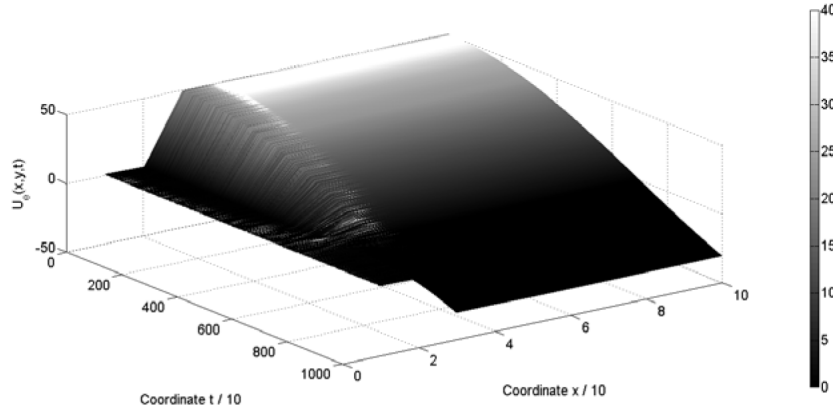


FIGURE 3. Trajectories of the DNN identifier obtained by the finite difference method. These trajectories were generated using the parameters defined in the identifier considering the problem.

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

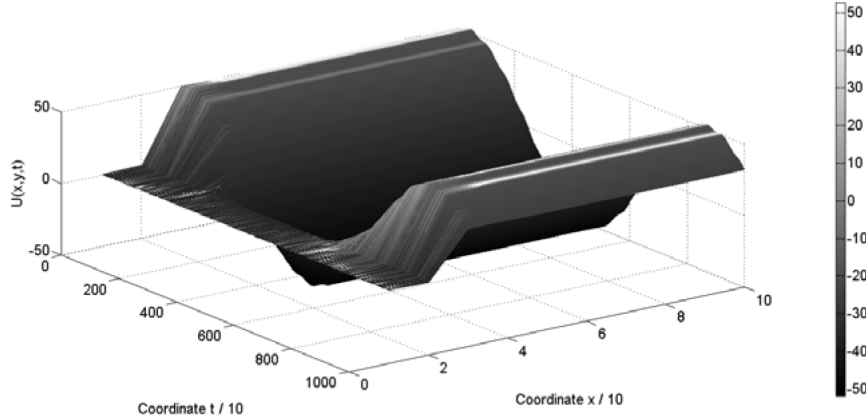


FIGURE 4. Trajectories of the hyperbolic partial differential equation obtained by the finite difference method. These trajectories were generated using the parameters defined in the second numerical problem with perturbations.

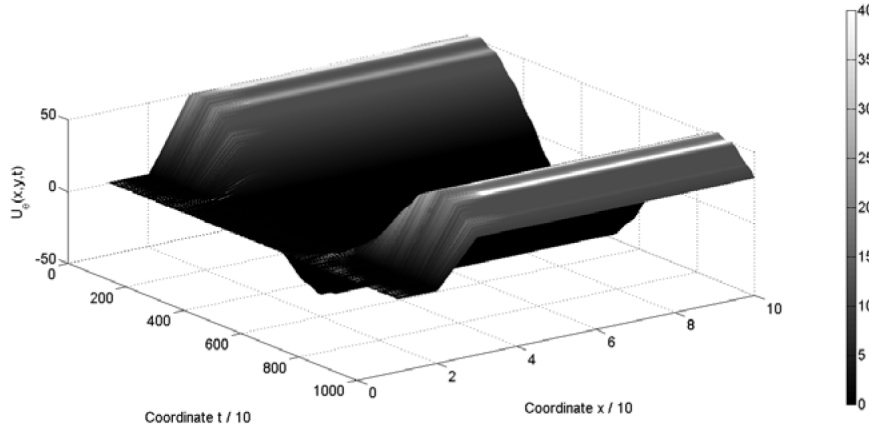


FIGURE 5. Trajectories of the DNN identifier obtained by the finite difference method. These trajectories were generated using the parameters defined in the identifier considering the problem with perturbations.

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. Bentrucia, 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 analysis 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 controller with 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 Heidelberg-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. Sanchez 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

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

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

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



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

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

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

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

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

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

where

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

One can see that

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

Here, the following identity has been used

$$\begin{aligned} \zeta_i(t) &:= N(\zeta_i(t) + C^\top \tilde{u}_{1,i}(t)) \\ N &= (I + C^\top C)^{-1}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{aligned}$$

The last term in the previous equations can be rearranged as

$$\zeta_i^\top(t) P_i \begin{bmatrix} 0 \\ 0 \\ G_i(t) - \beta_2 \text{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix} := \zeta_i^\top(t) P_i M [G_i(t) - \beta_2 \text{sign}(\tilde{u}_{1,i}(t))]$$

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

$$X^T Y + Y^T X \leq X^T \Lambda X + Y^T \Lambda Y$$

where

$$\begin{aligned} X, Y &\in \mathfrak{R}^{n \times m} \\ \Lambda &\in \mathfrak{R}^{n \times n}, \quad \Lambda^T = \Lambda > 0 \end{aligned}$$

By direct application of this inequality, one has

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

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

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

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

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

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

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

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

differential equation

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

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.

*Head note that is usually deleted*

## Adaptive Gain Identifier for Uncertain Hyperbolic Partial Differential Equations Based on Differential Neural Networks and Second Order Sliding Mode Algorithm

S. Juárez<sup>1</sup>, O. Camacho<sup>1</sup> and I. Chairez<sup>2</sup>

<sup>1</sup> Investigation Centre on Computation, National Polytechnic Institute

<sup>2</sup> Professional Interdisciplinary Unit of Biotechnology, National Polytechnic Institute

The date of receipt and acceptance will be inserted by the editor

**Abstract** There are many examples in science and engineering that may be described by a set of partial differential equations (PDE's). 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 (DNN) 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 representation are used to demonstrate the identification abilities achieved by the DNN designed in this paper..

---

*Send offprint requests to:* Offprints Assistant

*Correspondence to:* Address for offprint requests

**Key words** Adaptive Identifier, Variable Gain, Uncertain Hyperbolic Partial Differential Equation, Differential Neural Network, Super-Twisting Algorithm

## 1 Introduction

Partial differential equations (PDEs) are the most common method by which one can model physical problems in engineering. For example, many PDE describing flexible mechanical systems are hyperbolic. This is the main reason because the study of hyperbolic equations is of substantial contemporary attraction. A hyperbolic partial differential equation of order  $n$  is a PDE that, roughly speaking, has a well-posed initial value problem for the first  $n - 1$  derivatives. More precisely, the Cauchy problem can be locally solved for arbitrary initial data along any non-characteristic hypersurface. Interestingly, just knowing the type of PDE can give us insight into how smooth the solution is, how fast information propagates, and the effect of initial and boundary conditions. In this paper, the class of Hyperbolic PDE that will be considered is

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

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:

$$u_x(0, t) = 0, \quad u(0, t) = u_0, \quad u(x, 0) = c \quad (2)$$

The function  $v(x, t) \in \mathfrak{R}$  can be considered like the measurable external perturbation or a designed distributed control action. In (1) the following notation was used

$$u_x(x, t) = \frac{\partial u(x, t)}{\partial x}, \quad u_{xx}(x, t) = \frac{\partial^2 u(x, t)}{\partial x^2}$$

$$u_t(x, t) := \frac{\partial u(x, t)}{\partial t}, \quad u_{tt}(x, t) := \frac{\partial^2 u(x, t)}{\partial t^2}$$

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

The hyperbolic PDE can be represented by (using the state variable theory)

$$z_{1,t}(x, t) = z_2(x, t)$$

$$z_{2,t}(x, t) = f(z_1(x, t), z_{1,x}(x, t), z_{1,xx}(x, t), v(x, t)) \quad (3)$$

where  $z_1(x, t) := u(x, t)$ . Under the following variable definition  $z(x, t) = [z_1(x, t) \ z_2(x, t)]$ , the PDE (1) may be transformed to

$$z_t(x, t) = g(z(x, t), z_x(x, t), z_{xx}(x, t), v(x, t))$$

$$g(z(x, t), z_x(x, t), z_{xx}(x, t), v(x, t)) := \begin{bmatrix} z_2(x, t) \\ f(z(x, t), z_x(x, t), z_{xx}(x, t), v(x, t)) \end{bmatrix}$$

with adequate initial and boundary conditions. The PDE representation given in (3) explain why the Hyperbolic PDEs have been used to represent a lot of mechanical systems with distributed parameters. Actually, modelling methods used to obtain the hyperbolic PDE representing real mechanical systems can provide remarkable and accurate mathematical representation of reality. Nevertheless, a high degree of uncertainties may appear considering the big number of assumptions required to construct the aforementioned model. The class of uncertain nonlinear system that will be analyzed throughout this paper is described by the following structure

$$\begin{aligned} z_{1,t}(x,t) &= z_2(x,t) \\ z_{2,t}(x,t) &= f(z_1(x,t), z_{1,x}(x,t), z_{1,xx}(x,t), v(x,t)) + h(x,t) \end{aligned} \quad (4)$$

where the perturbations  $h(x,t)$  fulfills the following restriction  $\|h(x,t)\|^2 \leq \alpha_1 + \alpha_2 |z_1(x,t)|^2 + \alpha_3 |z_2(x,t)|^2$ , where  $\alpha_1, \alpha_2, \alpha_3 \in \mathfrak{R}^+$ . This last representation is actually a more general class of systems compared to those presented in (3).

## 2 Modelling second order uncertain systems using Neural Networks

A different option to classical model methods based on modelling rules comes from the so-called identification theory. The modern theory of identification regards to solve problem of determining a mathematical model for an unknown system by observing the input-output data generally refers to the adaptive system modelling. The purpose of system identification, among other things, is to predict the behavior of a system under analysis. Usually, the so-called least mean square technique has been used to approximate simple systems represented by ordinary differential equations as equal as PDE with complex structure. However, this method can be used just for very specific cases. A lot of variations have been presented to include nonlinear forms that can manage different approximations. Neural networks is one of the most active areas to model such class of complex nonlinear systems. 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 on its parameters or structure [1], [2].

Using 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 to reproduce the behavior of the uncertain system.

The selection for the ANN structure defines how 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 (mainly 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 [3].

The good performance of identifiers based on DNN [4] depends on its specific structure and the adaptive laws to adjust its parameters. In this paper, the DNN proposal given in [5], [6] and [7] will be used. These DNN have been successfully used in the identification of these unknown systems due to its massive parallelism, fast adaptation and learning capability. These 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 [8]. This approximation skill may be used to solve the numerical identification of partial differential equations.

*Despite the great capabilities showed by the DNN theory to identify full-state available systems, the class of uncertain PDE (4) has no the whole state available to construct the identifier. Actually, it is clear that  $z_2(x, t)$  is not available directly. The proposal introduced in this paper changes the identification scheme by a class of fake state estimation algorithm. Therefore, the identification of the uncertain Hyperbolic PDE is transformed into an observer design.*

State observers have been studied since the 60's. The work of Kalman and Luenberger solved the problem of state estimation for linear systems based on the property of observability, characterized by the condition of the range of the observability matrix. In regards to nonlinear systems, there is no general solution, as the property of observability of a nonlinear system. Observers also be classified depending on how you calculate the gain of the observer or, on the mathematical and physical principles used. Among others, observers using differential geometry techniques, passivity based observers, using bond-graph technique and sliding mode theories have shown relevant advantages to handle different types of second order systems.

Particularly, second order sliding modes (SOSM) observers have shown outstanding capabilities to reject perturbations, avoid the high frequency oscillations (chattering), etc. 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 [9], [10], 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 subspace 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 [11]. This discontinuous term enables the rejection of external matched disturbances [12].

Sliding modes observers are widely used because can provide finite-time convergence, robustness with respect to perturbations and uncertainties estimation ([13], [14]). A new generation of observers based on the so-called second-order sliding-modes has been recently developed [15]. In [9], 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 [13], 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 [16], 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 [17]. 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 (STA) is implemented adding a proportional term in its structure. This term helps the estimation process. The convergence proofs based on Lyapunov approach also are useful to design more advanced second order sliding modes algorithms. Moreover, the homogeneous nature of the standard STA does not allow to compensate uncertainties/disturbances growing together with the state variables. This means that the standard STA cannot ensure the sliding motions even for systems for which the linear part is not exactly known. That is why it is very important to design non homogeneous extension of the standard STA allowing the exact compensation of the smooth uncertainties/disturbances bounded together with their derivatives by the known functions, which could be grow together with the state, i.e. design the global variable gain STA(VGSTA).

This paper describes the pseudo-observer design for a class of uncertain hyperbolic PDE using a mixed structure considering the contribution given by the DNN theory to approximate gray-box systems and the super-twisting like observer method. Considering that neural network is actually a numerical algorithm, the approximation proposed here will be done following a common way to analyze stabilization, control and identification of PDE. DNN based identifier will done using a space discretization for the PDE domain. Here, several methods can be used such as finite differences, finite element and ortogonal collocation. This will be called mesh-based approximation.

### 3 Mesh-based approximation and finite differences method

This well known method consists of a partial approximation of algebraic expressions involving the values of the dependent variable in a limited number



of selected points. As result of this approach, the partial differential equation describing the problem is replaced by a finite number of algebraic equations, written in terms of the dependent variable values at selected points. These equations are usually linear if the PDE's are also linear. When the PDE is time dependent, the algebraic equations turn out to be ordinary differential equations (ODE) [18].

The method produces the value of the selected points to be converted into the set of ODEs, instead of the continuous spatial distribution of the dependent variable. The system of algebraic or differential equations must be solved and may involve a big number of arithmetic operations.

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

$$\begin{aligned} w(x+h, t) &= w(x, t) + hw_x(x, t) + \frac{1}{2}h^2w_{xx}(x, t) \\ &\quad + \frac{1}{6}h^3w_{xxx}(x, t) + l_1(x, t) \\ w(x-h, t) &= w(x, t) - hw_x(x, t) + \frac{1}{2}h^2w_{xx}(x, t) \\ &\quad - \frac{1}{6}h^3w_{xxx}(x, t) + l_2(x, t) \end{aligned} \quad (5)$$

Here  $l_1(x, t)$  and  $l_2(x, t)$  represent terms containing powers of  $h^4$  or/and higher. Adding these Taylor's expansions,

$$w(x+h, t) - w(x-h, t) = 2w_x(x, t) + h^2w_{xx}(x, t) + l(h)^4 \quad (6)$$

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

$$w_{xx}(x, t)_{t \text{ const}} \simeq \frac{1}{h^2} \{w(x+h, t)_{t \text{ const}} - 2w(x, t) + w(x-h, t)_{t \text{ const}}\} \quad (7)$$

with error of order  $h^2$ . Subtracting the equation (5) of equation (6), and neglecting the terms of order  $h^3$  is obtained:

$$w_x(x, t)_{t \text{ const}} \simeq \frac{1}{2h} \{w(x+h, t)_{t \text{ const}} - w(x-h, t)_{t \text{ const}}\} \quad (8)$$

with error de order  $h^2$ . Equation (8) approximates the slope of the tangent at the point  $x$ . 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  $(x, t, w(x, t))_{t \text{ const}}$  by the slope of the line through the points  $(x, t, w(x, t))$  and  $(x+h, t, w(x+h, t))$ , obtaining the backward difference approximation

$$w_x(x, t)_{t \text{ const}} \simeq \frac{1}{h} \{w(x, t) - w(x-h, t)\} \quad (9)$$

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

$$w_x(x, t)_{t \text{ const}} \simeq \frac{1}{h} \{w(x-h, t) - w(x, t)\} \quad (10)$$

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

$$\begin{aligned} w_i(t) &:= w(x_i, t), & w(x, t) &= \left. \frac{dw(x, t)}{dt} \right|_{x=x_i} \\ w_{i,tt}(x, t) &= \left. \frac{d^2w(x, t)}{dt^2} \right|_{x=x_i} \\ w_{i,x}(x, t) &= \left. \frac{\partial w(x, t)}{\partial x} \right|_{x=x_i}, & w_{i,xx}(x, t) &= \left. \frac{\partial^2 w(x, t)}{\partial x^2} \right|_{x=x_i} \end{aligned}$$

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

$$w_{i,x}(x, t) \approx \frac{w(x_i, t) - w(x_{i-1}, t)}{\Delta x}, \quad w_{i,xx}(x, t) \approx \frac{w_{i,x}(x, t) - w_{i-1,x}(x, t)}{\Delta x}$$

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

$$\begin{aligned} w_{i,t}(t) = \dot{w}_i(t) &= \Theta_i(w_i(t), w_{i-1}(t), w_{i-2}(t), v_i(x, t)) \\ &+ \tilde{f}(w_i(t), w_{i-1}(t), w_{i-2}(t), v_i(x, t)) \end{aligned}$$

Where

$$f_i \left( w(x_i, t), \frac{w(x_i, t) - w(x_{i-1}, t)}{\Delta x}, \frac{w_{i,x}(x, t) - w_{i-1,x}(x, t)}{\Delta x}, v_i(x, t) \right)$$

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 approximation and finite differences for uncertain PDE

Following the ideas presented above about the neural network approximation, by simple adding and subtracting the necessary terms, one can represent the hyperbolic PDE (1) as

$$\begin{aligned} u_{tt}(x, t) &= \bar{A}u(x, t) + \hat{V}_1 \bar{\sigma}(x)u(x, t) + \hat{V}_2 \bar{\varphi}(x)u_x(x, t) + \\ &\hat{V}_3 \bar{\gamma}(x)u_{xx}(x, t) + \hat{V}_4 \bar{\eta}(x)v(x, t) + \hat{f}(x, t) \end{aligned} \quad (11)$$

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

$$\begin{aligned} \|\bar{\sigma}(x) - \bar{\sigma}(x')\| &\leq L_{\bar{\sigma}}\|x - x'\| & \|\bar{\varphi}(x) - \bar{\varphi}(x')\| &\leq L_{\bar{\varphi}}\|x - x'\| \\ \|\bar{\gamma}(x) - \bar{\gamma}(x')\| &\leq L_{\bar{\gamma}}\|x - x'\| & \|\bar{\eta}(x) - \bar{\eta}(x')\| &\leq L_{\bar{\eta}}\|x - x'\| \end{aligned}$$

Also it can be shown they are bounded for all  $x \in [0, 1]$ , i.e.,

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

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 computationally 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 simply and efficiently way [19].

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 based 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.

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

$$u_{i,tt}(t) = f_i(x, v, t)_{x=x_i} = f_{0,i}(x, v, t)_{x=x_i} + \tilde{f}_i(x, t)_{x=x_i}, \quad i \in [1, N] \quad (12)$$

where  $u_{i,tt}(t) := u_{tt}(x, t)_{x=x_i}$ . The last decomposition is based on the approximation capabilities of neural networks. Here it should be noticed that  $f_i(\cdot, \cdot) \in \mathfrak{R}^n$  always could be presented (by the Stone-Weisstrass and the Kolmogorov theorems [20]) as the composition of nominal  $f_{0,i}(x, u | \Omega) : \mathfrak{R}^{n+m} \rightarrow \mathfrak{R}^n$  and a modeling error  $\tilde{f}_i : \mathfrak{R}^n \rightarrow \mathfrak{R}^n$  term (as is usual when a model-free approximation is applied). The nominal part  $f_{0,i}$  will be approximated using a nonlinear description based on neural network theory [21], [22] using general basis to reproduce the assumed unknown nonlinear function. Many possible sophisticated suggestions can be made here to design many 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  $f_i$ .

Here the term  $f_{0,i}(t)$  is usually referred to as the nominal dynamics or the DNN approximation. This structure obeys the basic regressor form described in [23]. Therefore, the  $f_{0,i}(t)$  has the form

$$\begin{aligned} f_{0,i}(t) := & \bar{A}u(x, t) \Big|_{x=x_i} + \hat{V}_1 \bar{\sigma}(x)u(x, t) \Big|_{x=x_i} + \hat{V}_2 \bar{\varphi}(x)u_x(x, t) \Big|_{x=x_i} \\ & + \hat{V}_3 \bar{\gamma}(x)u_{xx}(x, t) \Big|_{x=x_i} + \hat{V}_4 \bar{\eta}(x)v(x, t) \Big|_{x=x_i} \end{aligned}$$

On the other hand, the term  $\tilde{f}_i(t) := \tilde{f}(x, t) \Big|_{x=x_i}$  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

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

In this equation  $V_{1,i} \in R^{n_1}, V_{2,i} \in R^{n_2}, V_{3,i} \in R^{n_3}, V_{4,i} \in R^{n_4}$ . Hereafter, it will be assumed that the modeling error terms satisfy the followings assumptions:

**Assumption:** The modelling error is absolutely bounded in  $[0, 1]$  :

$$\left\| \tilde{f}_i(t) \right\|^2 \leq \tilde{f}_{1,i} \|u_i(t)\|^2 \quad (13)$$

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

$$\begin{aligned} f_{0,i}(t) &:= \bar{A}_i u_i(t) + \\ &\left[ \hat{V}_1 \bar{\sigma}(x_i) + (\Delta x)^{-1} \hat{V}_2 \bar{\varphi}(x_i) + (\Delta x)^{-2} \hat{V}_3 \bar{\gamma}(x_i) \right] u_i(t) \\ &- \left[ (\Delta x)^{-1} \hat{V}_2 \bar{\varphi}(x_i) + 2(\Delta x)^{-2} \hat{V}_3 \bar{\gamma}(x_i) \right] u_{i-1}(t) \\ &+ \left[ (\Delta x)^{-2} \hat{V}_3 \bar{\gamma}(x_i) \right] u_{i-2}(x, t) + \hat{V}_4 \bar{\eta}(x_i) v(x_i, t) \end{aligned}$$

Indeed by the approximation theory introduced above, one may represent this expression as  $f_{0,i}(t) := \Omega_{0,i}^\top(t) \Pi_i(t)$ . These new variables are defined as

$$\begin{aligned} \Omega_{0,i}^\top &:= [\bar{A}_i \ \hat{W}_1 \ \hat{W}_2 \ \hat{W}_3 \ \hat{W}_4] \\ \Xi_i^\top(u, t) &:= [u_i(t) \ \sigma(x_i) u_i(t) \ \varphi(x_i) u_{i-1}(t) \ \gamma(x_i) u_{i-2}(t) \ \eta(x_i) v(x_i, t)] \end{aligned}$$

where

$$\begin{aligned} \hat{W}_1 \sigma(x_i) &:= \left[ \hat{V}_{1,i} \bar{\sigma}(x_i) + (\Delta x)^{-1} \hat{V}_2 \bar{\varphi}(x_i) + (\Delta x)^{-2} \hat{V}_3 \bar{\gamma}(x_i) \right] \\ \hat{W}_2 \varphi(x_i) &:= - \left[ (\Delta x)^{-1} \hat{V}_2 \bar{\varphi}(x_i) + 2(\Delta x)^{-2} \hat{V}_3 \bar{\gamma}(x_i) \right] \\ \hat{W}_3 \gamma(x_i) &:= \left[ (\Delta x)^{-2} \hat{V}_3 \bar{\gamma}(x_i) \right] \\ \hat{W}_4 \eta(x_i) &:= \hat{V}_4 \bar{\eta}(x_i) \end{aligned}$$

Here  $\sigma(x_i)$ ,  $\varphi(x_i)$ ,  $\gamma(x_i)$  and  $\eta(x_i)$  are sigmoid functions fulfilling the same conditions asked previously.

This is the so-called finite difference DNN approximation of the uncertain hyperbolic partial differential equation. Considering the structure given in (12) and using the DNN description, finite differences expression of PDE may be approximated by the following second order DNN identifier

$$\begin{aligned} \dot{u}_{1,i}(t) &= u_{2,i}(t) \\ \dot{u}_{2,i}(t) &= \Omega_{0,i}^\top(t) \Xi_i(u, t) + \tilde{f}_i(x, t) \end{aligned} \quad (14)$$

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

## 5 Identification of hyperbolic equations with DNN and STA

Differential neural observers have been studied using the sliding mode approach to obtain algebraic learning procedures for online identification of nonlinear plant (design model) where the state is fully available. 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 are predefined.

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. Some attractive features of second order sliding mode (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, etc. In most cases, sliding modes are obtained by the injection of a non-linear discontinuous term, depending on 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. This discontinuous term enables the rejection of external matched disturbances. Another important contribution in the field of the stability analysis for SOSM observers was made by [17]. 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 (SOSML).

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

$$\begin{aligned}
\frac{d}{dt}\hat{u}_{1,i}(t) &:= \hat{u}_{2,i}(t) + \beta_{1,i}(t)\lambda(\tilde{u}_{1,i}(t))\text{sign}(\tilde{u}_{1,i}(t)) \\
\frac{d}{dt}\hat{u}_{2,i}(t) &:= \Omega_i^T(t)\Xi_i^T(\hat{u}, t) + \beta_{2,i}(t)\text{sign}(\tilde{u}_{1,i}(t)) \\
\tilde{u}_{1,i}(t) &:= u_{1,i}(t) - \hat{u}_{1,i}(t) \\
\Omega_i^T(t) &:= [A_i \ W_{1,i}(t) \ W_{2,i}(t) \ W_{3,i}(t) \ W_{4,i}(t)] \\
\Xi_i^T(\hat{u}, t) &:= [\hat{u}_i(t) \ \sigma(x_i)\hat{u}_i(t) \ \varphi(x_i)\hat{u}_{i-1}(t) \ \gamma(x_i)\hat{u}_{i-2}(t) \ \eta(x_i)v(x_i, t)]
\end{aligned} \tag{15}$$

The solution of equation (15) 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, where we are using

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

### 5.1 Learning laws for Identifier Weights

For each  $i = 0, \dots, N$ , lets define the vector-functions defining the error between the trajectories produced by the model and the DNN-identifier as well as their derivatives for each.

$$\tilde{u}_{1,i}(t) := \hat{u}_i(t) - u_i(t)$$

Let the weights matrices satisfy the following nonlinear matrix differential equations

$$\dot{W}_{k,i} := \Phi_{k,i}(\tilde{u}_{1,i}(t), \tilde{W}_{k,i}(t)), \quad k = \overline{[1, 4]} \tag{16}$$

where

$$\begin{aligned} \tilde{\Phi}_{k,i}(\tilde{u}_{1,i}(t), \tilde{W}_{k,i}(t)) &:= -k_{k,i} \Sigma_3^\top C^\top M_2 \zeta(t) \Xi_{k,i}^\top(\hat{u}, t) - \\ &- k_{k,i} \Sigma_3^\top A_c \Sigma_3 \tilde{W}_{k,i}(x, t) \Xi_{k,i}(\hat{u}, t) \Xi_{k,i}^\top(\hat{u}, t) - \alpha \tilde{W}_{k,i}(x, t) \end{aligned}$$

$$\begin{aligned} \Sigma_3 &= NP_i M_3, \quad M_2 := [0 \ 1 \ 0]^\top, \quad M_3 := [0 \ 0 \ 1]^\top \\ \Xi_{1,i}(t) &:= \sigma(x_i) \hat{u}_i(t), \quad \Xi_{2,i}(t) := \varphi(x_i) \hat{u}_{i-1}(t) \\ \Xi_{3,i}(t) &:= \gamma(x_i) \hat{u}_{i-2}(t), \quad \Xi_{4,i}(t) := \eta(x_i) v(x_i, t) \end{aligned}$$

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

$$Ric(P_i) := P_i \tilde{A}_i + \tilde{A}_i^\top P_i + P_i R_i P_i + Q_i^P = 0 \quad (17)$$

where

$$\begin{aligned} R_i^P &:= M_3 \tilde{W}_{1,i}(x) A_\alpha^{-1} \tilde{W}_{1,i}^\top M_3^\top + M_3 \tilde{W}_{2,i}(x) A_\beta^{-1} \tilde{W}_{2,i}^\top M_3^\top + \\ &M_3 \tilde{W}_{3,i}(x) A_\gamma^{-1} \tilde{W}_{3,i}^\top M_3^\top + M_3 \tilde{W}_{4,i}(x) A_\delta^{-1} \tilde{W}_{4,i}^\top M_3^\top + \frac{1}{2} N A_5 N^\top \\ &\quad + M_3 A_f M_3^\top + \lambda \max\{A_f\} f_{1,i} \\ Q_i^P &:= \alpha I_{n \times n} + \alpha_v I_{n \times n} + d^2 O^\top I_{n \times n} O \end{aligned}$$

and

$$\alpha := \min_{i=1:N} \left\{ \frac{1}{2} \alpha_{\min} \left\{ P_i^{-1/2} H_i P_i^{-1/2} \right\} \right\}$$

The learning laws (16) have been obtained using the Lyapunov methodology as it will be explained in the main theorem. 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.

**Theorem 1** *Consider the non linear model*

$$u_{tt}(x, t) = f(u(x, t), u_x(x, t), u_{xx}(x, t), v(x, t))$$

given by the system of PDE's with uncertainties (perturbations) in the states and the outputs, under the border conditions  $u_x(0, t) = 0$ ,  $u(0, t) = u_0$ ,  $u(x, 0) = c$ ,  $u_0, c \in \mathfrak{R}$ . Let also suppose that parameters in the DNN-identifier given by (15) are adjusted by the learning laws (16). If there exist positive definite matrices  $Q_i^P$  providing the existence of positive solutions  $P^i (\forall i = 1, \dots, N)$  to the Riccati equations (17), and for any positive scalar  $d \in \mathfrak{R}^+$  such that the following set of LMI have positive solutions

$$\left( P_i \begin{bmatrix} -\beta_{1,i}(t) & 0 & 1 \\ 0 & -d_1 & 0 \\ 0 & 0 & -d_2 \end{bmatrix} + \begin{bmatrix} -\beta_{1,i}(t) & 0 & 1 \\ 0 & -d_1 & 0 \\ 0 & 0 & -d_2 \end{bmatrix}^\top P_i \right) \leq H_i \quad (18)$$

and the observer gain are adjusted according to the following algorithms,

$$\begin{aligned}\dot{\beta}_{1,i}(t) &:= \frac{1}{2r_1} \zeta_i^\top(t) P_i M_2 \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) - \alpha \beta_{1,i}(t) \\ \dot{\beta}_{2,i}(t) &:= \frac{1}{2r_2} \zeta_i^\top(t) P_i M_3 \text{sign}(\tilde{u}_{1,i}(t)) - \alpha \beta_{2,i}(t)\end{aligned}$$

then the origin  $\hat{u}_i(t) - u_i(t) = 0$ ,  $\tilde{W}_{k,i}(t) = 0$  is an equilibrium point that is strongly asymptotically stable.

*Proof* This proof is based on the Lyapunov stability theory. The Lyapunov function candidate is

$$\begin{aligned}V(t) &= \sum_{i=1}^N \|\zeta_i(t)\|_{P_i}^2 + \sum_{i=1}^N \sum_{j=1}^4 \left[ k_{i,j} \text{tr} \left\{ \tilde{W}_{j,i}^T(t) \tilde{W}_{j,i}(t) \right\} \right] \\ &\quad + \sum_{i=1}^N \sum_{s=1}^2 r_{s,i} \beta_{s,i}^2(t), \quad s = 1, 2\end{aligned}$$

Here, the vector  $\zeta_i(t)$  is defined as follows

$$\zeta_i(t) := \left[ |\tilde{u}_{1,i}(t)|^{1/2} \text{sign}(\tilde{u}_{1,i}(t)) \tilde{u}_{1,i}(t) \tilde{u}_{2,i}(t) \right]^\top$$

In [17], 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.

Whereas Lyapunov approach

$$\begin{aligned}V(t) &= \sum_{i=1}^N \|\zeta_i(t)\|_{P_i}^2 + \sum_{j=1}^4 \left[ k_{i,j} \text{tr} \left\{ \tilde{W}_{j,i}^T(t), \tilde{W}_{j,i}(t) \right\} \right] + \sum_{i=1}^N \sum_{s=1}^2 r_{s,i} \beta_{s,i}^2(t) \\ \frac{d}{dt} \tilde{u}_{1,i}(t) &= \tilde{u}_{2,i}(t) - \beta_{1,i}(t) \lambda[\tilde{u}_{1,i}(t)] \text{sign}(\tilde{u}_{1,i}(t)) \\ \frac{d}{dt} \tilde{u}_{2,i}(t) &= f_{0,i}(t) - \bar{f}_{0,i}(t) + \tilde{f}_i(t) - \beta_{2,i}(t) \text{sign}(u_{1,i}(t) - \hat{u}_{1,i}(t))\end{aligned}$$

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

$$\dot{V}_i(t) = 2\zeta_i^T(t) P_i \dot{\zeta}_i(t) + \sum_{j=1}^4 k_{i,j} \text{tr} \left\{ \dot{W}_{j,i}^T(t), \tilde{W}_{j,i}(t) \right\} + 2 \sum_{i=1}^N \sum_{s=1}^2 r_{s,i} \beta_{s,i}(t) \dot{\beta}_{s,i}(t) \quad (19)$$

Following the ideas given in [17], 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

$$\begin{aligned} \dot{V}_i(t) &= 2\zeta_i^\top(t) P_i \dot{\zeta}_i(t) + \sum_{j=1}^4 \left[ k_{i,j} \text{tr} \left\{ \dot{W}_{j,i}^T(t) \tilde{W}_{j,i}(t) \right\} \right] \\ &\quad + 2 \sum_{i=1}^N \sum_{s=1}^2 r_{s,i} \beta_{s,i}(t) \dot{\beta}_{s,i}(t) \end{aligned}$$

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

$$2\zeta_i^\top(t) P_i \dot{\zeta}_i(t) := \zeta_i^\top(t) P_i \begin{bmatrix} \frac{1}{2} |\tilde{u}_{1,i}(t)|^{-1/2} [\tilde{u}_{2,i}(t) - \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t))] \\ \tilde{u}_{2,i}(t) - \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) \\ G_i(t) - \beta_{2,i}(t) \text{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix}$$

where

$$\begin{aligned} G_i(t) &= \tilde{\Omega}_i^\top(t) \tilde{\Pi}_i(t) - \Omega_{0,i}^\top(t) \Pi_i(t) := \tilde{A}_i \tilde{u}_i(t) + \\ &\quad \tilde{W}_{1,i} \sigma(x_i) \tilde{u}_i(t) + \tilde{W}_{1,i}(t) \sigma(x_i) \hat{u}_i(t) + \\ &\quad \tilde{W}_{2,i} \varphi(x_i) \tilde{u}_{i-1}(t) + \tilde{W}_{2,i}(t) \varphi(x_i) \hat{u}_{i-1}(t) + \\ &\quad \tilde{W}_{3,i} \gamma(x_i) \tilde{u}_{i-2}(t) + \tilde{W}_{3,i}(t) \gamma(x_i) \hat{u}_{i-2}(t) + \\ &\quad \tilde{W}_{4,i} \eta(x_i) v(x_i, t) + \tilde{W}_{4,i}(t) \eta(x_i) v(x_i, t) - \tilde{f}_i(t) \end{aligned}$$

One can see that

$$\begin{aligned} \zeta_i^\top(t) P_i \begin{bmatrix} \frac{1}{2} |\tilde{u}_{1,i}(t)|^{-1/2} [\tilde{u}_{2,i}(t) - \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t))] \\ \tilde{u}_{2,i}(t) - \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) \\ G_i(t) - \beta_{2,i}(t) \text{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix} &= \\ \frac{1}{2} |\tilde{u}_{1,i}(t)|^{-1/2} \zeta_i^\top(t) P_i \begin{bmatrix} -\beta_{1,i}(t) & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) \\ \tilde{u}_{1,i}(t) \\ \tilde{u}_{2,i}(t) \end{bmatrix} &+ \\ \zeta_i^\top(t) P_i \begin{bmatrix} 0 \\ \tilde{u}_{2,i}(t) - \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) \\ G_i(t) - \beta_{2,i}(t) \text{sign}(\tilde{u}_{1,i}(t)) \end{bmatrix} & \end{aligned} \quad (20)$$

Here, the following identity will be used

$$\begin{aligned} \zeta_i(t) &:= N(\zeta_i(t) + C^\top \tilde{u}_{1,i}(t)) \\ N &= (I + C^\top C)^{-1}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{aligned}$$

The last term in the previous equation will be analyzed in two parts. The first one is

$$\begin{aligned} \zeta_i^\top(t) P_i M_2 [\tilde{u}_{2,i}(t) - \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t))] &= \\ \zeta_i^\top(t) P_i M_2 \tilde{u}_{2,i}(t) - \zeta_i^\top(t) P_i M_2 \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) &= \\ \frac{1}{2} \zeta_i^\top(t) [P_i \tilde{A} + \tilde{A}^\top P_i] \zeta_i(t) - & \\ [N(\zeta_i(t) + C^\top \tilde{u}_{1,i}(t))]^\top P_i M_2 \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) & \end{aligned}$$

where  $\tilde{A} := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ . To analyze the other term, lets use the matrix inequality [24]  $X^\top Y + Y^\top X \leq X^\top \Lambda X + Y^\top \Lambda Y$  where  $X, Y \in \mathfrak{R}^{n \times m}$ ,  $\Lambda \in \mathfrak{R}^{n \times n}$ ,  $\Lambda^\top = \Lambda > 0$ . By direct application of this inequality, one has

$$\begin{aligned} & \zeta_i^\top(t) P_i M_3 [G_i(t) - \beta_{2,i} \text{sign}(\tilde{u}_{1,i}(t))] \leq \\ & \zeta_i^\top(t) P_i M_3 A_i \tilde{u}_i(t) + \tilde{u}_i^\top(t) A_i^\top M_3^\top P_i \zeta_i(t) + \\ & \zeta_i^\top(t) P_i M_3 \tilde{W}_{1,i} A_1^{-1} \tilde{W}_{1,i}^\top M_3^\top P_i \zeta_i(t) + \tilde{u}_i^\top(t) \sigma^\top(x_i) \Lambda_1 \sigma(x_i) \tilde{u}_i(t) + \\ & \zeta_i^\top(t) P_i M_3 \tilde{W}_{2,i} A_2^{-1} \tilde{W}_{2,i}^\top M_3^\top P_i \zeta_i(t) + \tilde{u}_{i-1}^\top(t) \varphi^\top(x_i) \Lambda_2 \varphi(x_i) \tilde{u}_{i-1}(t) + \\ & \zeta_i^\top(t) P_i M_3 \tilde{W}_{3,i} A_3^{-1} \tilde{W}_{3,i}^\top M_3^\top P_i \zeta_i(t) + \tilde{u}_{i-2}^\top(t) \gamma^\top(x_i) \Lambda_3 \gamma(x_i) \tilde{u}_{i-2}(t) + \\ & \zeta_i^\top(t) P_i M_3 \tilde{W}_{4,i} A_4^{-1} \tilde{W}_{4,i}^\top M_3^\top P_i \zeta_i(t) + v^\top(x_i, t) \eta^\top(x_i) \Lambda_4 \eta(x_i) v(x_i, t) + \\ & \zeta_i^\top(t) P_i M_3 A_f M_3^\top P_i \zeta_i(t) + \tilde{f}_i^\top(t) A_f^{-1} \tilde{f}_i(t) \\ & - [N(\zeta_i(t) + C^\top \tilde{u}_{1,i}(t))]^\top P_i M_3 \beta_2 \text{sign}(\tilde{u}_{1,i}(t)) + \\ & [N(\zeta_i(t) + C^\top \tilde{u}_{1,i}(t))]^\top P_i M_3 \tilde{W}_{1,i}(t) \sigma(x_i) \tilde{u}_i(t) + \\ & [N(\zeta_i(t) + C^\top \tilde{u}_{1,i}(t))]^\top P_i M_3 \tilde{W}_{2,i}(t) \varphi(x_i) \tilde{u}_{i-1}(t) + \\ & [N(\zeta_i(t) + C^\top \tilde{u}_{1,i}(t))]^\top P_i M_3 \tilde{W}_{3,i}(t) \gamma(x_i) \tilde{u}_{i-2}(t) + \\ & [N(\zeta_i(t) + C^\top \tilde{u}_{1,i}(t))]^\top P_i M_3 \tilde{W}_{4,i}(t) \eta(x_i) v(x_i, t) \end{aligned}$$

The integration of all results obtained previously leads and using this information, one gets by the simple inclusion of term  $\frac{1}{2} d_1 |\tilde{u}_{1,i}(t)|^{-1/2} \zeta_i^\top(t) P_i N \tilde{u}_{1,i}(t)$  +  $\frac{1}{2} d_2 |\tilde{u}_{1,i}(t)|^{-1/2} \zeta_i^\top(t) P_i N \tilde{u}_{2,i}(t)$  and the application of the Lambda inequality presented above

$$\begin{aligned} & 2\zeta_i^\top(t) P_i \dot{\zeta}_i(t) \leq \\ & -\frac{1}{2} |\tilde{u}_{1,i}(t)|^{-1/2} \zeta_i^\top(t) \left( P_i \begin{bmatrix} \beta_1(t) & 0 & 1 \\ 0 & d_1 & 0 \\ 0 & 0 & d_2 \end{bmatrix} + \begin{bmatrix} \beta_1(t) & 0 & 1 \\ 0 & d_1 & 0 \\ 0 & 0 & d_2 \end{bmatrix}^\top P_i \right) \zeta_i(t) \\ & + \zeta_i^\top(t) \left( P_i \tilde{A} + \tilde{A}^\top P_i + P_i M_3 A_i + A_i^\top M_3^\top P_i + P_i R_i P_i + Q_i \right) \zeta_i(t) + \\ & \left\| \Lambda_a^{1/2} \Sigma_2 \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) \right\|^2 + \left\| \Lambda_b^{1/2} \Sigma_3 \beta_{2,i} \text{sign}(\tilde{u}_{1,i}(t)) \right\|^2 - \\ & \tilde{u}_{1,i}^\top(t) C \Sigma_2 \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) - \tilde{u}_{1,i}^\top(t) C \Sigma_3 \beta_{2,i} \text{sign}(\tilde{u}_{1,i}(t)) + \\ & \left\| \Lambda_c^{1/2} \Sigma_3 \tilde{W}_{1,i}(t) \sigma(x_i) \tilde{u}_i(t) \right\|^2 + \tilde{u}_{1,i}^\top(t) C \Sigma_3 \tilde{W}_{1,i}(t) \sigma(x_i) \tilde{u}_i(t) + \\ & \left\| \Lambda_d^{1/2} \Sigma_3 \tilde{W}_{2,i}(t) \varphi(x_i) \tilde{u}_{i-1}(t) \right\|^2 + \tilde{u}_{1,i}^\top(t) C \Sigma_3 \tilde{W}_{2,i}(t) \varphi(x_i) \tilde{u}_{i-1}(t) + \\ & \left\| \Lambda_e^{1/2} \Sigma_3 \tilde{W}_{3,i}(t) \gamma(x_i) \tilde{u}_{i-2}(t) \right\|^2 + \tilde{u}_{1,i}^\top(t) C \Sigma_3 \tilde{W}_{3,i}(t) \gamma(x_i) \tilde{u}_{i-2}(t) + \\ & \left\| \Lambda_f^{1/2} \Sigma_3 \tilde{W}_{4,i}(t) \eta(x_i) v(x_i, t) \right\|^2 + \tilde{u}_{1,i}^\top(t) \Sigma_3 M \tilde{W}_{4,i}(t) \eta(x_i) v(x_i, t) \end{aligned}$$

Where  $\Sigma_j = NP_iM_j$  By the assumption on the existence of positive definite solutions of (17), this inequality is changed by

$$\begin{aligned} \dot{V}_i(t) \leq & -\frac{1}{2} |\tilde{u}_{1,i}(t)|^{-1/2} \zeta_i^\top(t) H_i \zeta_i(t) + \\ & \zeta_i^\top(t) \left( P_i \tilde{A} + \tilde{A}^\top P_i + P_i M_3 A_i + A_i^\top M_3^\top P_i + P_i R_i P_i + Q_i \right) \zeta_i(t) + \\ & \left\| \Lambda_a^{1/2} \Sigma_2 \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) \right\|^2 + \left\| \Lambda_b^{1/2} \Sigma_3 \beta_{2,i} \text{sign}(\tilde{u}_{1,i}(t)) \right\|^2 \\ & - \tilde{u}_{1,i}^\top(t) C \Sigma_2 \beta_{1,i}(t) \lambda(\tilde{u}_{1,i}(t)) \text{sign}(\tilde{u}_{1,i}(t)) - \tilde{u}_{1,i}^\top(t) C \Sigma_3 \beta_{2,i} \text{sign}(\tilde{u}_{1,i}(t)) + \\ & \left\| \Lambda_c^{1/2} \Sigma_3 \tilde{W}_{1,i}(x, t) \sigma(x_i) \hat{u}_i(t) \right\|^2 + \tilde{u}_{1,i}^\top(t) C \Sigma_3 \tilde{W}_{i,1}(x, t) \sigma(x_i) \hat{u}_i(t) + \\ & \left\| \Lambda_d^{1/2} \Sigma_3 \tilde{W}_{2,i}(t) \varphi(x_i) \hat{u}_{i-1}(t) \right\|^2 + \tilde{u}_{1,i}^\top(t) C \Sigma_3 \tilde{W}_{2,i}(t) \varphi(x_i) \hat{u}_{i-1}(t) + \\ & \left\| \Lambda_e^{1/2} \Sigma_3 \tilde{W}_{3,i}(t) \gamma(x_i) \hat{u}_{i-2}(t) \right\|^2 + \tilde{u}_{1,i}^\top(t) C \Sigma_3 \tilde{W}_{3,i}(x, t) \gamma(x_i) \hat{u}_{i-2}(t) + \\ & \left\| \Lambda_f^{1/2} \Sigma_3 \tilde{W}_{4,i}(t) \eta(x_i) v(x_i, t) \right\|^2 + \tilde{u}_{1,i}^\top(t) C \Sigma_3 \tilde{W}_{i,4}(x, t) \eta(x_i) v(x_i, t) \\ & + \sum_{j=1}^4 \sum_{i=1}^N \left[ k_{i,j} \text{tr} \left\{ \dot{W}_{j,i}^\top(t) \tilde{W}_{j,i}(t) \right\} \right] + 2 \sum_{s=1}^2 r_s \beta_{s,i}(t) \dot{\beta}_{s,i}(t) \end{aligned}$$

where  $H_i$  satisfy the inequality given in (18). Finally, if using the adjustment laws given in (16) one has

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

This last inequality can demonstrated following the procedure given in [17] and taking into account that  $|\tilde{u}_{1,i}(t)|^{1/2} \leq \|\zeta_i(t)\| \leq \lambda_{\max} \{P_i\} \sqrt{V_i(t)}$ . 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.

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 researchers 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 pseudo-observer constructed within the linear observers framework.

*Remark 1* The set of learning laws proposed in (16) can not be solved indeed. These equations demand the knowledge of  $\tilde{W}_{i,j}(x)$ ,  $j = 1..4$  that are actually unknown. Therefore, the learning laws can not be used to solve the approximation of the uncertain hyperbolic system. Nevertheless, a training algorithm can be used to overcome this problem. In this paper, the training follows the method proposed in [25]. The aforementioned method will produce a set of values  $\tilde{W}_{k,i}^{id}$  fulfilling the following condition (with  $0 < \mathcal{Y} < \infty$ )

$$\left\| \Delta \tilde{W}_{k,i}^{id} \right\|^2 := \left\| \tilde{W}_{k,i}^{id} - \tilde{W}_{k,i} \right\|^2 \leq \mathcal{Y}$$

if the state trajectories  $u_i(t)$  are bounded. Several methods have been proposed to solve this issue. Among others, matrix least mean square with several modifications have been used to solve that training algorithm. Moreover, nonlinear least mean square method based on a class of neural network also has been introduced.

Therefore, the learning laws can be seen as

$$\dot{\tilde{W}}_{k,i}(t) := \Phi_{k,i}(\tilde{u}_{1,i}(t), \tilde{W}_{k,i}^{id}(t)) + \Phi_{k,i}(\tilde{u}_{1,i}(t), \Delta \tilde{W}_{k,i}^{id})$$

This change is used to see how the training process can affect the quality of identification based on DNN for hyperbolic PDE. The utilization of this process is obligatory because no knowledge on  $\tilde{W}_{k,i}$  is assumed. Based on the result presented in the main theorem of this paper, the identification error  $\|\hat{u}_i(t) - u_i(t)\|$  is bounded too. Therefore, the identifier state is also bounded by the assumption presented in this remark. Using the conditions established in this part, changing  $\tilde{W}_{k,i}^{id}$  instead of  $\tilde{W}_{k,i}$  will produce a bigger but bounded deviation of the identifier state compared to the trajectories of the uncertain system. Therefore, the quality of the identification process will be reduced but the decrement will be measurable and bounded. Finally, the learning laws are become into

$$\begin{aligned} \Phi_{k,i}(\tilde{u}_{1,i}(t), \tilde{W}_{k,i}^{id}(t)) &:= -k_{k,i} \Sigma_3^T C^T M_2 \zeta(t) \Xi_{k,i}^T(\hat{u}, t) - \\ &k_{k,i} \Sigma_3^T \Lambda_c \Sigma_3 \tilde{W}_{k,i}^{id}(x, t) \Xi_{k,i}(\hat{u}, t) \Xi_{k,i}^T(\hat{u}, t) - \alpha \tilde{W}_{k,i}^{id}(x, t) \end{aligned}$$

These new versions for the learning laws are numerically solvable now. These solutions may be implemented to achieve the approximation of the uncertain nonlinear system representing the hyperbolic PDE.

The following section shows a numerical example where the identifier for hyperbolic PDE has been developed using DNN as main tool.

## 6 Simulation Results

This section shows the results for the identification algorithm tested in a benchmark hyperbolic PDE system. This system fulfills the necessary conditions required to achieving the results for the identification using the designed pseudo-sliding mode observer based on the DNN methodology.

Below, the numerical simulations show the qualitative illustration for a benchmark system. Let us consider the simplified problem of a vibrating string with continuous stopping dynamics. 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

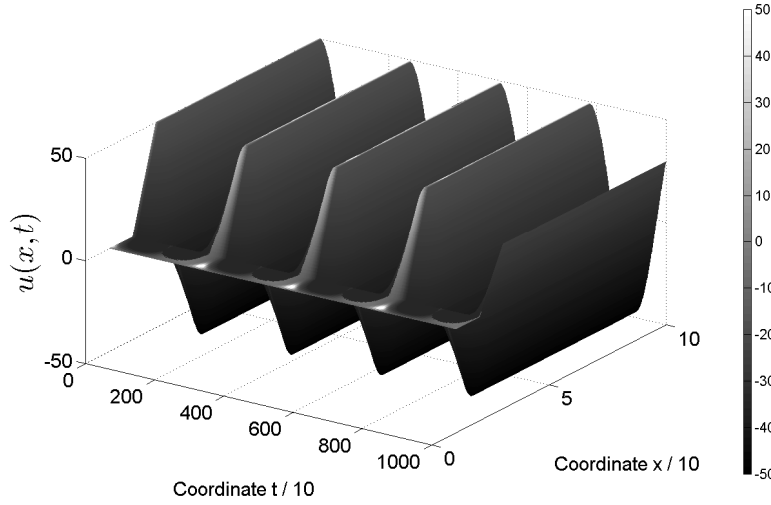
$$\begin{aligned} u_{tt}(x, t) &= 0.01u_{xx}(x, t) + 0.01u_x(x, t) + \\ &0.03u(x, t) + 0.01 \sin(0.1u(x, t)) \sin(u(x, t)) \\ u_x(0, t) &= 0, \quad u_x(1, t) = 0 \end{aligned} \quad (21)$$

This model assumes that you have access to discrete measures of the state  $u(x, t)$  along the entire spacial 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:

$$\begin{aligned} A &= \begin{bmatrix} 0 & 1 \\ -88.51 & -95.62 \end{bmatrix}, \quad P = \begin{bmatrix} 0.004 & 0 \\ 0 & 0.0028 \end{bmatrix} \\ S &= \begin{bmatrix} 5^{-7} & 0 \\ 0 & 5^{-7} \end{bmatrix}, \quad T = \begin{bmatrix} 2^{-7} & 0 \\ 0 & 2^{-7} \end{bmatrix} \\ W_{1,i}(0) &= W_{2,i}(0) = W_{3,i}(0) = W_{4,i}(0) = \begin{bmatrix} 31.78 \\ 39.2 \\ 23.8 \end{bmatrix} \end{aligned}$$

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 of the external perturbation  $A \sin(x, t)$  and the natural oscillating trajectory associated to this system. The solution of this PDE represented by (21) was obtained using a simple finite differences based program. This benchmark problem was selected because the high oscillations showed by the numerical solution (1). This model represents a good challenge for the identifier based on DNN.

DNN identifier generates the trajectories which are very close to the real trajectories of the benchmark algorithm (see Fig. 2). The numerical solution to produce the identifier trajectories is the same used to obtain the solution for the benchmark problem. Therefore, these both solution can be compared. The similitude between both figures (1 and 2) may demonstrate the



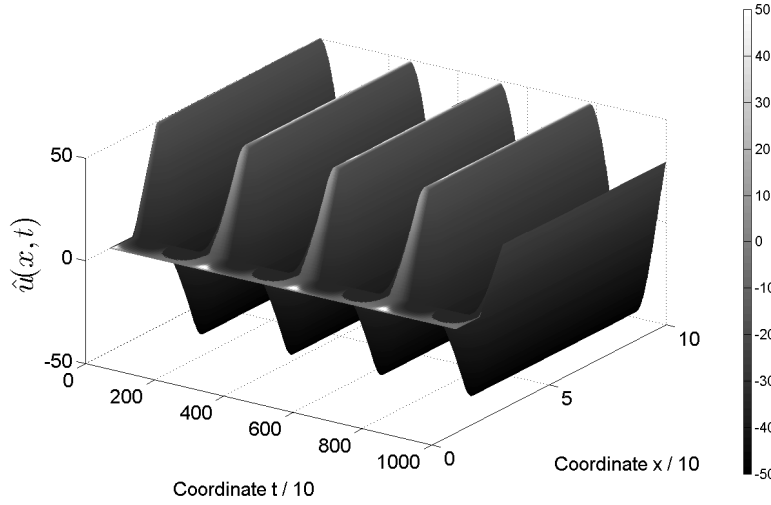
**Fig. 1** Simulation solution produced by the hyperbolic partial differential equation representing the general wave problem. This solution was obtained using a special software specially designed for this non parametric approximation problem.

effectiveness of the identification algorithm. The error between trajectories produced by the model and the proposed identifier is close to zero almost for all  $x$  and all  $t$  that shows the efficiency of the identification process provided by the suggested DNN algorithm (Fig. 3). The figure shows the evolution of the error between both trajectories with respect to time, that is, the figure 3 plots the following vector

$$Z = \log \left( \sqrt{\sum_{i=1}^N [u(x, t) |_{x=x_i} - \hat{u}(x, t) |_{x=x_i}]^2} \right)$$

that serves as a measure of the approximation quality achieved by the DNN identifier using variable gains. 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.

The benefits achieved by the variable gain in the identifier can be shown by a simple comparison between the measure demonstrated in the previous figure. The following figure demonstrates this comparison when the identifier uses the variable gain (dashed line) and fixed constant gains (solid line) (Fig. 4). This plot uses a y-semilog axis to find a better way to remark the benefits achieved by using variable gain in the identifier. One can see the similarities between both trajectories: the model and the identifier. Even when they are really similar, the grayscale defining the trajectories



**Fig. 2** Simulation solution produced by the non parametric identifier based on continuous neural networks to approximate a class of hyperbolic partial differential equations. This solution was obtained using a special software specially designed for this non parametric approximation problem.

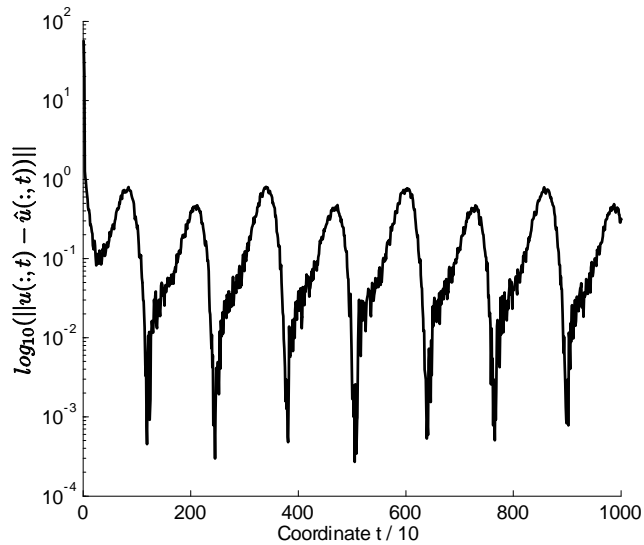
demonstrates small variations which are natural consequences of the adaptive scheme used by the identifier to produce its trajectory. The previous figure shows a validate comparison between the identification achieved when fixed and variables gains are used. An additional way to measure what the variable gains benefits are, comes from the well know square mean error defined by

$$\bar{Z}(t) = \int_{\tau=0}^t \left[ \log \left( \sqrt{\sum_{i=1}^N [u(x, \tau) |_{x=x_i} - \hat{u}(x, \tau) |_{x=x_i}]^2} \right) \right] d\tau$$

Here, the contribution of using variable gains is remarked here. The following figure (5) demonstrates how the variable gain improves the measure introduced in the previous equation. The dashed line shows the measure proposed before when variable gain is used versus fixed gain represented with the solid line.

## 7 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



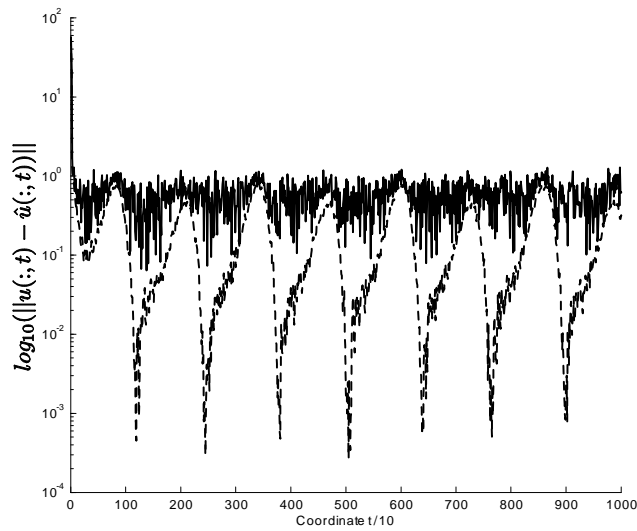
**Fig. 3** Evolution of the error between both trajectories with respect to time using the following measure:  $Z = \log \left( \sqrt{\sum_{i=1}^N [u(x, t)|_{x=x_i} - \hat{u}(x, t)|_{x=x_i}]^2} \right)$

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 pseudo-observer constructed within the linear observers framework. Numerical example showing the beam dynamics demonstrates the workability of this new methodology based on continuous neural networks.

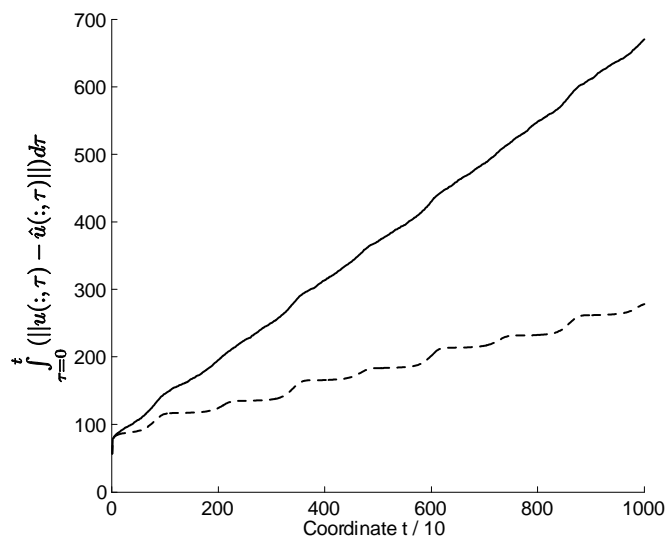
## 8 Acknowledgements

Authors want to acknowledge the support given by the projects IPN-SIP-COFAA-CIC\_20113090, ICYTDF\_243/2010 and 277/2010 and CONACYT\_I0110/127/08 (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.





**Fig. 4** Comparison of the error obtained when the identifier uses the variable gain (dashed line) and fixed constant gains (solid line) based on the measure defined as  $Z$ .



**Fig. 5** Comparison between the identification achieved when fixed and variables gains are used based on the square mean error.

## References

1. O. Nelles, *Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Berlin Heidelberg-New York: Springer-Verlag, 2001.

2. F. Lewis, A. Yesildirek, and K. Liu, "Multilayer neural-net robot controller with guaranteed tracking performance," *IEEE Transactions Neural Networks*, vol. 7, pp. 987–1000, 2001.
3. V. Petridis and S. Petridis, "Construction of neural network based lyapunov functions," in *IJCNN '06. International Joint Conference on Neural Networks*, 2006.
4. A. Poznyak, E. Sánchez, and W. Yu, *Differential Neural Networks for Robust Nonlinear Control (Identification, state Estimation an trajectory Tracking)*. Worl Scientific, 2001.
5. R. Fuentes, A. Poznyak, T. Poznyak, and I. Chairez, "Neural numerical modeling for uncertain distributed parameter systems," in *Proceedings of the International Joint Conference of Neural Networks*,, 2009.
6. R. Fuentes, A. Poznyak, I. Chairez, and T. Poznyak, "Neural numerical modeling for uncertain distributed parameter systems," in *International Joint Conference on Neural Networks*, pp. 909–916, 2009.
7. I. Chairez, R. Fuentes, A. Poznyak, T. Poznyak, M. Escudero, and L. Viana, "Neural network identification of uncertain 2d partial differential equations," in *6th International Conference on Electrical Engineering, Computing Science and Automatic Control, CCE*, 2009.
8. M. Biey, M. Gilli, and P. Checco, "Complex dynamic phenomena in space-invariant cellular neural networks," *IEEE Transaction on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, pp. 340–345, March 2002.
9. A. Levant, "Principles of 2-sliding mode designing," *Automatica*, vol. 43, pp. 823–830, 2005.
10. A. Levant and L. Fridman, *Sliding Mode in Control in Engineering*, ch. High Order Sliding Modes, pp. 53–101. Marcel Dekker, Inc., 2002.
11. V. Utkin, *Sliding Modes in Control and Optimization*. Springer-Verlag, 1992.
12. 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.
13. 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, June 5-7 2006.
14. J. Barbot, M. Djemai, and T. Boukhobza, *Sliding Mode observers in Sliding Mode Control in Engineering*. Springer-Verlag, 2002.
15. A. Levant, *High Order Sliding Modes*. Sliding Mode Control in Engineering, Marcel Dekker, Inc., 2002.
16. A. Levant, "Finite differences in homogeneous discontinuous control," *IEEE Transactions on Automatic Control*, vol. 52, no. 7, pp. 1208–1217, 2007.
17. J. Moreno and M. Osorio, "Lyapunov approach to second-order sliding mode controllers and observers," in *47th IEEE Conference on Decision and Control*, 2008.
18. R. Leveque, *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems (Classics in Applied Mathematics)*. SIAM, Society for Industrial and Applied Mathematics, July 2007.
19. T. Bentrícia, Y. Tamrabet, and M. Chahdi, "A low complexity finite-differences neural network combined approach for the modelling of pn junction electrical properties," *Journal of Electron Devices*, vol. 7, pp. 203–210, 2010.

20. N. E. Cotter, "The stone-weierstrass theorem and its application to neural networks," *IEEE Transactions on Neural Network*, vol. 1, pp. 290–295, 1990.
21. S. Haykin, *Neural Networks. A comprehensive Foundation*. New York: IEEE Press, 1994.
22. O. Omidvar and D. L. Elliott, *Neural systems for control*. Academic Press, 1997.
23. L. Ljung, *System Identification: Theory for the User*. Englewood Cliff, New Jersey: Prentice Hall, 1987.
24. A. Poznyak, *Advanced Mathematical Tools for Automatic Control Engineers. Vol. 1: Deterministic techniques*. Elsevier, 2008.
25. I. Chaírez, "Wavelet differential neural network," *IEEE Transactions on Neural Networks*, vol. 20, pp. 1439–1449, 2009.