



UNIVERSIDAD MICHOACANA DE SAN  
NICOLÁS DE HIDALGO

---

---

INSTITUTO DE FÍSICA Y MATEMÁTICAS

**REDES NEURONALES ARTIFICIALES  
APLICADAS A PROBLEMAS INVERSOS EN  
FÍSICA**

**T E S I S**

QUE PARA OBTENER EL GRADO DE:  
DOCTOR EN CIENCIAS EN EL ÁREA DE FÍSICA

PRESENTA:  
MAURICIO CARRILLO VALENCIA



DIRECTOR DE TESIS:  
DR. JOSÉ ANTONIO GONZÁLEZ CERVERA

MORELIA MICHOACÁN

AGOSTO, 2017

UNIVERSIDAD MICHOACANA DE SAN NICOLÁS DE HIDALGO.

INSTITUTO DE FÍSICA Y MATEMÁTICAS

**Redes Neuronales Artificiales  
Aplicadas a Problemas Inversos en Física**

Tesis sometida para recibir el grado de

Doctor en Ciencias en el Área de

Física

por

Mauricio Carrillo Valencia

Comité de evaluación

Dr. José Antonio González Cervera, Asesor

Dr. Francisco Siddhartha Guzmán Murillo

Dr. Umberto Cotti Gollini

Agosto, 2017

Derechos de autor  
Mauricio Carrillo Valencia, Agosto, 2017  
Derechos reservados.

El presente trabajo de tesis doctoral de Mauricio Carrillo Valencia es aprobado, y es aceptable en calidad y forma para su publicación:

---

---

---

Presidente

Instituto de Física y Matemáticas  
Universidad Michoacana de San Nicolás de Hidalgo.

Agosto, 2017



## DEDICATORIA

A mis padres, hermanos y amigos.



## EPÍGRAFO

*El único error verdadero es  
aquel con el que nada se aprende.*

John Powell





## TABLA DE CONTENIDO

	Índice de figuras . . . . .	XI
	Índice de tablas . . . . .	XIV
	Agradecimientos . . . . .	XV
	Vita y Publicaciones . . . . .	XVII
	Resumen . . . . .	XIX
	Abstract . . . . .	XXI
Capítulo 1.	Introducción y Motivación . . . . .	1
	1.1. Problemas inversos . . . . .	1
	1.2. Aprendizaje automático . . . . .	2
	1.3. Redes neuronales artificiales . . . . .	4
Capítulo 2.	Redes Neuronales Artificiales . . . . .	9
	2.1. Redes neuronales biológicas . . . . .	9
	2.2. Redes neuronales artificiales . . . . .	10
	2.3. Redes neuronales multicapa . . . . .	14
	2.3.1. Propagación hacia adelante . . . . .	14
	2.3.2. Funciones de activación . . . . .	16
	2.4. Entrenamiento . . . . .	19
	2.4.1. Error de entrenamiento . . . . .	19
	2.4.2. Backpropagation . . . . .	22
	2.5. Consideraciones y discusiones finales . . . . .	25
Capítulo 3.	Estimación de la Razón de Masas en Ondas Gravitacionales de la Colisión de Agujeros Negros. . . . .	27
	3.1. Ondas gravitacionales . . . . .	27
	3.2. Estimación de razón de masas de sistema binario de agujeros considerando las fases de fusión y decaimiento . . . . .	33
	3.2.1. Descripción y selección de datos para el análisis mediante las RNA . . . . .	34
	3.2.2. Resultados . . . . .	37
	3.3. Estimación de la razón de masas considerando fase orbital y ringdown . . . . .	42
	3.3.1. Estructura y entrenamiento de las RNAs . . . . .	45
	3.3.2. Resultados . . . . .	48
	3.4. Discusión del capítulo . . . . .	49

Capítulo 4.	Uso de RNAs en Problemas Inversos en el Flujo de un Fluido Clásico Alrededor de un Cilindro en un Tubo de Dos Dimensiones . . . . .	55
4.1.	Flujo alrededor de un cilindro en una tubería dos dimensional . . . . .	56
4.1.1.	Método de lattice Boltzmann . . . . .	60
4.2.	Estimación del número de Reynolds en el flujo de un fluido alrededor de un cilindro . . . . .	61
4.2.1.	Procesamiento datos . . . . .	63
4.2.2.	Resultados . . . . .	65
4.3.	Estimación de un obstáculo en el flujo dentro de un conducto de dos dimensiones . . . . .	73
4.3.1.	Metodología . . . . .	75
4.3.2.	Región objetivo . . . . .	78
4.3.3.	Estructura de la RNA . . . . .	78
4.3.4.	Resultados . . . . .	85
4.3.5.	Conclusiones de la sección . . . . .	98
4.4.	Discusión de capítulo . . . . .	98
Capítulo 5.	Conclusiones . . . . .	101
Apéndice A.	Implementación Algoritmo de Backpropagation . . . . .	105
Apéndice B.	Prueba de Código de la RNA . . . . .	109
B.1.	Reconocimiento de dígitos . . . . .	109
B.1.1.	Clasificación con función de costo de error cuadrático medio . . . . .	111
B.1.2.	Clasificación con función de costo de entropía cruzada . . . . .	112
B.1.3.	Resultados . . . . .	113
Apéndice C.	Ondas Gravitacionales . . . . .	117
C.1.	Ondas gravitacionales de sistemas binarios de agujeros negros . . . . .	121
Apéndice D.	Método de Lattice Boltzman . . . . .	123
Glosario	. . . . .	129
Bibliografía	. . . . .	131

## Índice de figuras

Figura 2.1.: Esquema de los valores de entrada y pesos de las conexiones de cada una de ellas. El estímulo que recibe la neurona es la combinación lineal del vector de entrada $\mathbf{x}$ y el vector de parámetros $\mathbf{w}$ . El estado activo o inactivo de la neurona dependerá si la suma de este producto interno es mayor al umbral $U$ . . . . .	12
Figura 2.2.: Esquema de una RNA tipo MLP . . . . .	15
Figura 2.3.: Funciones de activación tipo sigmoide . . . . .	18
Figura 2.4.: Muestra del error cuadrático medio calculado en el entrenamiento y validación . . . . .	22
Figura 2.5.: Ejemplificación de sobreentrenamiento en una RNA . . . . .	23
Figura 3.1.: Fases de una onda gravitacional de una sistema binario de agujeros negros . . . . .	29
Figura 3.2.: Señal generada por la onda gravitacional GW150914 . . . . .	31
Figura 3.3.: Formas de onda simuladas por relatividad numérica . . . . .	34
Figura 3.4.: Error de una RNA en función del número de neuronas ocultas para la estimación de razón de masas en ondas gravitacionales sin ruido . . . . .	37
Figura 3.5.: Predicciones de la razón de masas en señales sin ruido de dos agujeros negros y en función del número de neuronas ocultas . . . . .	38
Figura 3.6.: Error porcentual en la predicción de masas para una RNA con 10 neuronas ocultas en señales de GWs sin ruido . . . . .	40
Figura 3.7.: Parte real del <i>strain</i> de una forma de onda con razón de masas igual a 1 y ruido Gaussiano con $SNR = 10$ . . . . .	41
Figura 3.8.: Error de predicción de la razón de masas para formas de onda con ruido Gaussiano . . . . .	42
Figura 3.9.: Rango de parámetros cubiertos en el catálogo de ondas gravitacionales generado por el grupo de Georgia Tech . . . . .	43
Figura 3.10.: Parte real del <i>strain</i> y sus correspondientes transformadas de Fourier para algunas formas de onda . . . . .	46
Figura 3.11.: Predicciones de las razones de masa de dos agujeros negros con ruido Gaussiano en el análisis de serie de tiempo . . . . .	51
Figura 3.12.: Predicciones de las razones de masa de dos agujeros negros con ruido Gaussiano en el análisis de frecuencias . . . . .	52
Figura 4.1.: Flujos de Poiseuille . . . . .	57
Figura 4.2.: Flujos alrededor de un obstáculo cilíndrico para distintos números de Reynolds . . . . .	59
Figura 4.3.: Mallado de LBM con modelo D2Q9 . . . . .	61
Figura 4.4.: Esquema de detectores donde se extraen los perfiles de $v_x$ y $(\nabla \times \mathbf{v})_z$ . . . . .	63

Figura 4.5.: Perfiles de $v_x$ y $(\nabla \times \mathbf{v})_z$ en distintas ubicaciones del tubo de dos dimensiones . . . . .	64
Figura 4.6.: Perfiles de $v_x$ y $(\nabla \times \mathbf{v})_z$ para $Re = 30$ y $100$ en $x = 0.3m$ . . . . .	68
Figura 4.7.: RMSE en la estimación del número de Reynolds en función del número de sensores y distancia . . . . .	69
Figura 4.8.: Error porcentual en la estimación del número de Reynolds en función del número de sensores y distancia . . . . .	70
Figura 4.9.: Error porcentual de la estimación de $Re = 30$ en distintos lugares de medición . . . . .	71
Figura 4.10.: Error porcentual de la estimación de $Re = 99$ en distintos lugares de medición . . . . .	72
Figura 4.11.: Simulación en LBM del flujo alrededor de un cilindro . . . . .	74
Figura 4.12.: Esquema del problema del flujo alrededor del obstáculo cilíndrico y los sitios de medición utilizados . . . . .	75
Figura 4.13.: Diferencia de IRSL en la malla numérica y la malla objetivo . . . . .	79
Figura 4.14.: Serie de tiempo del flujo en tres sensores en un sitio de medición . . . . .	80
Figura 4.15.: Resultados de $R^2$ en la predicción de la RNA producidos para obstáculos de diferentes diámetros y localizados en el centro de la tubería . . . . .	87
Figura 4.16.: IRSL para un obstáculo objetivo en el centro del tubo y el predicho por la RNA . . . . .	88
Figura 4.17.: Resultados de $R^2$ para obstáculos en el centro, con análisis de serie de tiempo . . . . .	89
Figura 4.18.: Resultados de $R^2$ para obstáculos en el centro, con una sola medición en el tiempo . . . . .	90
Figura 4.19.: IRSL para un obstáculo objetivo en la parte superior de la tubería y el predicho por la RNA . . . . .	91
Figura 4.20.: Resultados de $R^2$ para obstáculos en distintas posiciones en el eje $y$ , en un análisis de serie de tiempo con un solo sensor . . . . .	92
Figura 4.21.: Resultados de $R^2$ para obstáculos en distintas posiciones en el eje $y$ , en un análisis de serie de tiempo con tres sensores . . . . .	93
Figura 4.22.: Simulación de LBM para el flujo alrededor de un obstáculo cuadrado . . . . .	95
Figura 4.23.: IRSL de obstáculo cuadrado y el predicho por la RNA . . . . .	97
Figura B.1.: Representación de los dígitos 0 y 7 extraídos del catálogo del UCI Machine Learning Repository . . . . .	110
Figura B.2.: Muestra del error cuadrático medio en el entrenamiento y validación en el reconocimiento de dígitos . . . . .	114
Figura C.1.: Efecto de una onda gravitacional en un anillo de partículas libres	120
Figura C.2.: Fases de la onda gravitacional GW150914 . . . . .	122
Figura D.1.: Esquema de los enfoques de modelación de dinámica de fluidos	124

Figura D.2.: Representación esquemática de la malla numérica de <i>lattice Boltzmann</i> y modelo D2Q9 . . . . .	126
Figura D.3.: Modelo de <i>bounceback</i> usado como condición de frontera en los métodos de <i>lattice Boltzmann</i> . . . . .	127

## Índice de tablas

Tabla 3.1.: Conjuntos de entrenamiento, validación y predicción con distintas razones de masa entre agujeros negros . . . . .	36
Tabla 3.2.: Predicciones de $q$ en señales sin ruido, considerando RNAs con 10, 50 y 100 neuronas ocultas . . . . .	39
Tabla 3.3.: Conjunto de simulaciones de ondas gravitacionales usadas a partir del catálogo generado por el grupo de Georgia Tech . . . . .	44
Tabla 3.4.: Error y coeficiente de ajuste en la predicción de $q$ en señales con ruido en los análisis de serie de tiempo y frecuencia . . . . .	49
Tabla 3.5.: Error porcentual en predicciones para señales con ruido en el análisis de serie de tiempo . . . . .	50
Tabla 3.6.: Error porcentual en predicciones para señales con ruido en el análisis de frecuencia . . . . .	50
Tabla 4.1.: RMSE en función del número de sensores en distintos lugares de medición utilizando $v_x$ . . . . .	68
Tabla 4.2.: RMSE en función del número de sensores en distintos lugares de medición . . . . .	80
Tabla 4.3.: Comparación de la precisión de $Re$ en distintos sitios de medición de $v_x$ mediante la prueba $\chi^2$ . . . . .	81
Tabla 4.4.: Comparación de la precisión de $Re$ en distintos sitios de medición de $(\nabla \times \mathbf{v})_z$ mediante la prueba $\chi^2$ . . . . .	82
Tabla 4.5.: Parámetros de los casos de estudio en el reconocimiento de un obstáculo cilíndrico . . . . .	83
Tabla 4.6.: Parámetros de los obstáculos con diferentes diámetros, velocidad de flujo entrante y ubicaciones de los sensores . . . . .	83
Tabla 4.7.: Resultados de $R^2$ para distintos obstáculos y velocidades de flujo entrante . . . . .	96
Tabla B.1.: Resultados para el reconocimiento de dígitos con RNAs con función de costo de error cuadrático medio . . . . .	115
Tabla B.2.: Resultados para el reconocimiento de dígitos con RNAs con función de costo de entropía cruzada . . . . .	116

## AGRADECIMIENTOS

Mi más profundo agradecimiento a mi asesor Dr. José Antonio González por su mentoría para realizar este trabajo de investigación, así como por su tiempo, paciencia y dedicación en mi crecimiento personal.

Al Dr. Francisco Siddhartha Guzmán por su apoyo y continua participación, comentarios y consejos para el trabajo de investigación. Agradezco también al Dr. Umberto Cotti y al Dr. Alfredo Raya por su interés y colaboración en el trabajo desarrollado.

A mis compañeros y todo el personal académico y administrativo del Instituto de Física y Matemáticas por ofrecer un ambiente de trabajo agradable. A la Universidad Michoacana de San Nicolás de Hidalgo, por ofrecer espacios y actividades para el desarrollo humano. Al Consejo Nacional de Ciencia y Tecnología por el apoyo otorgado en el programa de Doctorado y a lo largo del desarrollo profesional.





## VITA

- 2008                      Licenciado en Física,  
Facultad de Ciencias,  
Universidad Autónoma del Estado de México.
- 2013                      Maestro en Ciencias en el Área de Física,  
Instituto de Física y Matemáticas,  
Universidad Michoacana de San Nicolás de Hidalgo.

## PUBLICACIONES

- Carrillo, M., González, J.A., Hernández, S., López, C.E. & Raya, A., *Bloch oscillations in two-dimensional crystals: Inverse problem*. Computational Materials Science, 137, (2017).
- Carrillo, M., Que, U., González, J.A. & López, C., *Recognition of an obstacle in a flow using artificial neural networks*. Physical Review E, 96 (2), 2017.
- Carrillo, M., Que, U. & González, J.A. *Estimation of Reynolds number for flows around cylinders with lattice Boltzmann methods and artificial neural networks*. Physical Review E, 94(6), 2016.
- Carrillo, M., Gracia-Linares, González, J.A. & Guzmán., F.S. *Parameter estimates in binary black hole collisions using neural networks*. General Relativity and Gravitation, 48, 2016.
- Agüero, M.A., De Lourdes, M. & Carrillo, M., *Nonclassic solitonic structures in DNA's vibrational dynamics*. International Journal of Modern Physics B, 22(16), 2008.



# Resumen

En la Física experimental se necesita relacionar las observaciones con los parámetros físicos que describen al sistema en la teoría, es decir, obtener un modelo del sistema a partir de los datos. Este planteamiento es conocido como el problema inverso, y aunque existen técnicas analíticas para resolverlo, generalmente se utilizan métodos numéricos para inferir el valor de los parámetros del modelo físico. En este trabajo de tesis doctoral se plantea el uso de redes neuronales artificiales como método de resolución para dos problemas inversos en el área de Física generados mediante simulaciones numéricas.

El primero de ellos, motivado por las recientes detecciones de ondas gravitacionales GW150914, GW151226 y GW170104, se enfoca en estimar la razón de masas iniciales en la colisión de sistemas binarios de agujeros negros, analizando señales de ondas gravitacionales generadas a partir de simulaciones numéricas. El análisis consiste en introducir a una red neuronal artificial, fragmentos de las formas de onda asociados a la fase orbital, fusión y decaimiento del evento.

En el segundo problema utiliza las redes neuronales artificiales para analizar los perfiles de velocidad, vorticidad y presión dinámica en simulaciones de flujos de fluidos alrededor de un cilindro en un tubo de dos dimensiones. Con el reconocimiento de los patrones de los perfiles dichas las variables físicas se busca determinar: el número de Reynolds, que caracteriza al flujo; y estimar el tamaño y posición de la obstrucción dentro del tubo.

Con estas dos ejemplificaciones se muestra el potencial de utilizar los métodos aprendizaje automático en el reconocimiento de patrones en observaciones experimentales o en simulaciones numéricas, para la resolución de problemas inversos, particularmente en el campo de la Física.

Palabras clave: *Redes neuronales artificiales, problema inverso, simulaciones numéricas, ondas gravitacionales, flujo en un tubo de dos dimensiones.*



# Abstract

In experimental Physics, it is needed to relate the observations with the physical parameters that describe the system in the theory, this is, to obtain a model from the data. This approach is known as the inverse problem, and although there are analytical methods to solve it, generally are used numerical methods to infer the parameter values of the model. In this doctoral thesis is proposed the use of artificial neural networks as a method to solve two inverse problems in Physics generated through numerical simulations.

The first problem, motivated by the recent gravitational waves observations GW150914, GW151226 and GW170104, is focused in the estimation of the mass ratio in binary black hole collisions, by analyzing gravitational wave signals generated through numerical simulations. The analysis consists in introducing into an artificial neural network fragments from the wave forms related to the inspiral, merger and ringdown phases of the event.

The second problem uses artificial neural networks to analyze the velocity, vorticity and dynamic pressure profiles in numerical simulations of flows around a cylinder in a two dimensional pipe. Through the pattern recognition of the profiles of those physical variables it is desired to obtain: the Reynolds number, which characterizes the flow; and estimate the size and position of the obstruction inside the pipe.

With this two examples, it is shown the potential to use machine learning methods in pattern recognition for experimental data or numerical simulations, to solve inverse problems, particularly in the field of Physics.

*Keywords: Artificial neural networks, inverse problems, numerical simulations, gravitational waves, flow in a two dimensional pipe.*



# Capítulo 1

## Introducción y Motivación

### 1.1. Problemas inversos

En general, a lo largo del desarrollo académico de una persona en el área de la Física se le plantean problemas en los cuales debe obtener la solución en el tiempo o espacio sobre un sistema físico en el cual se le provee la especificación completa de la estructura interna de este. Sin embargo, en términos experimentales surge la situación contraria: obtener la descripción del sistema a partir de observaciones directas o indirectas sobre este. Este planteamiento se define como resolver el problema inverso. Un ejemplo de esta índole puede ser que aquel donde se desea obtener la distribución de la temperatura en un material en contacto con alguna fuente de calor. De forma teórica, a expensas de la dificultad geométrica y del cálculo involucrado, la solución del problema directo para conocer el campo de temperaturas en algún punto del sistema se puede obtener utilizando la ecuación de distribución de calor con condiciones iniciales o restricciones conocidas. Si al contrario se tienen los valores de la temperatura varios puntos del material, se desea conocer las condiciones del sistema, por ejemplo, la temperatura, el flujo de calor de la fuente o alguna otra propiedad del material como su coeficiente de conducción de calor [Hon & Wei, (2003)]. En general, los problemas inversos son reconocidos como mal planteados en el sentido de Hadarmard [Kaipio & Somersalo, (2005)], sí no cumple al menos una de las siguientes condiciones:



- Tiene solución.
- Existe una única solución
- La solución depende de los datos.

Los problemas inversos son por naturaleza inestables porque las soluciones son desconocidas y los parámetros tienen que ser determinados por medio de observaciones discretas que contienen errores en la medición. Además, la unicidad o existencia de la solución para las observaciones dadas, no está garantizada [Aster, (2003)].

Tradicionalmente el campo de resolución de problemas inversos se ha desarrollado a través de modelos estadísticos clásicos o de regularización como regresiones lineales, multivariantes o modelos generalizados. No obstante, con el avance de los recursos computacionales se ha incrementado el uso de análisis estadísticos en los cuales se pueden obtener distribuciones probabilísticas de dichas variables. Sin embargo, estas técnicas son computacionalmente extenuantes, ya que en algunos casos conlleva resolver el problema directo una gran cantidad de veces. Por lo tanto, resulta necesario encontrar otros métodos alternativos capaces de acelerar el proceso de la estimación de parámetros que puedan competir con los tradicionales en velocidad, robustez y calidad de resultados. Gracias a la conexión que existe entre la teoría de regularización y los algoritmos de aprendizaje automático (conocido también como *machine learning* por su nombre en inglés), es posible explotar esta metodología que resulta intuitiva en la introducción de datos así como la interpretación de los resultados [De Vito, E., et al.,(2005)].

## 1.2. Aprendizaje automático

El aprendizaje automático es una rama de la inteligencia artificial, capaz de aprender y adaptarse a los cambios que suceden en la información que se analiza. El aprendizaje automático significa construir programas con algoritmos capaces de resolver problemas de optimización y toma de decisiones a través del análisis de datos. El aprendizaje automático nació por la necesidad de reconocer patrones en

información donde los humanos no tienen la capacidad suficiente o el número de datos es tan grande que el análisis de estos se vuelve abrumador [Alpaydin, (2010)].

Un algoritmo de aprendizaje automático se modela por ciertos parámetros, y el aprendizaje se logra encontrando los valores de los parámetros del modelo mediante la información proporcionada. Las tareas que se pueden lograr a través del reconocimiento de patrones con estos métodos se pueden delimitar por:

- Clasificación. El algoritmo analiza la información de entrada y categorizarla en función de sus características asociándola a clases definidas previamente. Los algoritmos más representativos de este tipo de tareas son: *k-nearest neighbors*, *naïve Bayes*, máquinas de soporte vectorial (SVM), redes neuronales artificiales (RNAs), entre otros [Bishop, (2006)].
- Regresión. En este enfoque el algoritmo debe procesar los datos y reconocer propiedades en ellos tales que sea capaz de dar una estimación en términos de un valor numérico asociado a la respuesta deseada. En este caso los algoritmos más conocidos son la regresión lineal, logística y RNAs [Bishop, (2006)].

Los algoritmos de aprendizaje automático también se pueden distinguir por su proceso de entrenamiento:

- Aprendizaje supervisado: Este tipo de adiestramiento se basa en proporcionar al modelo un conjunto de ejemplos, preferentemente representativos del problema en cuestión. Estos ejemplos se componen del tipo de información que se quiere analizar y de las correspondientes respuestas que se desea obtener. Con este enfoque, el algoritmo debe comprender la correlación existente entre los dos subconjuntos (datos de entrada y de respuesta), para elaborar una serie de reglas de acción que debe realizar para futuros ejemplos de entrada desconocidos. El entrenamiento requiere que los parámetros que componen al modelo se adapten a los datos ejemplo de un modo iterativo. Esta metodología es la más usual en las redes neuronales artificiales y la mayoría de los algoritmos de aprendizaje automático.

- Aprendizaje no supervisado. A diferencia del anterior, los ejemplos presentados no contienen las respuestas deseadas, únicamente se le proporcionan los datos que se desea analizar, por lo que el algoritmo debe encontrar la relación entre ellas de forma automática, basándose en algunas reglas predefinidas. El resultado deseado son datos que se encuentran agrupados por ciertas características y que difieren en relación a otros grupos. Un ejemplo de este tipo de aprendizaje es por medio del algoritmo de *k-means clustering*, el cual particiona los datos en  $k$  grupos dentro de un determinado radio; donde el número de grupos y la métrica del radio para cada grupo son definidos por el usuario [Bishop, (2006)]. También las RNAs han sido desarrolladas en este tipo de aprendizaje, en los denominados *self-organizing maps* [Kohonen, (1995)].
- Aprendizaje reforzado. A diferencia del entrenamiento supervisado, en este tipo de entrenamiento no se le suministran las respuestas deseadas al algoritmo, sino una serie de reglas a seguir. El algoritmo aprende mediante recompensas cuando realiza acciones positivas que llevan al objetivo final y se le castiga cuando ejecuta acciones negativas. El algoritmo más prominente en este ámbito son nuevamente las RNAs, por ejemplo en el trabajo realizado por [Tesauro, (1995)], planteando un modelo de aprendizaje para una RNA que fue capaz de jugar *backgammon* a un nivel experto.

Una vez entrenado el modelo, este ya ha aprendido las reglas para relacionar los valores de entrada con las respuestas deseadas, por lo que en el futuro se espera que pueda hacer predicciones precisas. Como se ve de los puntos anteriores, las RNAs, están presentes en todas las formas de aprendizaje, acreditando su uso en el presente trabajo.

### 1.3. Redes neuronales artificiales

Las redes neuronales artificiales (RNAs), provienen de la analogía de las redes neuronales biológicas, en el cual un conjunto de elementos interconectados son

capaces de realizar operaciones de forma sincronizada. Su desarrollo fue provocado inicialmente por la descripción del funcionamiento del sistema nervioso por Ramón y Cajal, quien mostró que este se conforma por un ensamble de células llamadas *neuronas*, comunicándose de forma estructurada en un proceso llamado sinapsis [Ramón y Cajal, (1899)]. Con el paso del tiempo y posterior investigación sobre el sistema nervioso y aparición de las máquinas computacionales, se llegó a la pregunta de si este mecanismo era capaz de ser emulado por las máquinas.

Las redes neuronales artificiales suelen ser vistas como aproximadores de funciones de tipo *caja negra*, donde un dato de entrada provee una respuesta sin saber cual fue el mecanismo exacto por el cual la RNA llegó a tal solución. A pesar de ello, los algoritmos desarrollados se han sofisticado de tal manera que en la actualidad las RNAs son utilizadas para el reconocimiento de patrones en datos, clasificación de datos masivos (big data en inglés), categorización de objetos y/o reconocimiento facial en imágenes; operación y optimización de procesos industriales, etc.

Tomando en cuenta que en los problemas inversos el conocimiento de los parámetros del fenómeno físico es parcial o nulo, la búsqueda de ellos puede ser muy extenuante. Por ello, la implementación de métodos como las RNAs es muy valiosa, siendo capaces de extraer información de los datos. Algunas de las aplicaciones de las RNAs se pueden resumir en los campos de:

**Medicina.** Este campo es actualmente muy explotado por el aprendizaje automático, dado que el conocimiento exacto sobre los mecanismos fisiológicos no es totalmente conocido. En este ramo se pueden enlistar aplicaciones en la interpretación de electrocardiogramas, donde los especialistas pueden evaluar la existencia de una condición cardíaca [Bortolan, Degani & Williems, (1991)]. Las RNAs pueden revelar un tumor canceroso o daño de tejido a través de muestras biológicas [Dorsey, et al., (1997)], o incluso interpretar y reconstruir imágenes a partir de las proyecciones de tomografías computarizadas [Kerr & Bartlett, (1995), Knoll et al., (1999)]. Las RNAs son útiles en la lectura de electroencefalogramas y señales biométricas para un posible diagnóstico telemétrico [Begg, Kamruzzaman & Sarker, (2006)]. En las aplicaciones farmacéuticas también se emplean diseñando tratamientos médicos o correcciones en la dosificación de drogas, cantidad de radiación

en pacientes con cáncer en base a sus respuestas eméticas [Martín et al., (2004)], absorción o eliminación de fármacos en la sangre [Chen, Chen, Min, Fisher & Wu, (2003)], entre otros.

**Economía e Industria.** Las instituciones, entidades financieras y corporaciones se han interesado en la investigación y aplicación de RNAs a mercados financieros, los cuales son influenciados por la gran cantidad de datos en continuo cambio. La gente interesada en estas áreas pretenden predecir el comportamiento de los mercados y obtener ventajas para la inversión de opciones financieras, intercambio monetario, comercio de recursos, entre otros [McNelis, (2005)]. Las RNAs en la industria se basan en predicción de costos, capacidad de producción, almacenamiento y suministro de productos, y principalmente en la optimización de procesos [Eldon, (1994)]. Dado el incremento de información que se maneja a través del internet, se están mejorando los sistemas de seguridad con ayuda de RNAs que sean capaces de detectar virus, correo no deseado o programas maliciosos [Tesauro, Kephart & Sorkin, (1996), Tobiyama et al., (2016)].

**Ciencias básicas.** El potencial de las RNAs en diversas áreas de la Física, Química, Biología, Geología, entre otros, se enfoca en fenómenos con presencia de ruido e incertidumbre sobre las variables responsables. Entre las aplicaciones se pueden nombrar: análisis de espectroscopía [Cirovic & Dragan, (1997)], predicción sobre reacciones nucleares [Gernoth, (1999)], pronóstico del número de manchas en el ciclo solar [MacPherson, Conway & Brown, (1995)], y más recientemente en el análisis y/o desarrollo de materiales como el grafeno o termoeléctricos [Khanam et al., (2016), Zhao & Li, (2017)]. Son útiles en temas relacionados con el análisis de series de tiempo, procesamiento de señales y filtración de ruido tal y como se requiere en sistemas dinámicos cuasi-caóticos [Garliauskas, (1998)] y clasificación de fallas en experimentos de detección de ondas gravitacionales [Biswas et al., (2013), Graff, (2012)].

Los puntos mencionados, dan motivos suficientes para explotar el uso de recursos de aprendizaje automático en las distintas ciencias, justificando en particular el trabajo que se desarrolla en esta tesis. Para mostrar el uso y desempeño de las redes neuronales artificiales en algunos problemas inversos de Física, se investigaron dos temas:

1. Estimación de la razón de masas iniciales en la colisión de dos agujeros negros sin espín.
2. Estimación del parámetro que caracteriza el flujo de un fluido en un tubo de dos dimensiones y del tamaño y posición del obstáculo.

El primer punto ha sido motivado dada las recientes observaciones de ondas gravitacionales GW150914 [Abbot et al., (2016a)], GW151226 [Abbot et al., (2016b)] y GW170104 [Abbot et al., (2017)]. Por lo que es de suma importancia resolver el problema inverso y extraer la información que contiene alguna señal de una onda gravitacional, para poder inferir las propiedades físicas y parámetros sobre las fuentes que generaron dicho evento. En el caso particular de este trabajo, se han construido RNAs en las cuales la información a analizar corresponde a la serie de tiempo o el espectro de frecuencias en un fragmento de una simulación de una onda gravitacional ocasionada por la colisión de dos agujeros negros. Mientras que la respuesta esperada de la RNA es la razón de masas iniciales entre ambos agujeros negros. Las ondas gravitacionales utilizadas para este estudio fueron extraídas a partir de dos catálogos diferente y creados por medio de simulaciones numéricas [Jani et al., (2016), Gracia, (2017)].

En el segundo tema se han propuesto las RNAs para resolver de forma independiente dos problemas inversos y relacionados al flujo de un fluido clásico alrededor de un obstáculo cilíndrico en un tubo de dos dimensiones espaciales. En el primero de ellos se caracteriza el flujo del fluido por medio de la estimación del número de Reynolds asociado a dicho flujo. Mientras que en el segundo, se intenta reconstruir tanto el tamaño y posición de un obstáculo dentro del tubo. En ambos problemas, se utilizan las RNAs para el reconocimiento de los perfiles de velocidad, vorticidad

o presión dinámica del flujo en simulaciones numéricas realizadas por medio del método de *lattice Boltzmann*.

La tesis se ha redactado en función de los artículos enviados y publicados hasta el momento de ser presentada. Primeramente en el capítulo 2 se exponen los antecedentes sobre redes neuronales artificiales, así como los métodos de entrenamiento utilizadas a lo largo del trabajo. En el capítulo 3, se comienza hablando sobre los artículos relacionados con la estimación del parámetro de razón de masas en el choque de un sistema binario de agujeros negros. En el capítulo 4 se describe el uso de RNAs para resolver los dos problemas inversos del flujo de un fluido clásico en la simulación numérica de un tubo en dos dimensiones espaciales. Para terminar el cuerpo principal de esta tesis, en el capítulo 5 se dan las conclusiones globales de los trabajos realizados en la tesis doctoral así como una discusión final sobre la aplicación del aprendizaje automático en problemas de Física. Finalmente, se han agregado cuatro apéndices relacionados tanto con el aprendizaje y prueba del código base de las RNAs implementadas, así como de las bases teóricas sobre los problemas que se han abordado en los capítulos principales, esto es: ondas gravitacionales y el método de *lattice Boltzmann*. También se ha agregado un glosario con definiciones de conceptos que se utilizan a lo largo del trabajo.

# Capítulo 2

## Redes Neuronales Artificiales

Dado que las redes neuronales artificiales (RNAs) nacen como un análogo del procesamiento de información del sistema nervioso de los seres vivos, es natural dar una breve explicación del funcionamiento estas redes neuronales biológicas, para extender el concepto de aprendizaje en los seres vivos hacia el denominado aprendizaje automático al cual pertenecen las RNAs [Rosenblatt, (1962)].

### 2.1. Redes neuronales biológicas

Los distintos sistemas nerviosos se componen de un conjunto de neuronas en la cuales un grupo inicial de ellas se encarga de recibir la información externa para transmitirla hacia otras neuronas. Este nuevo conjunto de neuronas se encarga de procesar la información, y un grupo posterior de interpretar todo el conjunto de impulsos nerviosos generados para tomar una decisión.

La información se propaga por las neuronas gracias a la sincronía de un flujo de iones a través de la membrana celular. El cambio de potencial provocado por el aumento o disminución de iones provoca que la transmisión eléctrica del impulso nervioso viaje a través de la neurona y hacia otras neuronas. La fuerza de conexión entre las neuronas, llamada sinapsis, determina la importancia o participación de cada neurona en los procesos neuronales, en otras palabras, a una mayor fuerza sináptica, mayor es la participación de la conexión de dicha neurona en la toma de decisión, ya sea inhibitoria o exhitatoria [Hebb, (1949)].



El funcionamiento de las redes neuronales es bastante compleja y su estructura varía no solamente de especie a especie, sino de individuo a individuo. Estas variaciones estructurales y de funcionamiento dependen no solamente del desarrollo biológico del individuo, sino también del aprendizaje que se tiene a lo largo de su vida referido como la plasticidad de la neuronas. Este mecanismo descrito en la teoría *Hebbiana* [Hebb, (1949)], propone que la plasticidad de las neuronas depende del aprendizaje obtenido través de los procesos iterativos en la transmisión de impulsos nerviosos, es decir, la recurrencia de un impulso nervioso para realizar una acción provoca que la fuerza de conexión mejore. Este argumento, es una de las bases en el inicio del desarrollo de redes neuronales artificiales para emular dicho aprendizaje en el reconocimiento de patrones y toma de decisiones [Russell & Norvig, (1995)].

Tomando en cuenta los argumentos expuestos para las redes neuronales biológicas, se les puede relacionar con las RNAs en estructura, modelación matemática del flujo de información entre las conexiones de las neuronas y la metodología para el aprendizaje.

## 2.2. Redes neuronales artificiales

Al igual que en las redes neuronales biológicas, las RNAs se componen por un gran número de elementos o nodos simples densamente conectados, llamados “neuronas”. Las neuronas de las RNAs reciben información de sus vecinos y posteriormente envían sus propias señales positivas (exhitorias) o negativas (inhibitorias). La forma y magnitud de la respuesta que se obtiene de cada neurona depende de la función de “disparo” o de activación.

En analogía con las redes neuronales biológicas, las neuronas de una RNA se pueden definir en tres tipos de unidades principalmente:

- Neuronas de entrada. Encargadas de recibir los datos a analizar y enviarlos al resto de neuronas de la RNA.
- Neuronas de procesamiento. Responsables de realizar los cálculos matemáticos correspondientes.

- Neuronas de respuesta o de salida. Encomendadas a entregar la solución. En general, tanto en las neuronas de procesamiento como las de salida se realizan los cálculos pertinentes.

El primer modelo de una neurona computacional fue realizado por [McCulloch & Pitts, (1943)], proponiendo a la neurona como una compuerta lógica con dos posibles estados: activo o inactivo. El estado de este tipo de neuronas se determina con una función tipo escalón y es dependiente de las conexiones de entrada, las cuales son pesadas y sumadas. En otras palabras, si el valor del estímulo rebasa un valor de umbral  $U$  la neurona se considera activa, de otra forma inactiva. Por ejemplo, definiendo el vector  $\mathbf{x}$  como los valores de entrada y  $\mathbf{w}$  como el vector de pesos correspondiente, el estado de la neurona se define por:

$$S = \begin{cases} 1 & \text{si } \mathbf{w}^T \cdot \mathbf{x} > U, \\ 0 & \text{si } \mathbf{w}^T \cdot \mathbf{x} < U; \end{cases} \quad (2.1)$$

donde  $\mathbf{w}^T$  define el transpuesto del vector y al producto de  $\mathbf{w}^T \cdot \mathbf{x}$  se le llama comúnmente en el ámbito de las RNAs como función de transferencia. Este modelo (esquemático en la Figura 2.1), es aún la base del elemento primario de las RNAs, donde la interconexión de neuronas permite realizar las tareas para las cuales han sido entrenadas.

El primer algoritmo de aprendizaje para este tipo de RNAs fue propuesto de forma independiente en [Rosenblatt, (1962)] y [Widrow, (1962)]. Se basaron en el paradigma de la recompensa y castigo para desarrollar sus máquinas denominadas “perceptron” y “adaline”. Estos sistemas eran capaces de categorizar la información si se encontraban los parámetros (pesos) adecuados. Sin embargo, [Minsky & Papert, (1988)] mostraron que este tipo de RNAs estaban limitadas a operaciones lógicas linealmente separables. Actualmente las estructuras de las RNA han evolucionado desde su concepción pasando desde una estructura muy sencilla hasta ser capaces de aproximar funciones no-lineales y reconocer patrones más sutiles en la información proporcionada. En términos básicos, las RNAs actuales funcionan como un mapeo de valores de un espacio a otro donde es posible que la información sea “más fácil” de interpretar, algo similar a problemas donde se recurre a analizar-

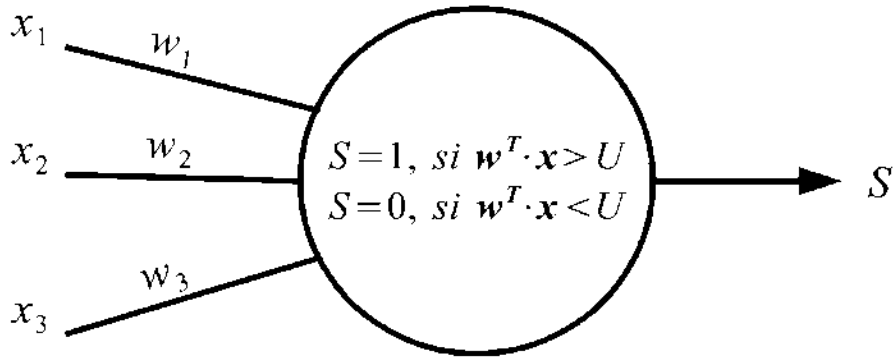


Figura 2.1: Esquema de los valores de entrada y pesos de las conexiones de cada una de ellas. El estímulo que recibe la neurona es la combinación lineal del vector de entrada  $\mathbf{x}$  y el vector de parámetros  $\mathbf{w}$ . El estado activo o inactivo de la neurona dependerá si la suma de este producto interno es mayor al umbral  $U$ .

los en el espacio de de frecuencias mediante transformadas de Fourier. Suponiendo por ejemplo, una RNA que recibe un vector de entrada  $\mathbf{x} \in \mathbb{R}^n$ :

$$\mathbf{x} = (x_1, x_2, \dots, x_n), \quad (2.2)$$

y se mapea a otro vector  $\mathbf{y} \in \mathbb{R}^l$ :

$$\mathbf{y} = (y_1, y_2, \dots, y_l). \quad (2.3)$$

Entonces la RNA se ve como una función  $f : f(\mathbf{x}) \rightarrow \mathbf{y}$ . Las estructuras de las RNA son actualmente muy variadas tanto en complejidad como flexibilidad respecto a los problemas que se desea resolver. Sin embargo, a grandes rasgos se puede englobar en 3 grupos:

- Redes Neuronales Multicapa: Aunque se abarcan distintas estructuras en esta rama, en general se trata de RNAs que se conforman por al menos tres capas de neuronas interconectadas, generalmente en su totalidad, donde la información se propaga en una sola dirección. Entre sus principales ventajas destaca su relativa sencillez en implementación con un rendimiento considerado como bueno en una gran aplicación de problemas. Sin embargo, esto también conlleva a que se vean limitadas en problemas muy complejos como

en el procesamiento de lenguaje, series de tiempo con gran cantidad de ruido, reconocimiento de objetos en imágenes de gran resolución, etc., [Rojas, (1996)]. Además, sí para resolver un problema en específico se consideran una gran cantidad de características de entrada (comúnmente denominadas *features* en inglés), la estructura de la RNA resultará en una gran número de interconexiones entorpeciendo la velocidad y calidad del aprendizaje. A pesar de esto, si el problema se plantea apropiadamente generalmente se obtienen soluciones adecuadas utilizando los algoritmos de aprendizaje actuales.

- Recurrentes: Este tipo de redes se asemejan a las anteriores en su estructura, sin embargo, la información también puede propagarse hacia capas anteriores, permitiendo reforzar las conexiones que sean apropiadas en una propagación. La recurrencia también puede ser temporal, lo que significa que son capaces de aprender de información introducida previamente, lo que las hace más apropiadas para análisis de señales de reconocimiento de lenguaje y texto. No obstante, estas mismas características hacen que sus algoritmos de entrenamiento sean sofisticados y a menudo su aprendizaje no llega a ser adecuado [Bogdan, (2009)].
- Convolucionales. Este tipo de RNAs, se pueden considerar como una ampliación de las multicapa con un esquema de interconexión y operaciones matemáticas más complejo, siendo utilizadas para un aprendizaje profundo. Su estructura está basada principalmente en el sistema visual de los seres vivos, por lo que son adecuadas para el reconocimiento de imágenes con un preprocesamiento de información mínimo. Las capas por las que se componen son divididas en distintos segmentos, aminorando el número de conexiones que se necesitarían en comparación con el de una multicapa normal con el mismo número de capas y elementos. Sus capas de convolución permiten que la generalización y el aprendizaje sean efectivos. Sin embargo, dependen altamente de un ajuste heurístico en su estructura y aún no es clara la razón de su eficiencia y como mejorarlas [Zeiler & Ferguson, (2013)].

En el trabajo se desarrollaron RNAs con estructura de multicapa, por lo que

será la única que se describe con una mayor profundidad a continuación.

## 2.3. Redes neuronales multicapa

El sistema de RNAs más común es la red neuronal multicapa o *Multilayer Perceptron* (MLP, por sus siglas en inglés), es una extensión del perceptron desarrollado por Rosenblatt. Son ampliamente utilizadas tanto por su simplicidad como por su desempeño satisfactorio en los problemas que se han abordado con ellas.

Las RNA tipo MLP, son llamadas así dado que las neuronas que la componen se encuentran estratificadas por capas, es decir, las neuronas iniciales que reciben la información se encuentran en la llamada capa de entrada, mientras que las neuronas de procesamiento y de salida se encuentran en capas distintas e independientes entre sí. Las neuronas en capas intermedias anteriores a las de salida son llamadas también neuronas ocultas, por lo que de aquí en adelante se hará la distinción en nombre entre las neuronas y/o capas ocultas con respecto a aquellas en la capa de salida, tal como se muestra en la Figura 2.2. En pocas palabras y considerando la misma figura, en las MLP la información viaja desde las neuronas de entrada hacia las neurona ocultas para posteriormente pasar a las neuronas de salida, realizando una serie de cálculos conforme esta viaja a través de la RNA. Este proceso es denominado propagación hacia adelante o *feedforward* (por su nombre en inglés), y a continuación se dan más detalles sobre este proceso.

### 2.3.1. Propagación hacia adelante

De ahora en adelante se dejará de utilizar el término MLP y se utilizará el término genérico de RNA. En los siguientes párrafos se describen las operaciones matemáticas que se realizan para la propagación hacia adelante suponiendo una RNA con una estructura de  $n$  entradas,  $m$  neuronas en una única capa oculta y  $l$  salidas, simplificando la notación para un solo vector de entrada como en la ecuación (2.2). En la siguiente capa con  $m$  unidades, la  $j$ -ésima neurona oculta recopila la información proveniente de todas las conexiones de las neuronas de entrada, siendo estos los mismos valores que el vector  $\mathbf{x}$  regulados en cada conexión

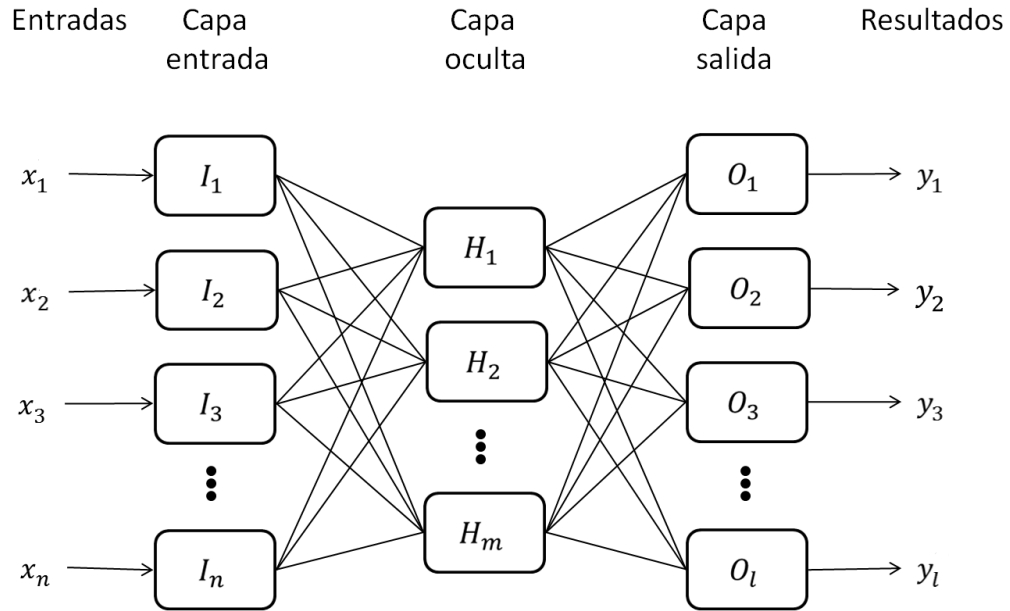


Figura 2.2: Esquema de una RNA tipo MLP o *Feedforward* con una sola capa oculta. En este esquema la RNA cuenta con  $n$  elementos de entrada,  $m$  neuronas ocultas y  $l$  de salida, indicadas con las letras  $I$ ,  $H$  y  $O$  respectivamente. Este tipo de RNA puede contar con más capas ocultas, las cuales se interconectan de forma seriada. Los elementos con los que se alimenta a la red se especifican en este caso por el vector  $\mathbf{x}$  mientras que los resultados con  $\mathbf{y}$ .

por un vector de pesos  $\mathbf{w}_j = (w_{0j}, w_{1j}, \dots, w_{nj})$ . La composición de estos dos vectores consiste en la combinación lineal entre el vector de entrada y los pesos de cada conexión. Es decir, la primera operación matemática que realiza esta unidad está dada por:

$$\sigma_j = \mathbf{w}_j^T \cdot \mathbf{x} = \sum_{i=1}^n w_{ij} x_i. \quad (2.4)$$

El escalar  $\sigma_j$  es entonces el resultado del mapeo de una función de  $\mathbb{R}^n \rightarrow \mathbb{R}$ , lo que permite hacer uso de funciones definidas en  $\mathbb{R}$  en la siguiente capa. Dichas funciones, en el contexto de las RNA, son llamadas funciones de activación. Suponiendo que  $F$  sea una función de activación para todas las neuronas en la

capa oculta, el resultado  $h_j$  en la  $j$ -ésima neurona se expresa como:

$$h_j = F(\sigma_j) = F\left(\sum_{i=1}^n w_{ij}x_i\right). \quad (2.5)$$

En general, las neuronas no reciben únicamente el valor proveniente de las neuronas de la capa anterior, sino también un parámetro umbral o *bias*, que juega un papel similar a  $U$  referido en la ecuación (2.1). Este parámetro puede ser considerado como una neurona adicional en la capa anterior con un valor siempre igual a 1 con un peso definido usualmente por  $b$ . Con esto, la expresión (2.5) se transforma en:

$$h_j = F\left(\sum_{i=1}^n w_{ij}x_i + b_j\right). \quad (2.6)$$

Finalmente, para la última capa el resultado de cada neurona de salida  $y_k$  se expresa como:

$$y_k = G(\sigma_k) = G\left(\sum_{j=1}^m w_{jk}h_j + b_k\right), \quad (2.7)$$

donde los pesos  $w_{jk}$  representan los pesos de la capa oculta a la de salida y  $G$  es la función de activación en la capa de salida. Las funciones utilizadas como base para las aproximaciones de las funciones en las RNA se han clasificado principalmente en tipo escalón, lineales o no-lineales. Las primeras nacieron como una completa analogía de las redes biológicas donde el potencial de acción que genera el impulso nervioso no es generado a menos que se rebase un umbral determinado, llamadas también como función de todo o nada. Sin embargo, como se refirió anteriormente, el tipo de problemas que se pueden resolver con este tipo de funciones es muy limitado.

### 2.3.2. Funciones de activación

La decisión sobre que funciones de activación utilizar, depende tanto de los datos de entrada como del vector de respuesta deseado. Considerando que la resolución de problemas, ya sea en términos de regresión o clasificación, tanto en las

RNAs como en otros métodos de aprendizaje automático se eligen funciones no-lineales que les brindan a dichos algoritmos propiedades analíticas y computacionales más complejas que los modelos de regresión lineal [Bishop, (1995), Rojas, (1996)]. En el caso particular de las RNAs, se han propuesto funciones de activación que cumplen ciertas condiciones en concordancia con el Teorema Universal de aproximación de funciones tipo sigmoide, tal y como se demostró en [Cybenko, (1989)]:

**Teorema.** *Sea  $\varphi$  una función continua, no-constante, acotada y monotonamente creciente. Sea  $I_m$  el espacio  $m$ -dimensional unitario en un hipercubo  $[0, 1]^m$ . El espacio de funciones continuas en  $I_m$  está definido por  $C(I_m)$ . Entonces, dada cualquier función sigmoide  $f$  en  $I_m$   $f \in C(I_m)$  y  $\varepsilon > 0$ , existe un entero  $N$ , con constantes reales  $c_i$  y  $b_i$  y vectores  $\mathbf{w}_i \in \mathbb{R}^m$ , donde  $i = 1, \dots, N$ , tal que:*

$$F = \sum_i^N c_i \varphi(\mathbf{w}_i^T \cdot \mathbf{x} + b_i), \quad (2.8)$$

como una aproximación de la función  $f$ , siendo independiente de  $\varphi$ , i.e.:

$$|F(\mathbf{x}) - f(\mathbf{x})| < \varepsilon, \quad (2.9)$$

para todos los  $\mathbf{x} \in I_m$ . En otras palabras, las funciones de la forma  $F(\mathbf{x})$  son densas en  $C(I_m)$  y esto se mantiene cuando se reemplaza  $I_m$  para cualquier subconjunto compacto de  $\mathbb{R}^m$ .

En términos de las funciones tipo sigmoide, en este trabajo se utilizaron las más usuales: la logística, la tangente hiperbólica y la normalización exponencial (conocida también como *softmax*). Las tres funciones son acotadas y tienden a un valor límite en sus extremos. La función logística estándar se define como:

$$F(\mathbf{x}) = \frac{1}{1 + \exp(-\sigma)} = \frac{1}{1 + \exp(-\mathbf{w}^T \cdot \mathbf{x} + b)}, \quad (2.10)$$

con un rango en  $(0,1)$ , ver Figura 2.3a. La función *softmax* se puede considerar como una extensión de la logística definida en el mismo rango, y que se utiliza generalmente en las capas de salida en una RNA. En concordancia con la ecuación (2.7), para la  $k$ -ésima neurona de salida, esta función se define como:

$$y_k = F(\sigma_k) = \frac{\exp(\sigma_k)}{\sum_o \exp(\sigma_o)} \quad (2.11)$$



donde  $o$  indiza a cada una de las neuronas en la capa de salida. El uso de esta función se muestra en el apéndice B. Por otro lado, la respuesta de salida de la tangente hiperbólica se encuentra en  $(-1,1)$ , tal y como se puede visualizar en la Figura 2.3b.

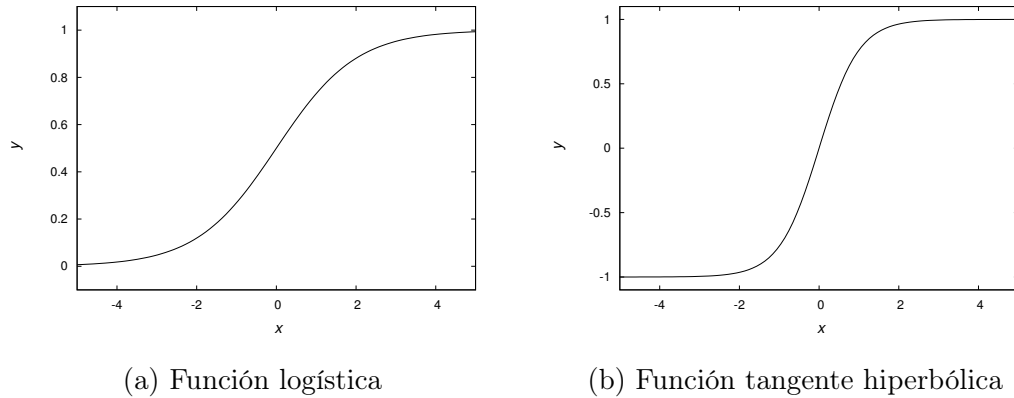


Figura 2.3: Funciones de activación tipo sigmoide más usuales en las RNAs. En (a) se puede apreciar el rango de la función logística con  $y \in (0, 1)$ . En (b) la tangente hiperbólica con  $y \in (-1, 1)$ .

En muchas ocasiones se prefiere la tangente hiperbólica dado que es una función simétrica centrada en el origen, por lo que en promedio la combinación lineal de los pesos y valores de entrada a la siguiente capa serán cercanos a 0. Esta y alguna otras más sugerencias heurísticas (*rules of thumb*, en inglés), que se siguieron en el trabajo de tesis sobre el uso de las funciones de activación, normalización de los vectores de entrada e inicialización de las RNA se encuentran en la referencia [LeCunn, Bottou, Orr, Muller,(2012)].

Tal y como se mencionó para las redes neuronales biológicas, la estimulación continua de las neuronas implica el fortalecimiento o debilitamiento de las conexiones sinápticas. Para el caso de las RNA sucede algo similar, lo que implica realizar un proceso de “entrenamiento”, esto es, enseñarle a la RNA el tipo de respuesta que uno desea.

## 2.4. Entrenamiento

En este trabajo se desarrollaron proyectos únicamente basados en un entrenamiento supervisado, como fue descrito brevemente en la sección 1.2, por lo que solamente se ahondará en este procedimiento. En este tipo de entrenamiento se necesitan los datos o **patrones ejemplo** a analizar, es decir, los **vectores de entrada** (*inputs*, por su nombre en inglés), y las respuestas esperadas, denominados **vectores objetivo** (*targets*, por su nombre en inglés). El par conformado por los vectores de entrada y objetivo deben ser representativos del problema a analizar, de tal manera que los datos reflejen todas las características de que son relevantes del problema en cuestión. De no ser así, la RNA puede llegar a ser ineficiente y no tener una conclusión apropiada a futuros datos. Bajo este enfoque, el ciclo de actividad de las RNA se puede definir en dos etapas:

1. La fase durante la cual la red neuronal adquiere la información necesaria a través de los patrones ejemplo para determinar lo que debe hacer y como debe hacerlo.
2. La etapa en la cual la RNA debe resolver nuevas tareas que le son presentadas, basándose en el aprendizaje adquirido.

La manera de medir y controlar el aprendizaje de las RNAs es mediante el cálculo del error de entrenamiento, esto es, medir la diferencia entre las respuestas esperadas y las hechas por la RNA y disminuirlo a través de algún procedimiento o algoritmo. En las subsecuentes secciones y al ir desarrollando el contenido sobre el entrenamiento de las RNAs, se describen las estrategias necesarias para que su aprendizaje sea más eficaz.

### 2.4.1. Error de entrenamiento

Como en el entrenamiento supervisado se tiene disponible tanto la respuesta de la RNA como el objetivo deseado, se puede estimar un error dada la diferencia entre ambos. Existen diversas formas de medir el error para cada uno de los patrones ejemplo del conjunto de entrenamiento. El error de salida de un único patrón

ejemplo, puede ser utilizado para modificar los parámetros de la RNA y calcular el desempeño global sobre el conjunto de vectores objetivo (que se escriben de ahora en adelante por  $\mathbf{t}$ ), proporcionando una estimación del aprendizaje a través de una función de error o de costo. Una de estas funciones de error es la función de tipo error cuadrático medio (MSE, por sus siglas en inglés). Siguiendo con la notación usada hasta ahora esta función de error se define como:

$$E = \frac{1}{lN} \sum_{p=1}^N \sum_{k=1}^l \frac{1}{2} (t_k^p - y_k^p)^2 = \frac{1}{2lN} \sum_{p=1}^N (\mathbf{t}^p - \mathbf{y}^p)^2 \quad (2.12)$$

donde  $p$  indiza cada uno de los  $N$  patrones del conjunto par  $\{\mathbf{x}, \mathbf{t}\}$  y  $l$  el número de elementos en los vectores de respuesta de la RNA  $\mathbf{y}$  y objetivo  $\mathbf{t}$ .

Aunque el error definido en la ecuación (2.12) es la medida de error más común para el entrenamiento, en realidad existen una gran cantidad de funciones de costo que se pueden emplear, dependiendo del objetivo del problema. Por ejemplo, una función de costo de valor absoluto, exponencial o de entropía cruzada, esta última utilizada en el apéndice B de este escrito en un problema de clasificación. En este contexto el objetivo del aprendizaje es minimizar la función de error 2.12 sobre todo el conjunto de entrenamiento, intentando reducirlo para cada uno de los patrones ejemplo. Aunque existen métodos que logran cambiar tanto la estructura, como las funciones de activación de las neuronas, generalmente los algoritmos se enfocan en modificar los pesos de las conexiones entre las neuronas de la RNA. Estas modificaciones, no suceden en un solo paso sino sobre múltiples iteraciones, cuantificando la respuesta de la RNA hacia los valores deseados sobre el conjunto de entrenamiento en cada iteración. La forma en que esto sucede depende del método de aprendizaje y los hiper-parámetros que utiliza cada algoritmo.

En el aprendizaje automático a fin de tener resultados o predicciones precisas de los datos que se le presenten a posteriori, se suele dividir el conjunto de patrones ejemplo disponible en dos subconjuntos:

- El conjunto de entrenamiento propiamente dicho, que será utilizado en la modificación de los pesos de la RNA.
- Un conjunto de validación que sirve como medida de error para evitar el sobre-ajuste hacia los patrones ejemplo con los que fue entrenada la RNA.

El objetivo de dividir el conjunto total consiste en detener el entrenamiento o guardar el valor de los parámetros de la RNA en cierta iteración de manera que se guarden los parámetros de la RNA cuando se alcanzó el error mínimo en el conjunto de validación. Este procedimiento se llama validación cruzada y una ejemplificación de esto se muestra en la Figura 2.4, que se ha extraído de una prueba descrita en el apéndice B. En dicha figura, se ve que el error de validación incrementa a partir de una iteración determinada, momento en el cual se han guardado los parámetros de la RNA, mientras que el error en el conjunto de entrenamiento continua disminuyendo como es de esperar. Esta técnica evita el sobre-entrenamiento de la RNA, lo que permite que se generalice sobre los patrones desconocidos y evitando que aprenda con una excelente precisión únicamente sobre los datos del conjunto de entrenamiento, algo como lo que ocurre en la Figura 2.5.

Siguiendo el esquema de un entrenamiento supervisado, los algoritmos que siguen esta metodología pueden resumirse a grandes rasgos en los siguientes pasos:

1. Inicio. Definir los subconjuntos de entrenamiento y validación tanto vectores de entrada como objetivo.
2. Medir el error sobre el conjunto de validación.
3. Presentar cada uno de los patrones de entrada y medir el error en base a la respuesta deseada del conjunto de entrenamiento.
4. Ajustar los pesos de la RNA con el algoritmo seleccionado.
5. Si el error en el subconjunto de validación actual es menor que el error anterior se guardan los valores de todos los pesos, definiendo así la mejor configuración de parámetros hasta el momento.
6. Si el error de validación no ha llegado a un valor mínimo deseado o notable y/o no se han alcanzado el número de iteraciones máximas predefinidas, regresar al paso 3. De otra forma continuar al siguiente paso.
7. Salir del algoritmo, guardando la RNA con la configuración de pesos con el menor error obtenido.

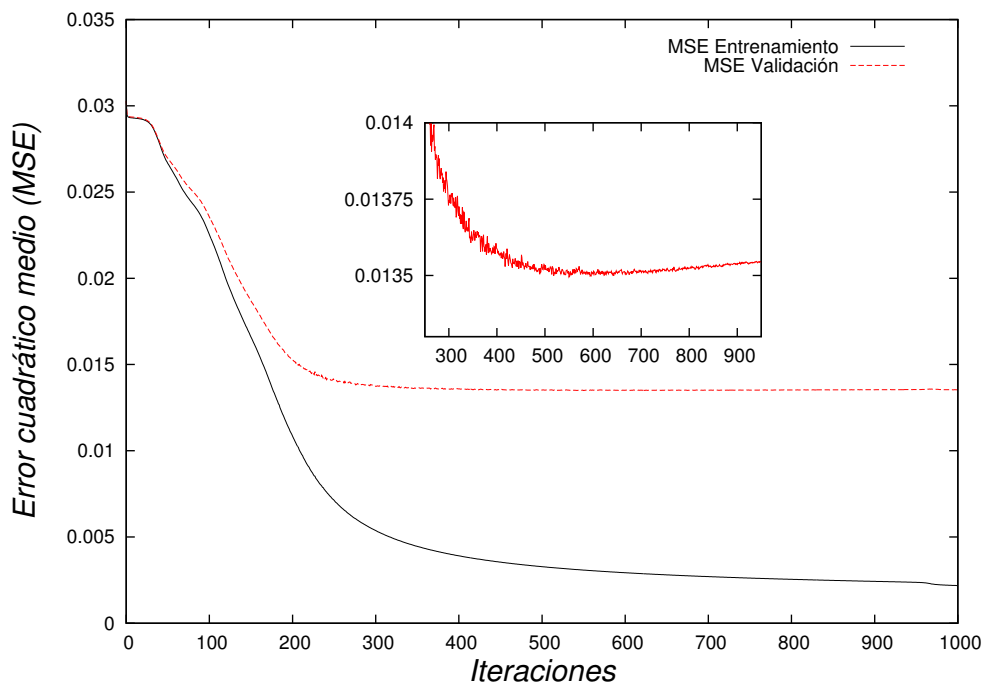


Figura 2.4: Error cuadrático medio calculado sobre el conjunto de entrenamiento (línea negra continua) y en el de validación (línea roja punteada) en función del número de iteraciones en la fase de aprendizaje de una RNA. El recuadro dentro de la gráfica es un acercamiento a la curva de error en el conjunto de validación. El sobre-entrenamiento comienza cuando este error comienza a incrementar continuamente. La explicación sobre el ejemplo hecho para esta gráfica se encuentra en la Figura B.2 del apéndice B.

En general, los algoritmos son mucho más complicados y requieren emplear condiciones más sutiles como la aleatoriedad al presentar los ejemplos del conjunto de entrenamiento, así como distinguir si el mínimo alcanzado es suficiente para obtener buenas predicciones a posteriori. A continuación se revisará un método en particular que se utilizó principalmente en los trabajos publicados: el algoritmo de propagación hacia atrás o más comúnmente conocido por su nombre en inglés como *backpropagation*.

## 2.4.2. Backpropagation

El algoritmo de *backpropagation* tiene como base principal de entrenamiento, la modificación de los valores de todos los pesos de la RNA siguiendo la dirección

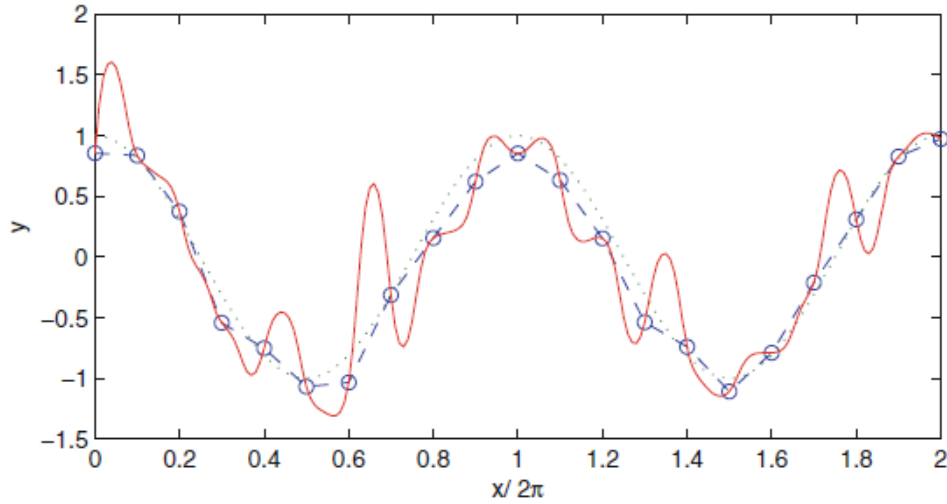


Figura 2.5: Ejemplo del sobre-ajuste de los parámetros de la RNA a los patrones de entrenamiento. Cuando esto sucede la RNA no generaliza de forma apropiada la función a la cual debería aproximar. En círculos azules se presentan los datos usados para el aprendizaje de la RNA al agregar un poco de ruido en una función coseno (línea negra punteada). La RNA llegó a la solución representada por la línea roja que claramente se encuentra alejada de la solución adecuada para los puntos desconocidos en el entrenamiento. Imagen tomada de [Du, Ke-Lin & Swamy, M.N.S., (2013)].

en la cual la función de error se minimiza localmente de forma más rápida, en un proceso iterativo de tipo gradiente descendente [Rumelhart, Hinton & Williams, (1986)]. Por ejemplo, considerando  $\mathbf{W}$  como el conjunto de pesos de una RNA:

$$\mathbf{W} = \{w_1, w_2, w_3, \dots, w_{n \cdot m + m \cdot l}\}. \quad (2.13)$$

donde por simplificación de notación no se han considerado los términos de *bias* de la RNA. A manera de continuar con la ejemplificación de una RNA con  $n$  entradas,  $m$  neuronas ocultas y  $l$  salidas. Por tanto, el gradiente que minimiza la función de error  $E$  se obtiene buscando el conjunto de parámetros  $\mathbf{W}$  tal que:

$$\nabla_{\mathbf{W}} E = \left( \frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_\nu}, \dots, \frac{\partial E}{\partial w_{n \cdot m + m \cdot l}} \right) \rightarrow 0, \quad (2.14)$$

donde  $w_\nu$  indiza cada uno de los pesos en la RNA. El propósito de este algoritmo es cambiar el valor de cada  $w_\nu$  a la siguiente iteración  $t + 1$ , propagando el error desde

las neuronas de salida hasta las neuronas de entrada en la iteración  $t$  mediante la regla de la cadena:

$$w_\nu(t+1) = w_\nu(t) + \Delta w_\nu(t) = w_\nu(t) - \gamma \frac{\partial E}{\partial w_\nu(t)}, \quad (2.15)$$

donde  $\Delta w_\nu(t) = -\gamma \frac{\partial E}{\partial w_\nu(t)}$  y  $0 < \gamma < 1$  es la constante de aprendizaje, que debe ser elegida por el usuario. Por ejemplo y continuando con la configuración de RNA de la sección 2.3.1, suponiendo un único patrón ejemplo, representado como el vector en la ecuación (2.2). Mientras que cada vector objetivo se define por  $\mathbf{t} = (t_1, t_2, \dots, t_l)$ . Considerando además a  $F$  como la función de activación tanto en la capa oculta como en la capa de salida, y reescribiendo por simplicidad las ecuaciones (2.6) y (2.7) sin considerar el término de *bias* como:

$$y_k = F(\sigma_k) = F\left(\sum_j w_{jk} h_j\right); \quad (2.16)$$

$$h_j = F(\sigma_j) = F\left(\sum_i w_{ij} x_i\right). \quad (2.17)$$

El término  $\Delta w_\nu$  en la ecuación (2.15), para el peso de la conexión entre la  $j$ -ésima neurona en la capa oculta y el  $k$ -ésimo elemento en la capa de salida se escribe de la forma:

$$\Delta w_{jk} = -\gamma \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial \sigma_k} \frac{\partial \sigma_k}{\partial w_{jk}} h_j \quad (2.18)$$

Mientras que para el peso de la conexión entre la  $i$ -ésima neurona en la capa de entrada y la  $j$ -ésima en la capa oculta,  $\Delta w_\nu$  se escribe como:

$$\Delta w_{jk} = -\gamma \sum_{k=1}^l \left[ \delta_{jk} \frac{\partial \sigma_k}{\partial h_j} \right] \frac{\partial h_j}{\partial \sigma_i} \frac{\partial \sigma_i}{\partial w_{ij}} x_j \quad (2.19)$$

donde:

$$\delta_{jk} = \frac{\partial E_k}{\partial y_k} \frac{\partial y_k}{\partial \sigma_k}. \quad (2.20)$$

Suponiendo  $F$  la función logística, definida en la ecuación (2.10), cuya derivada se puede escribir de la forma:

$$\frac{\partial F(\sigma)}{\partial \sigma} = F(\sigma) (1 - F(\sigma)) \quad (2.21)$$

y el error cuadrático medio definido como en la ecuación (2.12), para un único vector ejemplo es:

$$\frac{\partial E}{\partial y_k} = (y_k - t_k), \quad (2.22)$$

por lo que las ecuaciones (2.18) y (2.19) se escriben como:

$$\Delta w_{ij} = -\gamma(y_k - t_k)y_k(1 - y_k), \quad (2.23)$$

y

$$\Delta w_{jk} = -\gamma \sum_{k=1}^l [(y_k - t_k)y_k(1 - y_k)w_{jk}] y_j(1 - y_j)x_i, \quad (2.24)$$

respectivamente. Uno de los inconvenientes de usar un método de gradiente descendente, es la posibilidad de quedar atorado en un mínimo local, lo cual conlleva un mal ajuste para futuros datos. Para mejorar el algoritmo y reducir el riesgo de estancamiento en un mínimo local, se agrega generalmente un hiper-parámetro  $\alpha$ , proveyendo así un cierto “momento” al ajuste de los parámetros [Rumelhart, Hinton & Williams, (1986)]. Este cambio de velocidad se logra en el algoritmo haciendo que las modificaciones de los pesos dependan también de sus cambios anteriores:

$$\Delta w_\nu(t) = -\gamma \frac{\partial E}{\partial w_\nu(t)} + \alpha \Delta w_\nu(t - 1) \quad (2.25)$$

con  $0 < \alpha < 1$  la constante de momento. Otras mejoras y extensiones se han hecho al algoritmo de *backpropagation* [LeCunn, Bottou, Orr, Muller,(2012)]. También existen otros métodos de optimización de segundo orden para el ajuste de los pesos como Newton, Gauss-Newton, Levenberg Marquardt, Quasi-Newton (BFGS) y Gradiente Conjugado [Bishop, (1995)].

## 2.5. Consideraciones y discusiones finales

La aplicación de las distintas RNAs presentadas en este trabajo, fueron desarrolladas siguiendo reglas de normalización, basados generalmente en los rangos de valores de los vectores de entrada. Esta técnica junto con la de validación cruzada permite obtener un ajuste más apropiado de los pesos en las RNAs y así evitar el sobreajuste a los patrones de entrenamiento. Además, se ha utilizado el



error cuadrático medio definido en la fase de entrenamiento definido en la ecuación (2.12). Sin embargo, los resultados para las predicciones se muestran generalmente utilizando la raíz del error cuadrático medio (RMSE, por sus siglas en inglés):

$$RMSE = \sqrt{\sum_{p=1}^N \frac{(\mathbf{t}^p - \mathbf{y}^p)^2}{N}} \quad (2.26)$$

donde  $p$  indiza el par  $\{\mathbf{x}, \mathbf{t}\}$  de un patrón ejemplo del conjunto de entrenamiento. En el apéndice A se muestra un pseudo-código con la implementación de una versión del algoritmo de *backpropagation on line*. Un procedimiento similar se siguió en este trabajo donde la modificación de los pesos se realiza para cada patrón seleccionado aleatoriamente del conjunto de ejemplos de entrenamiento que se le muestran a la RNA en su fase de aprendizaje. Esta versión del algoritmo de *backpropagation* converge más rápidamente hacia un mínimo [LeCunn, Bottou, Orr, Muller, (2012)], además que al presentar aleatoriamente cada uno de los patrones de entrenamiento es más probable evitar un mínimo local. A diferencia del *backpropagation offline* o por lotes, donde se presenta el conjunto completo de los patrones de entrenamiento, donde ahora el error es el promedio de los errores entre cada los vectores objetivo y los resultados de la RNA, como el error definido en la ecuación (2.12).

Aunque en este capítulo se han descrito los elementos necesarios para la construcción de las RNAs generadas para todos los trabajos realizados, en los siguientes capítulos se remarcan las consideraciones y la notación propia para cada problema a partir de algunas de las ecuaciones que en este capítulo se han reseñado.

## Capítulo 3

# Estimación de la Razón de Masas en Ondas Gravitacionales de la Colisión de Agujeros Negros.

Este capítulo se basa principalmente en el artículo *Parameter estimates in binary black hole collisions using neural networks*, publicado en *General Relativity and Gravitation* [Carrillo, González, Gracia & Guzmán, (2016)]. Algunos detalles de los antecedentes teóricos sobre las ondas gravitacionales (GWs) se encuentran en el apéndice C, enfocado especialmente a la radiación producida por el colapso de un sistema binario de agujeros negros (BBH, por sus siglas en inglés). En este capítulo se da una reseña breve sobre ellas y centrando el análisis con el objetivo de dar una estimación de un parámetro intrínseco del evento a partir de una señal de onda gravitacional ocasionada por la colisión entre dos agujeros negros. Se resume la capacidad y practicidad de las RNAs en resolver el problema inverso: dada una señal de una GW se desea determinar la razón de masas iniciales ( $q$ ) del sistema binario de los agujeros negros.

### 3.1. Ondas gravitacionales

Hasta el año 2015, no se había detectado ninguna onda gravitacional (GW por sus siglas en inglés) de forma directa y la prueba de su existencia se infería a partir

del decaimiento del momento angular de sistemas binarios, como por ejemplo, de las observaciones del pulsar PSR 1913+16 [Hulse & Taylor, (1975)], el cual cierra su órbita debido a la radiación gravitacional.

A excepción de la observación del pulsar PSR 1913+16, hasta entonces la única forma que se tenía de obtener información sobre la radiación gravitacional, era a través de simulaciones computacionales resolviendo las ecuaciones de Einstein mediante relatividad numérica para los distintos fenómenos capaces de producir las ondas gravitacionales, tales como estrellas binarias de neutrones, colapso de estrellas de supernovas o procesos de acreción de altas energías o sistemas binarios de estrellas o de agujeros negros [Sathyaprakash & Schutz, (2009)]. Estos últimos son hasta el momento, los mejores candidatos a ser detectados por parte de los distintos proyectos experimentales como el Laser Interferometer Gravitational Wave Observatory (LIGO), Advance Virgo en Italia, KAGRA en Japón, entre otros [Caltech: Educational resources, (2017)]. Las señales de GWs de sistemas binarios se caracterizan por tres fases: orbital, la fusión y el amortiguamiento o decaimiento, más conocidas por sus nombre en inglés como las fase *inspiral*, *merger* y *ringdown*. Un bosquejo de la dinámica de este sistema, así como las correspondientes fases se pueden visualizar mejor con ayuda de la Figura 3.1, en la que se representa el prototipo de señal detectada o *strain* (como se le conoce técnicamente, ver apéndice C), a través del interferómetro al interactuar con una GW, producto del colapso de un sistema binario de agujeros negros.

Con la primera observación directa de la onda gravitacional GW150914 [Abbot et al., (2016a)], el 14 de septiembre de 2015, finalmente se pudo confirmar sin duda alguna la hipótesis de la existencia de las GWs. La GW150914 fue producto del colapso de un sistema binario de agujeros negros a una distancia de la Tierra de  $410_{-180}^{+160}$  Mpc, con masas iniciales (expresadas en masas solares  $M_{\odot}$ ), aproximadas de  $m_1 = 36_{-4}^{+5} M_{\odot}$  y  $m_2 = 29_{-4}^{+4} M_{\odot}$ , resultando en un solo agujero negro con una masa estimada de  $62_{-4}^{+4} M_{\odot}$ , y una energía gravitacional radiada de  $3.0_{-0.5}^{+0.5} M_{\odot} c^2$ . La señal generada en los detectores y posteriormente procesada en ambos interferómetros de LIGO en Hanford, Washington y Livingston, Louisiana en Estados Unidos se puede apreciar en la Figura 3.2. La comparativa del *strain* en ambos interferómetros, así

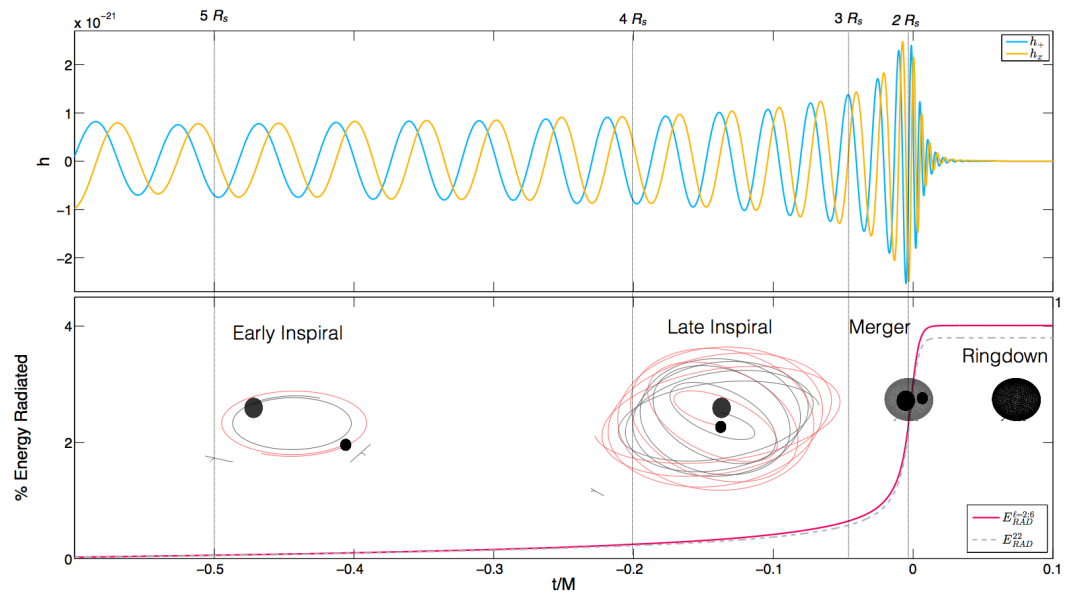


Figura 3.1: Fases orbital, de fusión y decaimiento (*inspiral*, *merger* y *ringdown* en inglés) de una onda gravitacional producida por la colisión de dos agujeros negros (arriba). El *strain* dibujado tanto por  $h_+$  y  $h_\times$  representa el tipo de señal que se extrae en los detectores. Como se puede apreciar en el gráfico inferior, en la fase orbital el porcentaje de energía radiada en la fase orbital es relativamente menor al momento en que ambos agujeros negros se encuentran muy próximos; alcanzando un máximo al momento de la fusión. Después de la fusión el sistema aún se encuentra inestable por lo cual continúa radiando energía gravitacional en la fase de decaimiento o *ringdown*. Imagen tomada de [Georgia Tech, (2017)]

como el incremento de frecuencia asociado al inminente colapso del sistema binario, es evidencia contundente que la observación fue en verdad una señal real y no una falla de los detectores.

A partir de la detección de la GW150914, comenzó una nueva era en el estudio de la Física al permitir obtener información del Universo de otro tipo de radiación distinta a la electromagnética. Esto ha llevado a que los distintos grupos de investigación en astrofísica se hayan volcado en investigar de manera más intensa el estudio de procesos fuente de ondas gravitacionales. Desde entonces se han detectado dos nuevas señales de GW, también producto de sistemas binarios de agujeros negros: la GW151226 y más recientemente la GW170104, detectadas también gracias al experimento LIGO [Abbot et al., (2016b), Abbot et al., (2017)].

La detección de la GW es solo el comienzo de la investigación en este aspecto, dada la importancia de conocer las propiedades físicas de las fuentes que las generaron. Para solucionar este problema inverso en particular, generalmente se resuelve computacionalmente el sistema, barriendo un gran espectro de parámetros físicos como las masas iniciales entre los agujeros negros, espín y precesión del sistema, resolviendo para cada uno de los casos las combinaciones posibles de los parámetros involucrados, lo que hace este método bastante extenuante. A pesar de tales restricciones, se han realizado esfuerzos para elaborar catálogos con una colección formas de onda con el *strain* de GWs generadas por la colisión de sistemas BBH, ya sea solucionando las ecuaciones de Einstein mediante relatividad numérica como en los trabajos de [Ajith et al., (2012), Mroue, (2013), Jani et al., (2016)]; modelos fenomenológicos [Khan et al., (2016)] o modelos surrogados [Field, Galley, Hesthaven, Kaye & Tiglio, (2014)] en las cuales se incluyen tanto las formas de onda de las señales de ondas gravitacionales como las condiciones y parámetros iniciales de las fuentes con las que se obtuvieron dichas soluciones.

La resolución del problema inverso, es decir, la estimación de parámetros de la fuente de una GW es abordada mediante diferentes estrategias incluyendo cadenas de Markov [Van der Sluys et al., (2009)] y la interpolación de formas de onda [Smith et al., (2013)]. Por ejemplo, la biblioteca LALInference desarrollada por el grupo de investigación principal de LIGO [Veitch et al., (2015)] usa métodos estocásticos

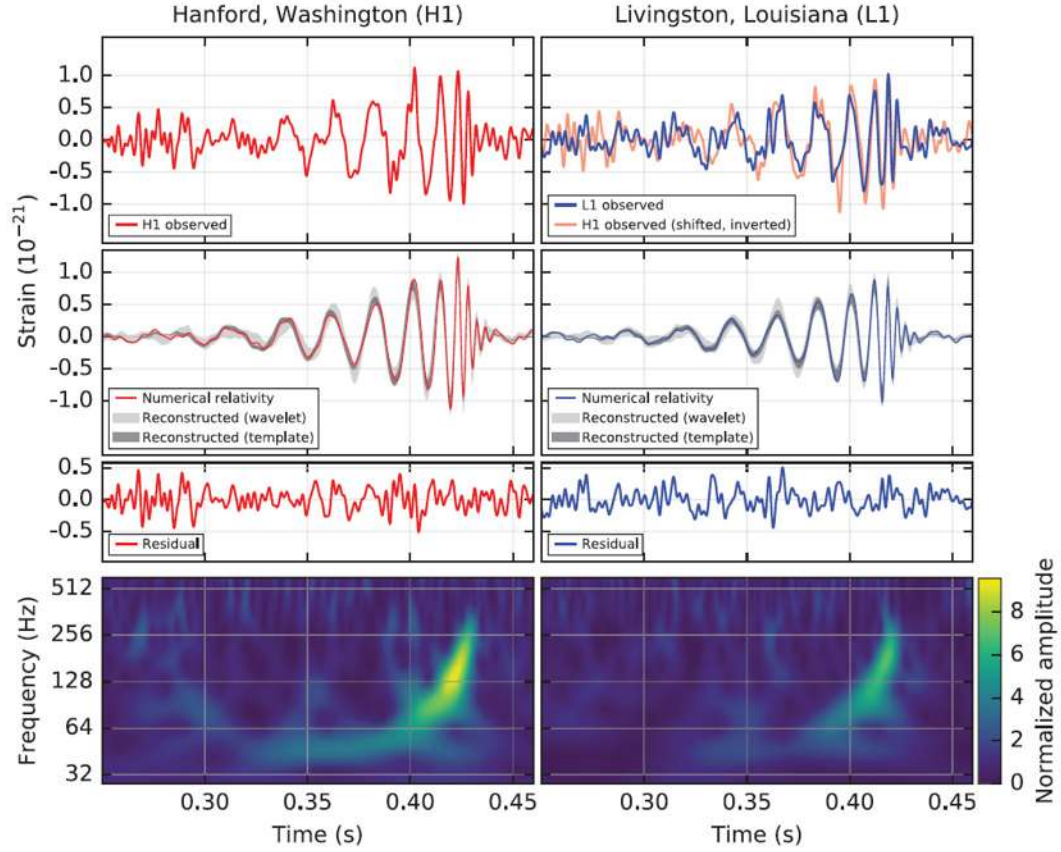


Figura 3.2: Señal observada por la onda gravitacional GW150914 en los detectores de LIGO en Hanford, Washington (H1, gráficos en la columna izquierda) y Livingston, Louisiana en Estados Unidos (L1, gráficos en columna derecha). En los dos paneles superiores, se muestra el *strain* observado en ambos detectores y su gran similitud es prueba de la detección de una GW. En los paneles intermedios se muestran las simulaciones numéricas y los residuos al compararlas con las observaciones reales, de tal manera que las señales puedan ser interpretadas y los parámetros de la fuente sean inferidos, en este caso se determinó que la GW es producto de un sistema binario de agujeros negros con masas de  $m_1 = 36_{-4}^{+5}M_\odot$  y  $m_2 = 29_{-4}^{+4}M_\odot$ , resultando en un solo agujero negro con una masa estimada en  $62_{-4}^{+4}M_\odot$ , y una energía gravitacional radiada de  $3.0_{-0.5}^{+0.5}M_\odot c^2$ . En los paneles inferiores se presentan los espectrogramas del incremento de frecuencia en función del tiempo, conocido como *chirp*, provocado por el colapso del sistema. Imagen tomada de [Abbot et al., (2016a)].

tipo Monte Carlo y Cadenas de Markov (MCMC). A diferencia de estos métodos, en este trabajo se ha intentado ser más eficiente al reducir la búsqueda estocástica y el espacio de parámetros mediante RNAs.

Ya se han aplicado métodos de aprendizaje automático en el análisis de GWs, por ejemplo en [Biswas et al., (2013)] se utilizan para clasificar y remover fallas (*glitches* en inglés) de las señales recolectadas en los detectores; en [Baker et al., (2015)] se utilizan para clasificar mejoras de sensibilidad en términos de las masas del sistema binario de agujeros negros. En el contexto relatado hasta ahora, el trabajo que se describe en este capítulo se enfocó en analizar el tipo de GWs ocasionadas por la colisión de un sistema binario de agujeros negros y estimar la razón de masas  $q$  por medio de RNAs. Para ello, se planeó construir una RNA, que ayudará a estimar  $q$  a partir de una GW en ambos una señal libre de ruido y de una señal con ruido Gaussiano considerando dos catálogos distintos. Este análisis se ha dividido en dos casos de estudio principalmente:

1. De un primer catálogo se consideró únicamente tanto la fusión como el decaimiento de la señal de la GW. Tomando en cuenta tanto señales limpias, así como contaminándolas con ruido Gaussiano.
2. De otra serie de simulaciones, en un catálogo independiente al primero, se consideró tanto la fase orbital, fusión y decaimiento de la forma de onda, con lo que se ha incrementado la complejidad de la RNA en información para analizar. De la misma forma que el caso anterior, se ha inspeccionado el desempeño de las RNAs para señales con y sin la inclusión de ruido Gaussiano.

Comenzando con el primer caso de estudio, a continuación se definirá la metodología usada para la generación y análisis de las formas de onda. Posteriormente se hará lo mismo para el segundo analizando sus variantes.

### 3.2. Estimación de razón de masas de sistema binario de agujeros considerando las fases de fusión y decaimiento

En el primer caso, todas las formas de onda han sido extraídas a partir del catálogo generado para el artículo [Carrillo, González, Gracia & Guzmán, (2016)], que corresponden a ondas gravitacionales generadas por sistemas binarios de agujeros negros sin espín y con una razón de masas  $q$  que va desde  $q = 1$  hasta  $q = 5$ . Además, el caso de estudio se ha dividido en dos subcasos: formas de onda libres de ruido y señales con ruido Gaussiano para una razón señal-ruido (SNR, por sus siglas en inglés) que va desde 0.5 a 25. Por lo que se probará el desempeño de las RNA bajo distintas condiciones de entrenamiento y prueba.

La generación numérica de las formas de onda del catálogo se realizó usando el código `EINSTEINTOOLKIT` [Löffler, (2012)], variando las condiciones iniciales del sistema BBH, considerando la separación inicial entre los dos agujeros negros tal que se obtenga una órbita cuasi-circular. Para ello, se resolvieron las ecuaciones de Einstein bajo el formalismo BSSN [Baumgarte & Shapiro, (1999), Shibata & Nakamura, (1995)] en un dominio definido de  $[-120M, 120M]^3$ , con  $M$  la masa del sistema en el formalismo ADM [Arnowitt, Deser & Misner, (1959)] y verificando que en todas las configuraciones iniciales la masa ADM del espacio-tiempo sea el mismo. De dichas simulaciones se extrajeron los modos  $l = 2$  y  $m = 2$  en  $r = 30M$  del escalar de Weyl  $\psi_4$ , convirtiéndolas posteriormente en la parte real del *strain*, i.e.,  $r(h_+)_{22/M}$ , fijando la resolución del tiempo de salida a  $0.125M$ . Para más detalles sobre la generación de este catálogo checar la referencia [Gracia, (2017)]. En la Figura 3.3 se muestra la parte real para algunas de las señales con distintas razones de masa, que han sido utilizadas como serie de tiempo en el trabajo.



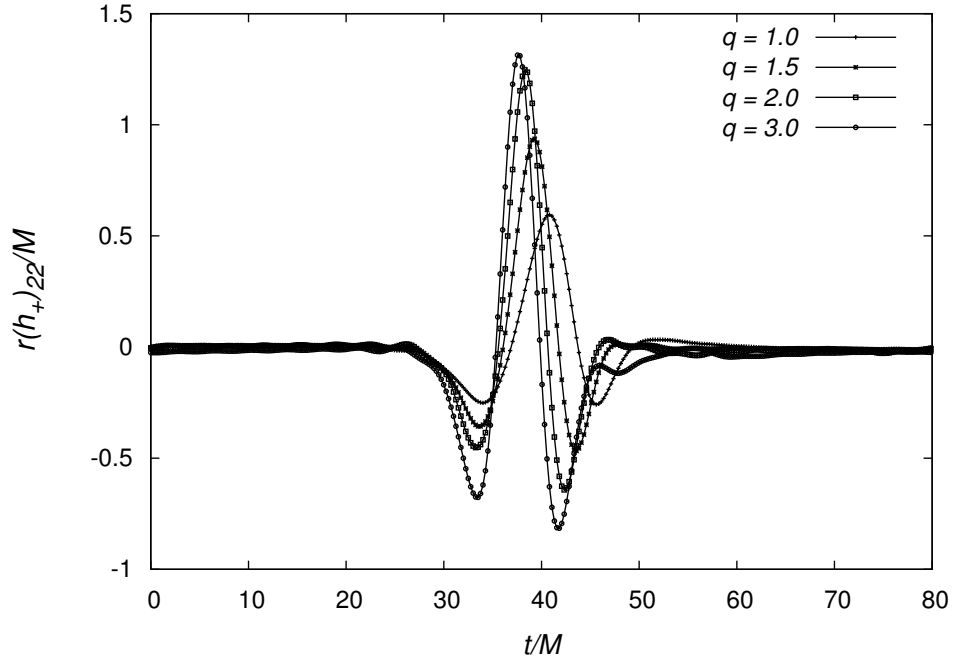


Figura 3.3: Series de tiempo de las ondas gravitacionales simuladas con razón de masas  $q = 1.0, 1.5, 2.0, 3.0$ , de las cuales se extrajo la parte real de los modos  $l = 2, m = 2$  del *strain*, es decir,  $r(h_+)_{22}/M$ . Estas señales fueron escaladas en el radio de extracción  $r$  y la masa  $M$  del sistema.

### 3.2.1. Descripción y selección de datos para el análisis mediante las RNA

Como se mencionó anteriormente, en este caso se toma en cuenta la señal de la GW en las fases de fusión y decaimiento del sistema binario, las cuales se comprenden a partir del punto máximo hasta que terminan las oscilaciones de las formas de onda. Con esto en mente, se ha definido el punto máximo como el inicio de la serie de tiempo que forma los vectores de entrada para la RNA.

Considerando que la estimación de las propiedades físicas de las GWs recae en la selección adecuada de la información que se introduzca a la RNA; la fase de fusión y decaimiento de las señales simuladas se traduce en términos adimensionales a tomar en cuenta los primeros 130 valores de la secuencia temporal a partir del punto máximo. Esta selección del tamaño de la ventana de la serie de tiempo es tal

que contiene los puntos suficientes para todas las formas de onda seleccionadas en el conjunto de entrenamiento, pudiendo no ser así tanto para las formas de onda en los conjuntos de validación o predicción. Por otro lado, la serie de tiempo extraída de  $r(h_+)_{22/M}$  ha sido escalada en el rango  $(-1,+1)$ , tomando como referencia el valor máximo y mínimo de las formas de onda del conjunto de entrenamiento exclusivamente.

En términos de la RNA, dados los valores tanto positivos como negativos en las formas de onda, la tangente hiperbólica resulta más apropiada a utilizar en la capa de oculta. Para la neurona de salida se consideró una función lineal para la regresión permitiendo una extrapolación en la parte superior del dominio. En este caso, también se evidencía el desempeño de distintas estructuras de RNAs, variando el número de neuronas ocultas desde 10 hasta 100 elementos. Los resultados reportados no se han hecho considerando una única RNA, sino tomando la media de los resultados hechos por un ensamble de 10 RNAs que han sido inicializadas de forma aleatoria tal que:

$$q = \frac{1}{10} \sum_{i=1}^{10} q_i. \quad (3.1)$$

Con todas las consideraciones anteriores y siguiendo la ecuaciones (2.4-2.7), la razón de masas  $q$  estimado por cada una de las 10 RNAs está determinado por:

$$q = \sum_{j=1}^m w_{jk} \tanh \left( \sum_{i=1}^{130} w_{ij} r(h_+)_{22/M}(t=i) + b_j \right) + b_0, \quad (3.2)$$

donde  $m = 10, 20, \dots, 100$  dependiendo del caso de estudio.

Como se ha descrito en la teoría de las RNAs, el conjunto de simulaciones realizadas se debe dividir en los conjuntos de entrenamiento, validación y predicción. En este caso, del conjunto de las 32 señales disponibles se seleccionaron 10 para el entrenamiento, 3 para la evaluación y 19 elementos para la predicción. Dicho conjunto se muestra en la Tabla 3.1. La selección del conjunto de predicción permite inspeccionar el rendimiento de la RNA para la estimación de  $q$  en una zona de interpolación y otra de extrapolación en relación a las señales con las que fue entrenada la RNA. El entrenamiento se ha hecho siguiendo el algoritmo de *back-propagation* descrito en la sección 2.4.2, usando validación cruzada y minimizando

Conjunto de valores de $q$		
Entrenamiento	Validación	Predicción
1.00	2.75	1.2, 1.4
1.10	3.00	1.6, 1.8
1.25	3.25	1.9, 2.1
1.30		2.2, 2.3
1.50		2.4, 2.6
1.70		2.7, 2.8
1.75		2.9, 3.5
2.00		3.4, 3.75
2.25		4.0, 4.5
2.5		5.0

Tabla 3.1: Subconjuntos de las señales generadas de simulaciones de GWs para distintos  $q$ , definido aquí por cada uno de los números. Para el subconjunto de predicción, el rango de  $1.0 \leq q \leq 2.5$  se define como el régimen de interpolación, mientras que para  $2.6 \leq q \leq 5.0$  es la región de extrapolación.

la función de error cuadrático medio. Sin embargo, los resultados son presentados para cada conjunto siguiendo la ecuación (2.26) del RMSE, en este caso escrito como:

$$RMSE = \sqrt{\sum_{p=1}^N \frac{(q_t^p - q^p)^2}{N}} \quad (3.3)$$

donde  $q_t$  representa la verdadera razón de masas y  $p$  es el índice de cada señal en el conjunto correspondiente con  $N$  elementos.

Como se refirió anteriormente y de la Tabla 3.1, el conjunto de predicción se puede dividir en dos secciones: la sección donde se incluyen las razones de masa que se encuentran en la zona de interpolación en relación al conjunto de entrenamiento que va desde  $q = 1.2$  hasta  $q = 2.4$ ; y la sección de extrapolación partiendo de  $q = 2.6$  hasta  $q = 5.0$ .

### 3.2.2. Resultados

#### Resultados para señales sin ruido

Para este subcaso, se ha medido el desempeño de las RNAs en base al RMSE obtenido en el conjunto de validación y en función del número de neuronas desde 10 elementos hasta 100. Los resultados sobre esto se muestran en las Figuras 3.4 y 3.5, donde es evidente la diferencia de precisión entre los conjuntos de entrenamiento, validación y predicción. Por otro lado, el incremento de neuronas ocultas no parece ser tan relevante en términos de una disminución significativa del error.

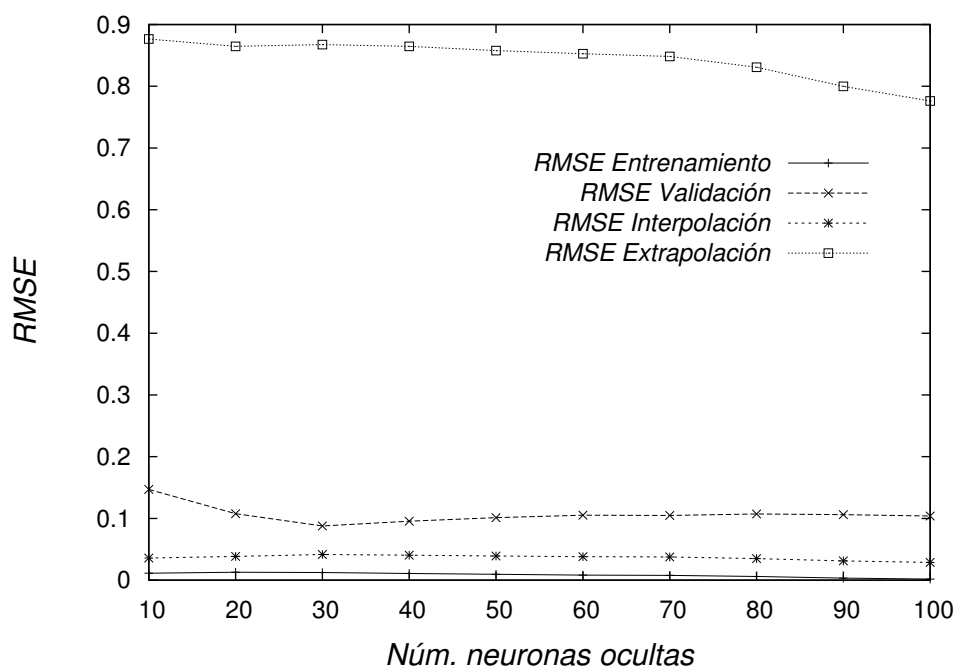


Figura 3.4: Error (RMSE) en la fase de entrenamiento, validación y predicción de señales sin ruido y en función del número de neuronas. Como es de esperarse el error de entrenamiento es menor que los otros dos. Igualmente, el error de predicción en el régimen de interpolación es menor que en el de extrapolación. El error es prácticamente el mismo aun cuando se incrementan el número de neuronas ocultas. Por lo que se ha decidido perseguir el esquema más sencillo, es decir, con 10 neuronas ocultas.

Como muestra de las estimaciones de las RNAs para algunas estructuras, en la Tabla 3.2 se presentan los resultados para el conjunto de predicción con formas

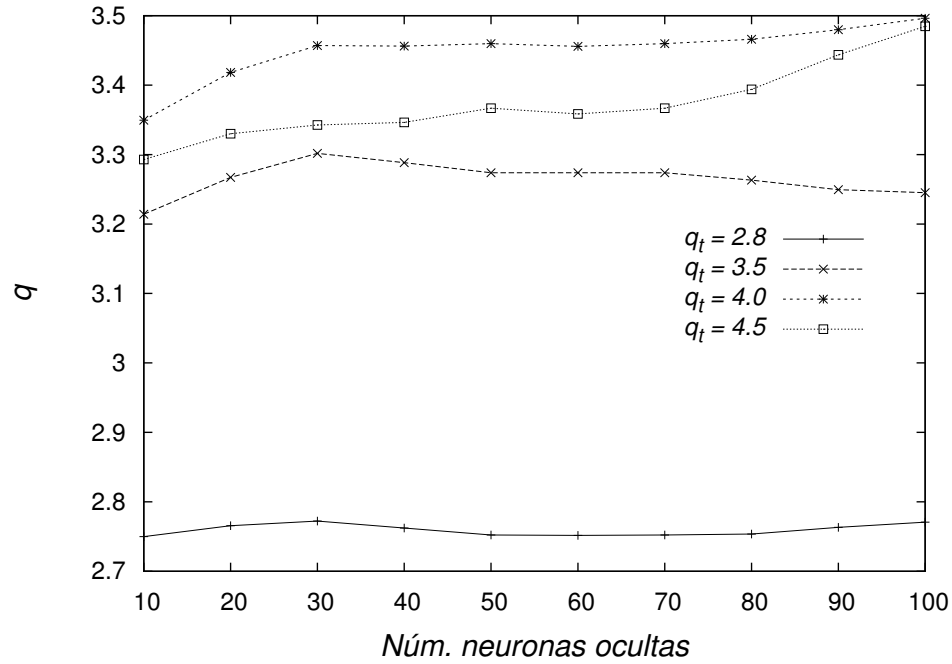


Figura 3.5: Valores predichos de la razón de masas  $q$ , en GWs sin ruido en el régimen de extrapolación, para los valores objetivo  $q_t = 2.8, 3.5, 4.0$  y  $4.5$ , en función del número de neuronas desde 10 hasta 100. La estimación de  $q$  por parte de las RNA no es tan buena como en la zona de interpolación, sin embargo parece no mejorar sustancialmente aun cuando se incrementa el número de elementos en la capa oculta.

de onda libre de ruido para tres casos de 10, 50 y 100 neuronas ocultas. Además como era de esperarse, para el régimen de extrapolación la magnitud del error es más significativo tal y como se observa en las gráficas de la Figura 3.6, donde en el peor de los casos el error porcentual alcanza un 50% para  $q = 5$ .

### Resultados señales con ruido Gaussiano

Para el subcaso, donde las señales han sido contaminadas con ruido Gaussiano, se han estudiado diversas amplitudes de la razón señal-ruido (SNR), que van desde  $SNR = 0.5$  hasta  $SNR = 25$ . Esta contaminación se agrega a cada forma de onda para los tres conjuntos: entrenamiento, validación y predicción. Por ejemplo, en la Figura 3.7 se presenta  $r(h_+)_{22/M}$  para  $q = 1$  y  $SNR = 10$ .

En la Figura 3.8 se muestra el RMSE para ambos regímenes de interpolación y

Predicciones de $q$			
$q$	RNA-10	RNA-50	RNA-100
1.2	$1.182 \pm 0.003$	$1.183 \pm 0.0001$	$1.186 \pm 0.001$
1.4	$1.384 \pm 0.007$	$1.384 \pm 0.0008$	$1.399 \pm 0.004$
1.6	$1.560 \pm 0.008$	$1.593 \pm 0.003$	$1.621 \pm 0.007$
1.8	$1.796 \pm 0.003$	$1.799 \pm 0.001$	$1.798 \pm 0.002$
1.9	$1.936 \pm 0.005$	$1.930 \pm 0.001$	$1.920 \pm 0.004$
2.1	$2.056 \pm 0.009$	$2.042 \pm 0.002$	$2.059 \pm 0.005$
2.2	$2.188 \pm 0.001$	$2.192 \pm 0.0007$	$2.193 \pm 0.0004$
2.3	$2.261 \pm 0.004$	$2.263 \pm 0.001$	$2.270 \pm 0.002$
2.4	$2.333 \pm 0.012$	$2.314 \pm 0.003$	$2.338 \pm 0.007$
2.6	$2.481 \pm 0.013$	$2.475 \pm 0.003$	$2.514 \pm 0.013$
2.7	$2.632 \pm 0.012$	$2.639 \pm 0.001$	$2.667 \pm 0.012$
2.8	$2.750 \pm 0.016$	$2.756 \pm 0.004$	$2.771 \pm 0.008$
2.9	$2.733 \pm 0.023$	$2.788 \pm 0.003$	$2.825 \pm 0.013$
3.4	$3.163 \pm 0.032$	$3.229 \pm 0.008$	$3.216 \pm 0.004$
3.5	$3.214 \pm 0.036$	$3.278 \pm 0.013$	$3.245 \pm 0.013$
3.75	$3.358 \pm 0.035$	$3.441 \pm 0.012$	$3.432 \pm 0.009$
4.0	$3.350 \pm 0.034$	$3.455 \pm 0.010$	$3.496 \pm 0.007$
4.5	$3.293 \pm 0.029$	$3.353 \pm 0.007$	$3.485 \pm 0.032$
5.0	$2.663 \pm 0.090$	$2.647 \pm 0.022$	$2.873 \pm 0.063$

Tabla 3.2: Promedio de las predicciones calculadas con 10 RNAs inicializadas de forma aleatoria, en el caso de señales sin ruido. En la tabla se presentan las variaciones para 10, 50 y 100 neuronas ocultas, escritas como RNA-10, RNA-50 y RNA-100. En general, las tres estructuras presentan valores muy similares. Además, es evidente como la incertidumbre, que se ha asociado a la desviación estándar, de las predicciones de cada ensamble aumenta al alejarse de la zona de interpolación, es decir, entre  $2.6 \leq q \leq 5.0$ .

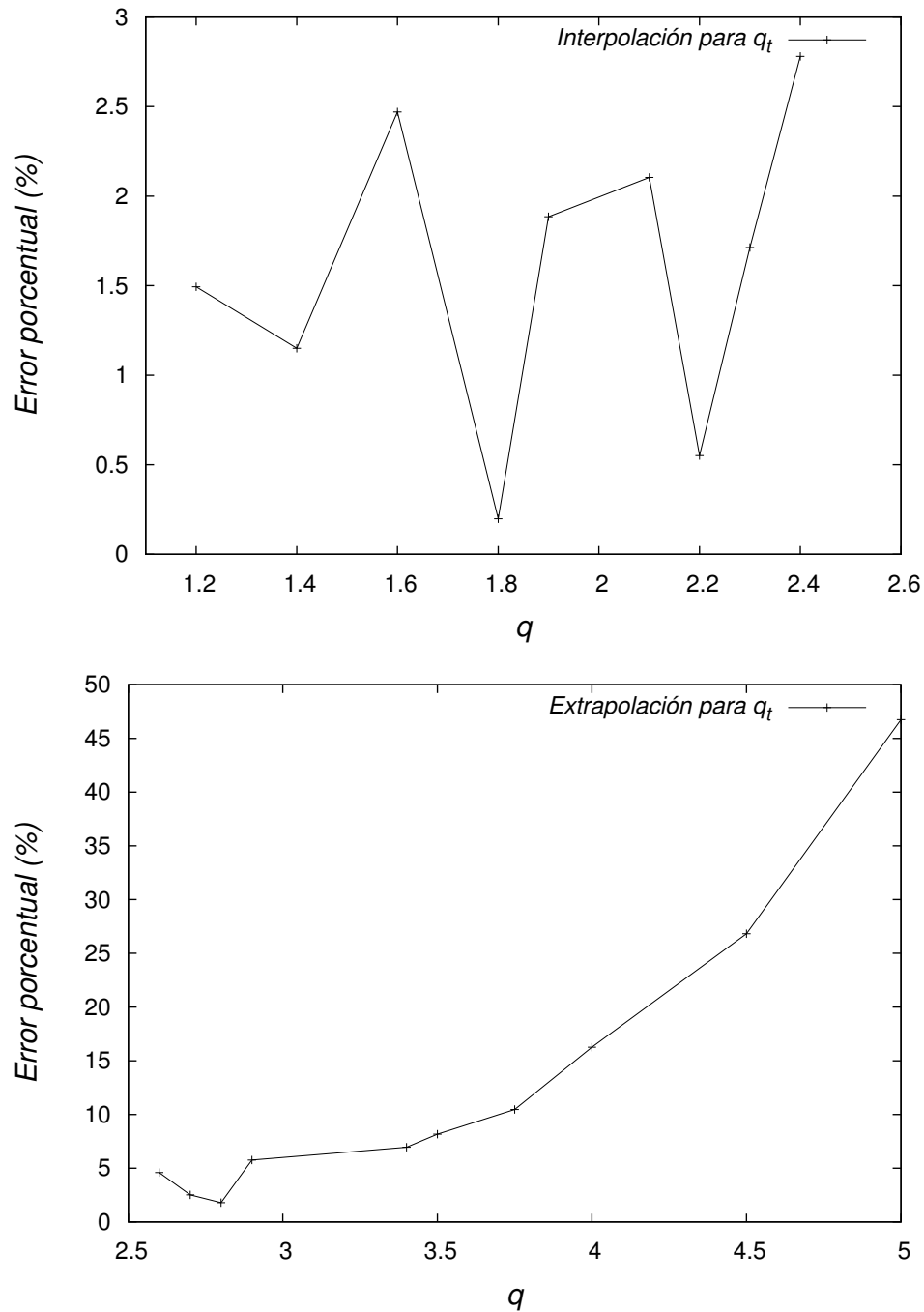


Figura 3.6: Porcentaje de error, en señales de GWs sin ruido, entre la predicción de la razón de masas por parte de la RNA compuesta por 10 neuronas ocultas y el valor objetivo  $q_t$ . En el gráfico de arriba se presenta la incertidumbre en la zona de interpolación en la cual fue entrenada la RNA, mientras que el gráfico de abajo se utilizaron señales de GW con  $q$  en el rango de extrapolación. Es clara la respuesta de la RNA en los dos regímenes, en la zona de extrapolación, la incertidumbre incrementa a medida que se aumenta  $q$ .

extrapolación en función del  $SNR$ , en los cuales es notorio el efecto del ruido en el error de predicción. También es de notar que la magnitud de los errores es mucho mayor en comparación con el análisis entre señales limpias. Este comportamiento es de esperarse dado que al disminuir el  $SNR$ , es más difícil para las RNAs encontrar la correlación aprendida en el entrenamiento para las señales dentro del ruido.

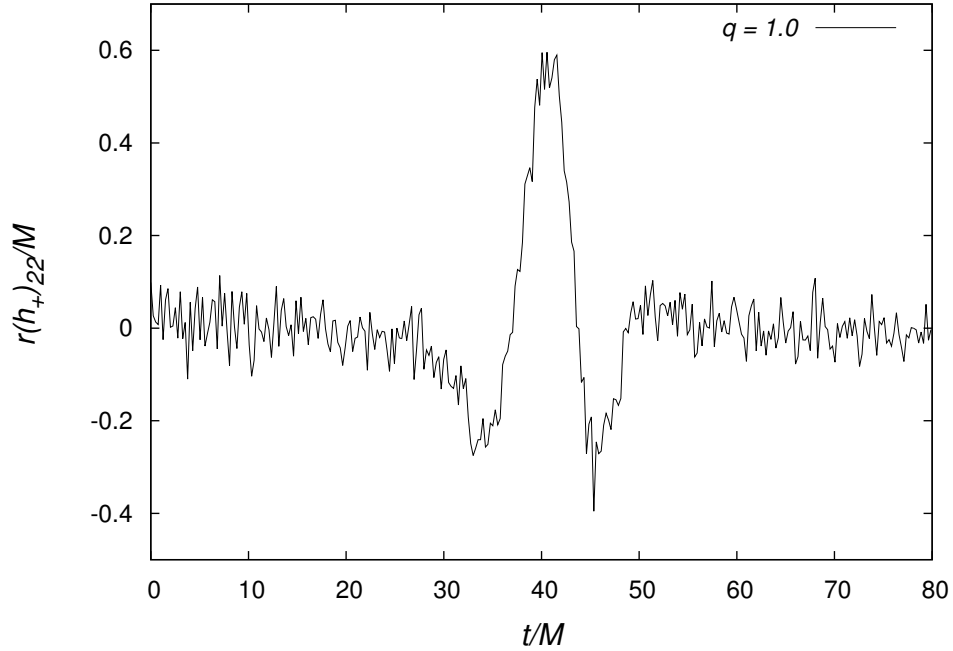


Figura 3.7: *Strain*  $r(h_+)_{22}/M$  en los valores escalados como se introducen en la RNA para un sistema binario de agujeros negros con la misma masa contaminado con ruido Gaussiano con un  $SNR = 10$ .

Los resultados indican que el conjunto de entrenamiento, aunque escaso, provee predicciones con un error de 3% en el régimen de interpolación, el cual crece rápidamente a valores de  $q$  que se encuentran lejos de dicha zona.



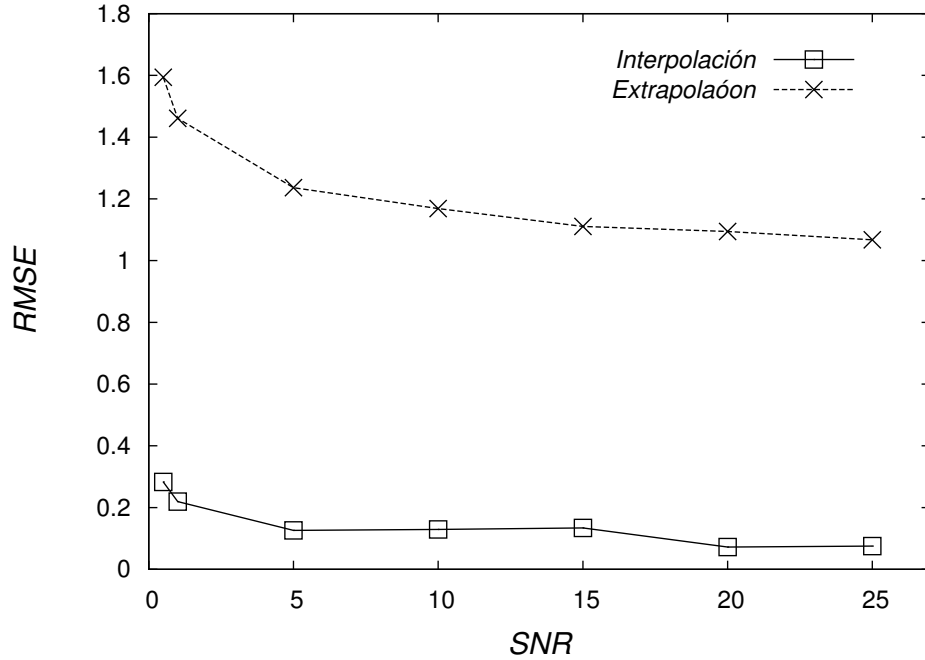


Figura 3.8: Error (RMSE) en el conjunto de predicción para señales con ruido Gaussiano que va desde un  $SNR = 0.5$  hasta 25. Es evidente la diferencia en la precisión entre el régimen de interpolación y extrapolación, así como también se incrementa el error en función del aumento de ruido.

### 3.3. Estimación de la razón de masas considerando fase orbital y ringdown

El catálogo de ondas gravitacionales generadas numéricamente considerado en esta sección del trabajo fue generado por el grupo de investigación de Georgia Tech [Jani et al., (2016), Georgia Tech, (2017)]. Las formas de onda fueron generadas empleando el código `MAYA`, basado en el código `EINSTEINTOOLKIT` que se utilizó en la generación de ondas del catálogo utilizado en la sección anterior. Este catálogo consiste en 452 formas de onda<sup>1</sup>, de las cuales 128 de ellas pertenecen a sistemas binarios de agujeros negros sin espín de precesión, esto es, o no tienen spin alguno o la relación entre sus espines están alineados paralela o antiparalelamente, y en los cuales la razón de masas se encuentra acotada para  $q \leq 15$ . El espacio de

<sup>1</sup>Al momento de la publicación del artículo citado.

parámetros de este caso en dicho catálogo puede visualizarse en la Figura 3.9. El restante de las 324 simulaciones son configuraciones de sistemas con distintos espines relativos entre ambos agujeros negros y con una razón de masas máxima tal que  $q \leq 8$ .

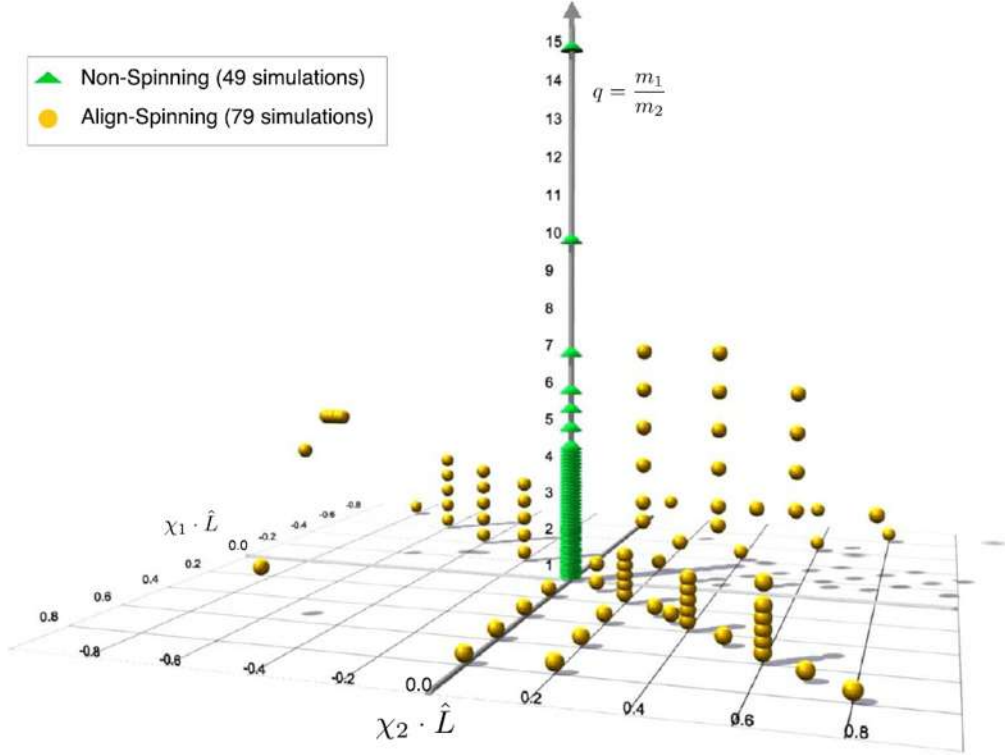


Figura 3.9: Esquemización del rango cubierto en el espacio de parámetros para las simulaciones de ondas gravitacionales realizadas y publicadas en el catálogo público de Georgia Tech. En esta gráfica se aprecia los distintos valores utilizados tanto para la razón de masas entre los agujeros negros  $q = m_1/m_2$  sin espín ( $\Delta$ ) y con espín alineado ( $\circ$ ). Imagen modificada de [Jani et al., (2016)].

Para esta parte del trabajo, nuevamente se inspeccionó encontrar el valor del parámetro  $q$  para sistemas donde ambos agujeros negros carezcan de espín. Debido que en el catálogo las formas de onda con esta características se vuelven escasas a partir de  $q > 5$ , se ha decidido únicamente considerar las simulaciones numéricas con una razón de masas  $1.0 \leq q \leq 5.0$ . Por lo que el conjunto se reduce de 128 formas de onda a 43, el cual ha sido dividido en los conjuntos para las fases de

Conjunto de valores de $q$		
Entrenamiento	Validación	Predicción
1.00, 1.10	1.20	1.15
1.30, 1.40	1.70	1.25
1.50, 1.60	1.90	1.75
1.80, 2.00	2.10	2.05
2.20, 2.30	2.70	2.35
2.50, 2.60	2.90	3.40
2.80, 3.00	3.10	3.90
3.20, 3.30	3.60	
3.50, 3.70	4.10	
3.80, 4.00		
4.20, 4.30		
4.40, 4.50		
5.00		

Tabla 3.3: Valores de  $q$  correspondientes a los utilizados para las tres fases en el aprendizaje de las RNA, prueba y predicciones.

entrenamiento, validación y predicción tal y como se muestra en la Tabla 3.3.

Como un nuevo enfoque se explora el impacto de considerar un segmento de la fase orbital del sistema binario de agujeros negros antes de la fusión. Adicionalmente, se examinan dos enfoques en el análisis de la señal: considerando la serie de tiempo de cada GW en forma análoga a la sección anterior 3.2, cuyo subcaso se identifica como **análisis de serie de tiempo** (*AST*) y como **análisis en el dominio de frecuencias** (*AF*).

A continuación se describirá de forma detallada el pre-procesamiento de las formas de onda tanto en el análisis como serie de tiempo como en el espectro de frecuencias antes de ser introducidas en las RNAs.

1. Análisis como serie de tiempo de las señales de las ondas gravitacionales. Al igual que en el enfoque utilizado en la sección anterior (3.2), en cada una de las señales se identifica el punto máximo. A diferencia que ahora a partir del punto máximo de la onda se han considerado 100 valores del *strain* anteriores, es decir, la parte orbital antes de la fusión, reduciendo el muestreo cada 5 valores respecto a la onda original, esto con la finalidad de reducir el número

de elementos de entrada que se utilizan en la RNA. Además se consideran también los 100 valores posteriores al punto máximo de la forma de onda en un muestreo cada 2 puntos en el tiempo, considerando nuevamente tanto la fase de fusión como decaimiento. Algunos ejemplos de los segmentos del *strain* utilizados para el análisis de las formas de onda se muestran en la Figura 3.10a.

2. Análisis de las señales de las ondas gravitacionales en el espectro de frecuencias. En esta propuesta se aplica la transformada discreta de Fourier a cada una de las formas de onda:

$$\hat{r}(h_+(k))_{22/M} = \sum_{t=0}^{N-1} r(h_+(t))_{22/M} e^{-2\pi tk/n}, \quad (3.4)$$

con  $N$  el número de valores de la serie de tiempo de la forma de onda, y  $k$  las frecuencias asociadas. Este análisis evita hacer una reducción en el muestreo de la serie de tiempo. Posteriormente, se ha calculado la potencia del espectro de las frecuencias calculado como:

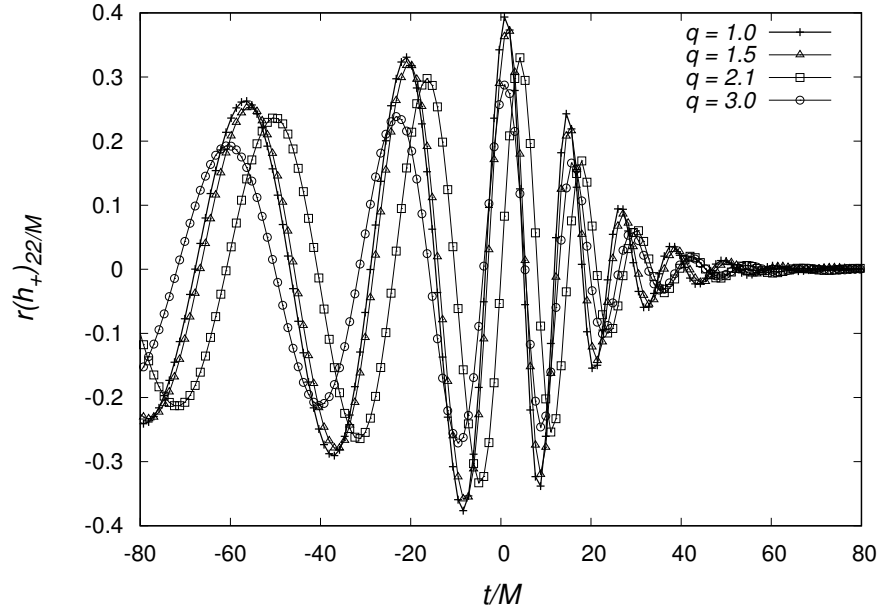
$$|\hat{r}(h_+(k))_{22/M}|^2, \quad (3.5)$$

permitiendo que se consideren únicamente los primeros 50 puntos de los perfiles correspondientes a las bajas frecuencias, descartando los demás valores que en amplitud son despreciables. Algunas ejemplificaciones con este enfoque se muestran en la Figura 3.10b.

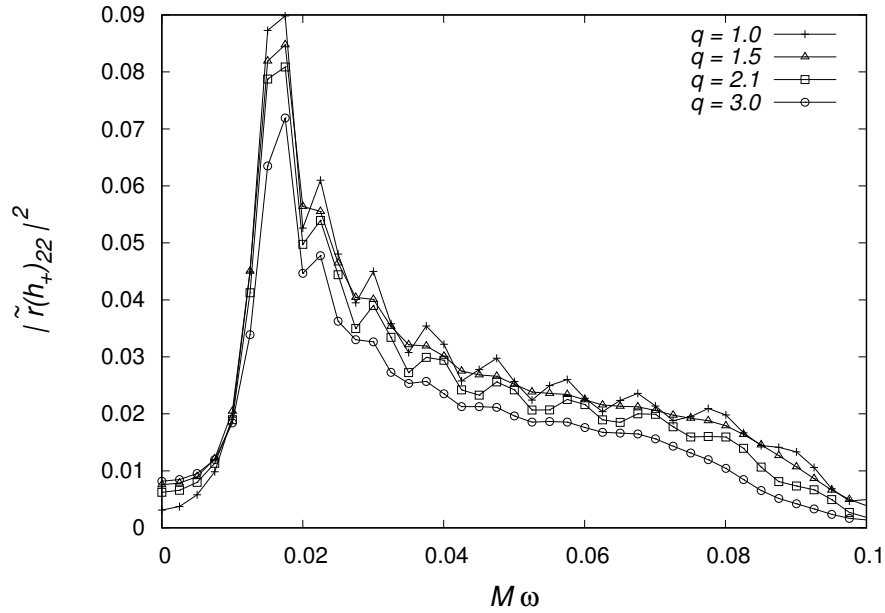
De forma similar a la sección 3.2, se ha examinado también como afecta la inclusión de ruido Gaussiano en la señales obtenidas, extrayendo los valores a utilizar como vectores de entrada siguiendo los criterios descritos en los puntos 1 y 2 con cinco diferentes valores de  $SNR = 5, 10, 15, 20$  y  $25$ .

### 3.3.1. Estructura y entrenamiento de las RNAs

Bajo los argumentos de los puntos 1 y 2 de la sección anterior, la estructura general de las RNAs generadas es la siguiente: doscientas neuronas de entrada y



(a)



(b)

Figura 3.10: (a) *Strain* de  $r(h_+)_{22}/M$  para algunas señales utilizadas en el caso de estudio sin ruido. (b) Potencia del espectro de frecuencias de la transformada de Fourier  $|\tilde{r}(h_+)_{22}|^2$  para los mismos casos que en (a). Aquí  $M = m_1 + m_2$ , con  $m_j = \sqrt{\mathcal{A}_j/16\pi}$  la masa aparente irreducible y  $\mathcal{A}_j$  es el área del horizonte de eventos de cada agujero negro.

una capa oculta con cincuenta elementos para el *AST*; mientras que se utilizaron cincuenta y veinte elementos en la capa de entrada y oculta respectivamente para el *AF*. En ambos casos, se ha utilizado la tangente hiperbólica como función de activación en las neuronas ocultas y una sola neurona con una función lineal como salida de respuesta asociada a  $q$ . La fase de aprendizaje de la RNA, se ha realizado bajo el esquema de entrenamiento supervisado, minimizando el error cuadrático medio usando el algoritmo de *backpropagation*.

Para el análisis de series de tiempo, un vector de entrada  $\mathbf{I}$ , relacionado directamente con los valores de la simulación de  $r(h_+)_{22/M}$  se escribe como:

$$\mathbf{I} = x_{\max-500}, x_{\max-495}, \dots, x_{\max-5}, x_{\max}, x_{\max+2}, x_{\max+4}, \dots, x_{\max+198}, \quad (3.6)$$

donde  $x_{\max}$  define el punto máximo de la forma de onda  $r(h_+)_{22/M}$ . Para el análisis de frecuencias los elementos de entrada son los primeros 50 puntos del perfil del espectro de frecuencias que corresponden al rango de frecuencias entre  $0M\omega$  y  $0.1M\omega$ . Nuevamente en concordancia con la ecuación (2.7) para ambas metodologías la propagación hacia adelante de las RNA es de la forma:

$$q = \sum_{j=1}^m w_j \tanh \left( \sum_{i=1}^n w_{ij} x_i + b_j \right) + b_0, \quad (3.7)$$

donde  $w_{ij}$  y  $w_j$  son los pesos de la RNA de la capa de entrada a la capa oculta y de la capa oculta a la neurona de salida respectivamente. Es decir,  $n = 200$  y  $m = 50$  son el número de entradas y neuronas ocultas en el *AST* y  $n = 50$  y  $m = 20$  para el *AF*. Para la determinación puntual de  $q$  se ha entrenado un conjunto de RNAs siguiendo los mismos procedimientos descritos hasta ahora, y la predicción final que se presenta es el promedio de todas las predicciones individuales del comité siguiendo un procedimiento similar al descrito en [Heskes, (1997)]. Esto también da la ventaja de dar intervalos de confianza del 95 %, relacionados con la varianza de los resultados del comité de las diez RNAs determinados por:

$$q = \langle q \rangle \pm 1.96(\sigma), \quad (3.8)$$

donde  $\langle q \rangle$  es el promedio de las predicciones respecto a  $q$  del comité y  $\sigma$  es la

varianza de dichos resultados dado por:

$$\sigma^2 = \frac{1}{nn - 1} \sum_{i=1}^{nn} (q_i - \langle q \rangle)^2. \quad (3.9)$$

En nuestro caso el número de RNAs en el comité  $nn = 10$  y  $q_i$  es el resultado de la  $i$ -ésima RNA.

### 3.3.2. Resultados

Nuevamente la precisión de las RNA se ha evaluado de tres formas: 1) calculando la raíz del error cuadrático medio (RMSE), 2) mediante el error porcentual de cada señal con diferente  $SNR$  y, 3) calculando el coeficiente  $R^2$  como métrica de ajuste para todo el conjunto de valores reales de  $q$ . El coeficiente  $R^2$  se ha calculado siguiendo la fórmula:

$$R^2 = 1 - \frac{\sum_i (q^i - q_t^i)^2}{\sum_i (q_t^i - \langle q_t \rangle)^2}, \quad (3.10)$$

donde  $q^i$  y  $q_t^i$  son la  $i$ -ésima predicción de la RNA y el correspondiente vector objetivo respectivamente, mientras que  $\langle q_t \rangle$  es el promedio de los  $q_t$  sobre el conjunto completo de predicción. Nótese que en el caso perfecto  $R^2 = 1$ , mientras que si la respuesta del comité de RNAs se acerca al promedio  $\langle q_t \rangle$ , entonces  $R^2 \rightarrow 0$ .

En el primer experimento que se ha nombrado *Limpias-Ruido*, se ha entrenado a la RNA con señales sin ruido en el conjunto de entrenamiento y validación, mientras que para las señales del conjunto de predicción se presentan los resultados tanto para señales limpias, como las formas de onda con ruido Gaussiano. Los resultados tanto del RMSE como del coeficiente  $R^2$  se presentan en la Tabla 3.4 para ambos *AST* y *AF*.

Para ambos enfoques, los resultados muestran valores de  $R^2$  cercanos a 1, aún en un  $SNR = 5$ . De las Tablas 3.5 y 3.6, se puede ver que los errores porcentuales son más fluctuantes en el caso *AST*, donde para el  $SNR = 5$  se ha alcanzado un error cercano al 13% en una zona de interpolación; mientras que para *AF* el error porcentual máximo es menor al 8%. Sin embargo, se puede ver que la predicción sobre señales limpias es relativamente mejor para *AST*.

La principal diferencia entre ambos enfoques, proviene de una menor dependencia del *AF* a la inclusión de ruido Gaussiano. Se puede observar que para el

*AST* las barras de confianza son más amplias para un  $SNR = 25$  que en el caso de *AF*. Sin embargo, en el *AST* la precisión es mejor que del *AF* para las señales limpias.

Métricas de error – Limpias - Ruido				
$SNR$	<i>AST</i>		<i>AF</i>	
	RMSE	$R^2$	RMSE	$R^2$
5	0.112	0.985	0.065	0.995
10	0.080	0.992	0.059	0.996
15	0.053	0.997	0.056	0.996
20	0.057	0.996	0.056	0.996
25	0.064	0.995	0.054	0.997
$\infty$	0.033	0.999	0.059	0.996

Tabla 3.4: RMSE y valores de  $R^2$  calculados sobre el conjunto de predicción por una RNA entrenada con señales limpias y probadas con señales con ruido. Utilizando el enfoque de *AST* el RMSE que se tiene cuando la RNA predice sobre señales limpias es casi la mitad del *AF*. Sin embargo, esto no sigue ocurriendo a medida que se incluye más ruido en las señales. A pesar de ello, los valores promedio del comité de RNA respecto a  $R^2$  son cercanos a 1.

De las Figuras 3.11 y 3.12 se puede apreciar como la inclusión del ruido Gaussiano ocasiona que las barras de los intervalos de confianza sean más amplias en comparación con la predicción sobre señales limpias tanto en el *AST* como en el *AF*. De esto podemos ver los resultados del *AF* son más consistentes en función del  $SNR$ . Esto es porque la transformada de Fourier reduce la sensibilidad al ruido Gaussiano.

### 3.4. Discusión del capítulo

A pesar de que en los dos catálogos estudiados en este capítulo, la muestra de patrones de entrenamiento es pequeña, se han obtenido para señales limpias y en la zona de interpolación predicciones con errores generalmente menores entre al 3%. De igual forma, en ambos casos con la inclusión de ruido Gaussiano en las formas de onda, las predicciones fluctúan como se esperaba, esto es, las RNAs



Error porcentual $AST$ – Limpias - Ruido						
$q$	$\infty$	25	20	15	10	5
1.15	3.038	7.748	2.526	9.536	4.170	2.749
1.25	0.075	3.784	1.647	5.768	10.870	3.317
1.75	1.855	8.473	2.357	0.277	3.878	12.740
2.05	0.809	0.428	2.507	0.320	4.977	9.457
2.25	0.721	0.792	2.272	0.688	1.223	1.171
2.35	1.438	0.745	3.100	1.337	4.514	2.355
3.4	0.034	0.461	2.694	1.930	0.977	0.199
3.9	0.636	0.216	1.674	0.211	1.385	2.176

Tabla 3.5: Error porcentual entre el valor predicho  $q$  y el real  $q_t$  (primer columna) para las señales en el conjunto de predicción usando el enfoque de  $AST$ . Los resultados incluyen tanto el caso de control donde las señales están libres de ruido (segunda columna), como para diferentes casos de  $SNR = 5, 10, 15, 20$  y  $25$ . Se puede ver como el error se incrementa en la mayoría de los casos a medida que se disminuye el  $SNR$ .

Error porcentual $AF$ – Limpias - Ruido						
$q$	$\infty$	25	20	15	10	5
1.15	2.859	2.371	0.761	4.389	2.997	5.0623
1.25	3.214	3.667	1.633	3.813	4.458	6.067
1.75	7.227	5.396	7.662	6.817	7.019	6.339
2.05	1.235	0.544	0.039	0.195	1.013	1.144
2.25	3.450	3.969	3.186	2.362	2.989	3.245
2.35	0.106	0.424	0.185	1.831	0.195	0.838
3.4	0.956	1.042	0.979	0.890	1.589	1.468
3.9	1.062	1.056	0.642	0.453	0.688	1.435

Tabla 3.6: Error porcentual entre el valor predicho  $q$  y el real  $q_t$  (primer columna) para las señales en el conjunto de predicción usando el enfoque de  $AF$ . En la segunda columna se muestra el error obtenido para las señales limpias y en el resto a diferentes  $SNR$ . A diferencia del caso  $AST$  mostrado en la tabla 3.5, el error se mantiene prácticamente del mismo orden para todos los  $SNR$ . Por lo que se puede concluir que este análisis es menos sensible a la inclusión de ruido Gaussiano.

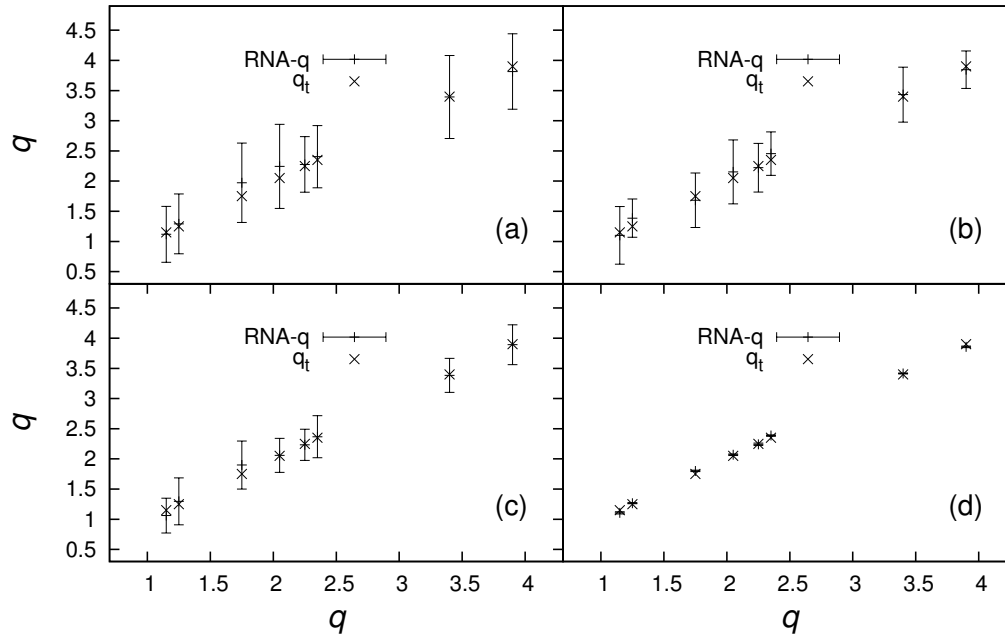


Figura 3.11: Estimación de  $q$  analizando las GWs como series de tiempo (*AST*). En este caso, las razones de masas reales del sistema ( $q_t$ ) son representadas con  $\times$  y las predicciones por RNA- $q$ . En (a), (b) y (c) se muestran los resultados para las formas de onda del conjunto de predicción que han sido contaminadas con ruido Gaussiano con un  $\text{SNR} = 5, 10$  y  $25$  respectivamente. En los tres casos las barras de confianza son mucho más amplias que las del caso de control (d), donde las señales del conjunto de predicción no tienen ruido Gaussiano alguno, aunque en algunos de los casos  $q_t$  está fuera del rango de las barras de confianza, que calculadas siguiendo la ecuación (3.8).

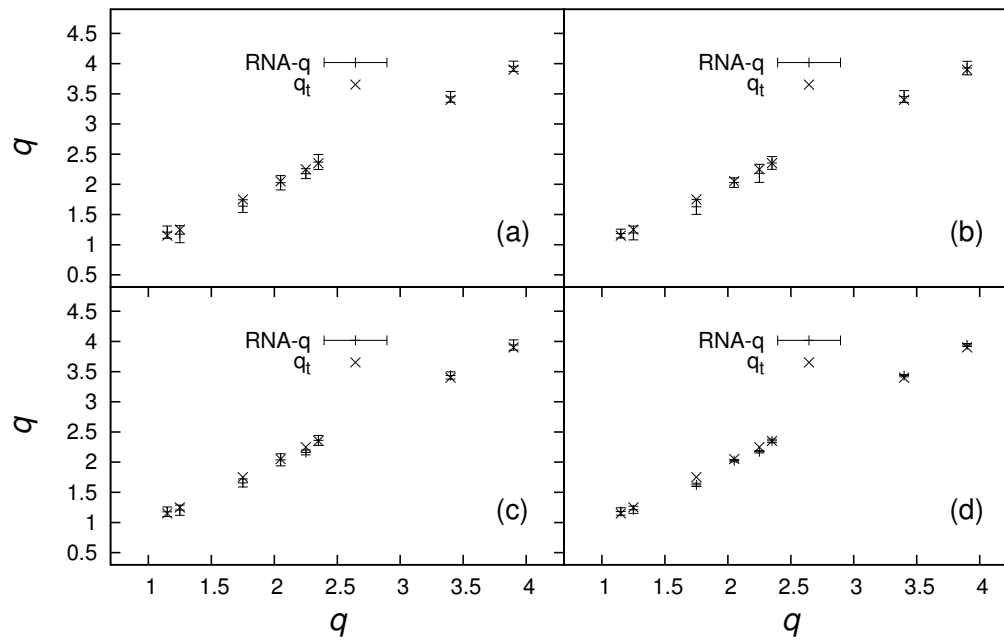


Figura 3.12: Valores de  $q$  predichos en  $AF$  por el comité de RNAs tanto para señales con ruido Gaussiano con  $SNR = 5, 10$  y  $25$  en (a), (b) y (c) respectivamente. En el caso de control para las señales sin ruido (d) se observa que las barras de confianza son notablemente más angostas que el resto de los casos de estudio. Sin embargo, la verdadera razón de masas  $q_t$  (representada con  $\times$ ) no se encuentra dentro de ellas en algunos casos. Esto se puede deber principalmente a la escasa base de datos para entrenar las RNAs. También se puede notar que dichos márgenes de error no se incrementan considerablemente al incrementar el ruido, al menos en para los rangos de SNR que se han manejado.

tienen una tolerancia hasta cierta magnitud de ruido con predicciones menores al 10% de error hasta llegar a una  $SNR \leq 10$ . Para  $SNR$  aún menores el error porcentual puede crecer tan grande como el 50%, como en algunos casos aislados cuando se han analizan las formas de onda como series de tiempo y en la zona de extrapolación.

Es de notar que el uso de un análisis en el espectro de frecuencias es bastante útil para disminuir la repercusión del ruido en las formas de onda, al menos del ruido Gaussiano empleado en ambas secciones, que puede ser de utilidad para aplicaciones posteriores. Los resultados descritos aquí, dan la confianza a analizar la implementación de RNAs o algún otro algoritmo de aprendizaje automático, para encontrar los valores intrínsecos en GWs, como aquellos para otro tipo de sistemas BBH anti-alineados de distintas magnitudes o en precesión.



## Capítulo 4

# Uso de RNAs en Problemas Inversos en el Flujo de un Fluido Clásico Alrededor de un Cilindro en un Tubo de Dos Dimensiones

Este capítulo se basa en la aplicación de RNAs en el área de dinámica de fluidos, que ha sido publicado en *Estimation of Reynolds number for flows around cylinders with lattice Boltzmann methods and artificial neural networks* [Carrillo, Que & González, (2016)]; y en *Recognition of an obstacle in a flow using artificial neural networks* [Carrillo, Que, González & López, (2017)], publicados ambos en *Physical Review E*.

En el trabajo realizado se ha aprovechado la capacidad de las RNAs para el reconocimiento de patrones generados en el flujo de un fluido, en específico para el flujo en un tubo de dos dimensiones en el que se encuentra inmerso un obstáculo cilíndrico. El capítulo aborda dos problemas: 1) la estimación de un parámetro que caracteriza el flujo en este tipo de sistemas, es decir, el número de Reynolds y 2) el reconocimiento del tamaño y posición del obstáculo.

Para entender el fenómeno del flujo alrededor de un obstáculo circular en una tubería de dos dimensiones y del método de simulación utilizado, primeramente se

revisarán algunas cuestiones generales sobre los dos temas.

## 4.1. Flujo alrededor de un cilindro en una tubería dos dimensional

En forma simple, este problema se describe por la presencia de un cuerpo inmerso en un fluido hidrodinámico con un flujo libre de fuerzas externas, donde el cuerpo ha experimentar los efectos del desplazamiento del fluido al pasar alrededor de él. Este fenómeno tiene múltiples aplicaciones desde el estudio de las propiedades mecánicas de espumas [Dollet, Durth & Graner, (2006)], aerodinámica de estructuras como puentes o alas [Larsen & Walther, (1998)]; patrones de ondas generadas en el flujo atmosférico [Lacaze et al.,(2013)], entre otras. Siendo interesante el estudio en los últimos dos por los efectos de las capas límite y zonas de recirculación presentes en el flujo.

La dinámica del fluido se puede categorizar en dos tipos: por tener un flujo laminar o uno turbulento. Comúnmente se busca que el flujo en un canal o tubería sea del tipo laminar tal que se evite la pérdida de energía o esfuerzos dentro del conducto que produzcan defectos sobre la superficie. Estos dos tipos de flujo y sus estados intermedios se pueden cuantificar en términos de un parámetro adimensional: el número de Reynolds ( $Re$ ). Este cuantifica los patrones de flujo similares en un problema específico y en términos de la razón entre las fuerzas inerciales y viscosas que experimenta el fluido. Para el caso particular de este trabajo de tesis, se ha considerado un flujo entrante (anterior al obstáculo) del tipo Poiseuille dentro de una tubería de dos dimensiones. Hay que remarcar que para el flujo de Poiseuille, el  $Re$  se puede definir en términos de la velocidad máxima ( $v_{\max}$ ) del perfil de velocidades a través de la expresión:

$$Re = \frac{v_{\max} L}{\nu}, \quad (4.1)$$

donde  $L$  es el ancho de la tubería y  $\nu$  es la viscosidad cinématica del fluido. Ver por ejemplo la Figura 4.1 donde se han gráfico distintos flujos de Poiseuille en un tubo de dos dimensiones de 0.5m de ancho con distintas velocidades entrantes

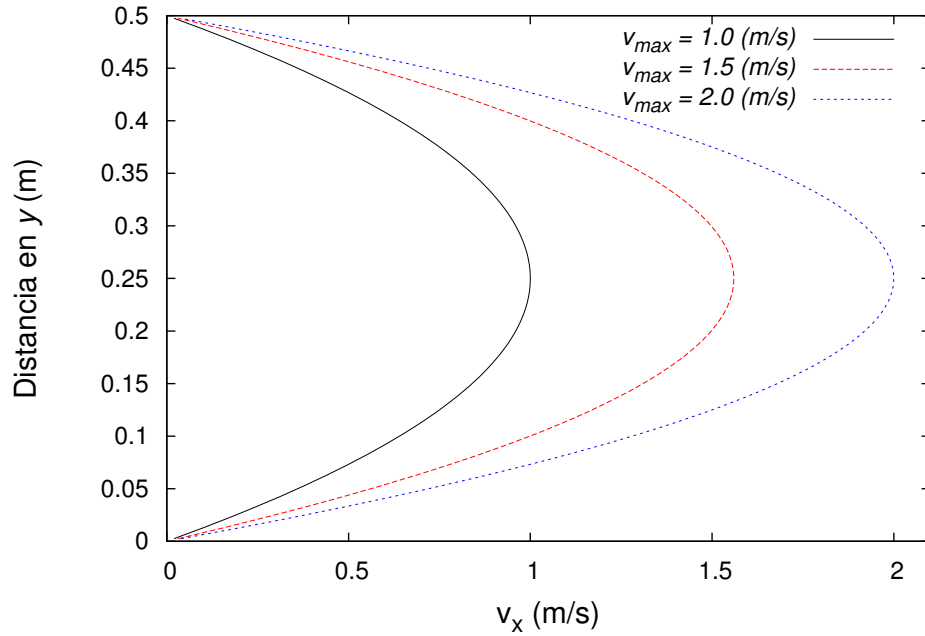


Figura 4.1: Flujos de Poiseuille con distintas velocidades entrantes en un tubo de 0.5m de ancho. Los diferentes flujos están fueron gráficos en función de sus velocidades características  $v_{\max} = 1.0(\text{m/s})$ ,  $1.5(\text{m/s})$  y  $2.0(\text{m/s})$  por las líneas negra, roja y azul respectivamente. Se ve claramente como cambian los perfiles de velocidad  $v_x$  en función de la velocidad en función de  $v_{\max}$ , lo que también determina la magnitud del  $Re$ .

caracterizadas por  $v_{\max} = 1.0, 1.5$  y  $2.0$  m/s. Para los casos específicos que se estudian en las siguientes secciones el  $Re$  del flujo después del obstáculo de forma circular se determina por:

$$Re = \frac{v_c l_c}{\nu}, \quad (4.2)$$

donde  $v_c$  es la llamada velocidad característica, relacionada con el flujo de Poiseuille por  $v_c = 2/3v_{\max}$ , y  $l_c$  es la longitud característica determinada por el diámetro del obstáculo.

Como se mencionó, la caracterización del flujo se determina por los valores de  $Re$ . Por ejemplo, para un  $Re < 40$  la presión del fluido aumenta desde el flujo libre hasta el punto de estancamiento, es decir, la zona donde el fluido golpea el obstáculo y fluye alrededor de este. Al incrementar la velocidad inicial sobre el tubo, el incremento de presión sobre el obstáculo restringe el movimiento del



fluido a lo largo de la superficie del cilindro, generando capas límite en ambos lados y provocando dos capas rasantes justo después del cilindro separando el flujo por ambos lados. Ver por ejemplo la Figura 4.2a donde se grafica el valor de la componente  $x$  del campo de velocidades  $\mathbf{v}$  o  $v_x$  en un corte transversal sobre el tubo para un  $Re = 30$ . Para  $Re$  más altos, las capas rasantes se enrollan en sí mismas generando vórtices rotando en sentidos contrarios. A medida que el flujo del fluido supera el valor de  $Re = 60$ , la estela generada posterior al cilindro se vuelve inestable, produciendo el fenómeno conocido como *calle de vórtices de Karman*, caracterizado por una proyección alternada y aperiódica de vórtices. En la Figuras 4.2b y 4.2c donde se muestra este fenómeno graficando tanto  $v_x$  como la componente  $z$  de la vorticidad  $(\nabla \times \mathbf{v})_z$ . En este fenómeno los vórtices generados son acarreados por la corriente del fluido circundante disipándose conforme estos avanzan a lo largo de la tubería.

El caso del flujo alrededor de un obstáculo de un fluido incompresible en dos dimensiones es un problema investigado en distintos trabajos en el área de la dinámica computacional de fluidos (CFD) [Catalano, et al., (2003), Grucelski & Pozorski, (2013), Karabelas et al., (2012)]. Para la elaboración de este trabajo se utilizó un código numérico específico para la simulación de la dinámica de fluidos: el método de *lattice Boltzmann* (LBM). Con las simulaciones elaboradas con este código se ha podido explorar el comportamiento del flujo bajo la variación de los distintos parámetros: diámetro del obstáculo, velocidad característica entrante del fluido y posiciones del obstáculo. Con esto en mente, se ha decidido medir el rendimiento de las RNA en la rama de la dinámica de fluidos como un identificador de patrones del flujo para dos problemas distintos:

1. Estimación del  $Re$  para un flujo alrededor de un cilindro con distintas velocidades entrantes.
2. Estimación de la morfología y posición del obstáculo.

Los principios básicos de implementación del modelo de LBM se menciona en el apéndice D. No obstante, por razones de autocontenido se dará aquí una breve

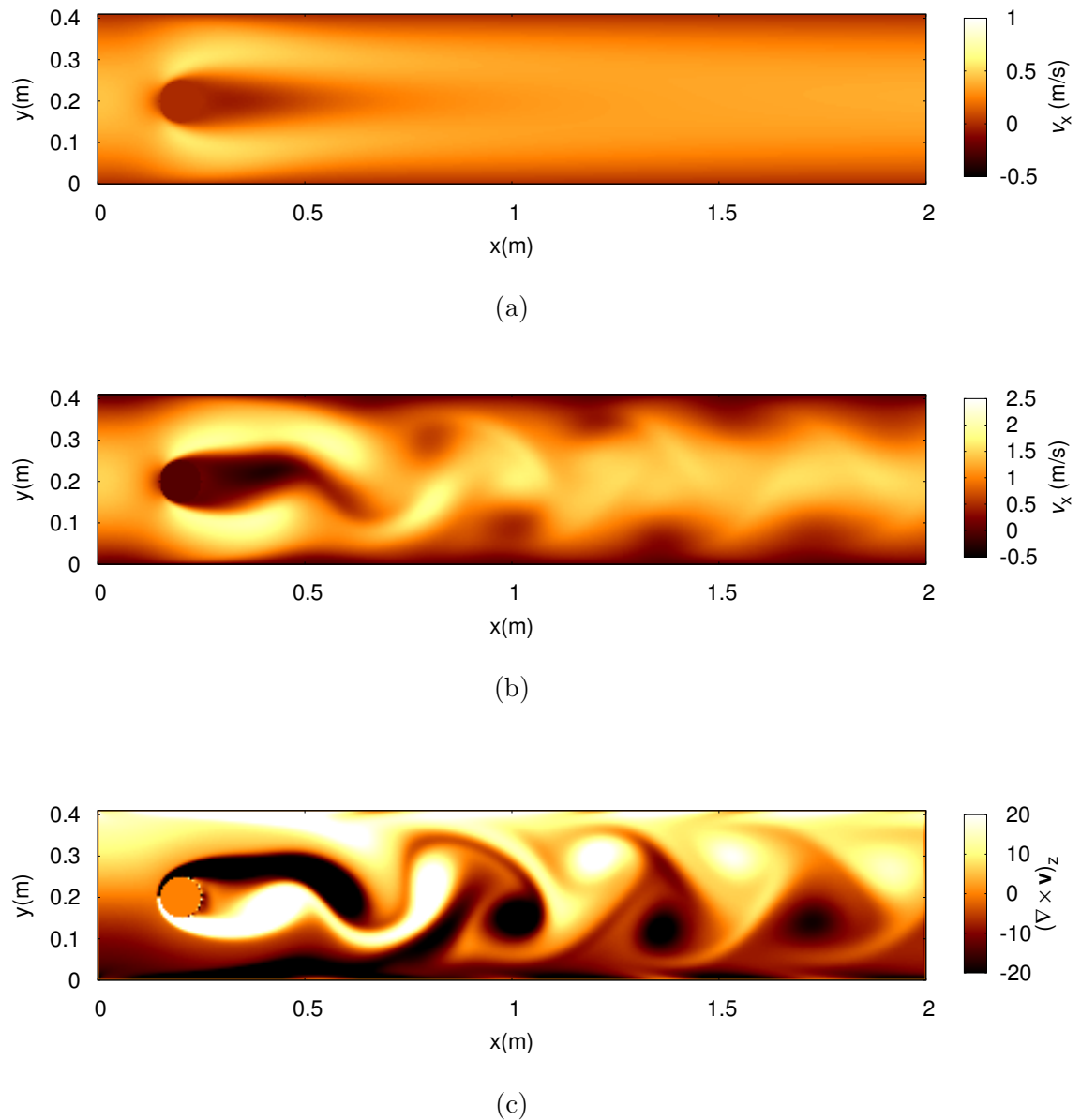


Figura 4.2: El obstáculo es perpendicular al plano  $x - y$ , por lo que se representa como un círculo, tanto en las simulaciones como en las gráficas presentadas. En (a) se gráfica  $v_x$  para un flujo con un  $Re = 30$ , se observa como la perturbación provocada por el obstáculo es menos notoria conforme el fluido se aleja de este, siendo apenas perceptible al final del dominio. En (b) la alteración del flujo es evidente para este caso, donde el flujo se caracteriza por un  $Re = 100$ , donde se pueden apreciar los vórtices provocados por el obstáculo. Finalmente en (c) se muestra la simulación para un flujo con  $Re = 100$ , pero graficando en este caso  $(\nabla \times \mathbf{v})_z$ . Las tres figuras se muestra el corte en el plano  $x - y$ .

descripción del método y en cada subsección se remarcan los parámetros utilizados en las simulaciones.

#### 4.1.1. Método de lattice Boltzmann

En LBM los fluidos se consideran como una gran colección de pequeños elementos del fluido, reduciendo las posiciones espaciales y momentos de todas las partículas que componen el fluido, en una discretización del tiempo. Bajo esta consideración, el método resuelve la dinámica del flujo del fluido a través de la ecuación de Boltzmann la cual considera el intercambio de energía y momento debido al flujo y colisión del ensamble de las partículas del fluido por medio de una función de distribución  $f$  como:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f = \Omega(f) \quad (4.3)$$

donde  $\mathbf{v}$  es la velocidad del ensamble y  $\Omega$  es llamado el operador de colisión, que se define como el cambio entre el estado inicial y final de la distribución  $f$ , el cual debe ser determinado para poder resolver el sistema. Al considerar la discretización de la ecuación de Boltzmann (4.3), se confinan los elementos de fluido en celdas en un arreglo tipo red, que de ahora en adelante se denominará como malla numérica (*lattice* por su nombre en inglés). En dichas celdas, las variaciones del momento que en el continuo debidas a una infinidad de direcciones de velocidad y cantidad de masa se reducen a un número fijo de posibilidades por una única partícula. Es decir, que el ensamble de partículas dentro de la celda se pueden describir por una función de densidad de distribución  $f$ , con una masa uniforme en tiempos discretos. Ver por ejemplo la Figura 4.3, donde se representa una malla de simulación en donde se representa el modelo de celda D2Q9, usada para este trabajo, en la cual se definen 9 posibles direcciones  $\mathbf{c}_i$  en dos dimensiones.

En cada una de los nodos del arreglo se preservan las cantidades físicas como la velocidad y densidad que son calculadas a partir de la función de distribución  $f$ . Por ejemplo, la densidad dependiente de la posición y el tiempo se calcula como:

$$\rho(x, t) = \sum_i f_i, \quad (4.4)$$

mientras que la velocidad  $\mathbf{v}(x, t)$  se define por:

$$\mathbf{v}(x, t) = \frac{1}{\rho} \sum_i f_i \mathbf{c}_i. \quad (4.5)$$

Para mayor detalle sobre la parte teórica del modelo o método de LBM, checar el apéndice D de este trabajo o también se puede consultar otras referencias como [Mohamad, (2011), Sukop & Thorne, (2010)], para dar paso al empleo de dicha técnica en los problemas antes mencionados.

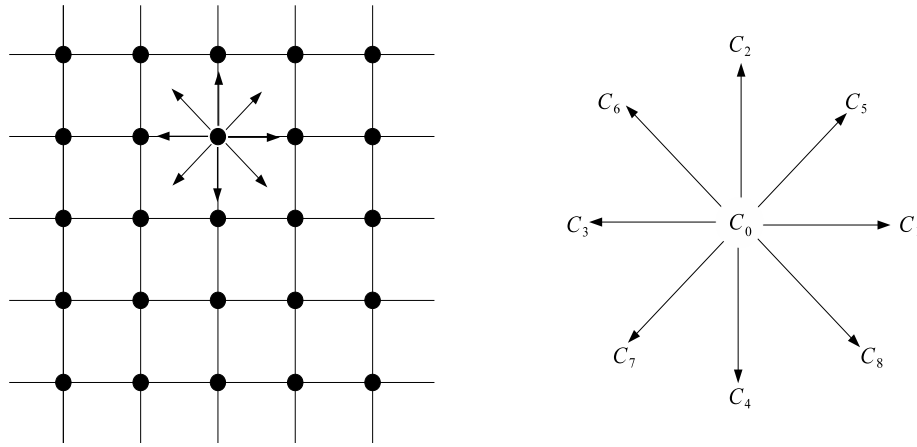


Figura 4.3: Ejemplificación de la malla numérica utilizada en LBM. El sistema físico se ha dividido en un arreglo de celdas y donde uno de los nodos son los puntos donde se extraen los valores físicos (izquierda). En esta figura se ejemplifica el modelo para el mallado utilizado en este trabajo, el llamado D2Q9, en el cual existen 9 posibles direcciones denotadas en este caso por  $\mathbf{c}_i$  con  $i = 0, 1, 2, \dots, 8$ , para el conjunto de elementos del fluido representado por la función de distribución  $f$  (derecha). Para una explicación más profunda revisar apéndice D.

## 4.2. Estimación del número de Reynolds en el flujo de un fluido alrededor de un cilindro

Esta sección se basa en el artículo *Estimation of Reynolds number for flows around cylinders with lattice Boltzmann methods and artificial neural networks* publicado en la revista *Physical Review E*.

Como se mencionó anteriormente, una forma de caracterizar un flujo es a través de la estimación del número de Reynolds. En esta sección del trabajo se describirá el desarrollo e implementación de una RNA bajo un entrenamiento supervisado que usa por un lado la componente  $x$  del perfil del campo de velocidades ( $v_x$ ) y por el otro la componente  $z$  del perfil de la vorticidad ( $(\nabla \times \mathbf{v})_z$ ) en algún punto del tubo como información de entrada a la RNA, mientras que el  $Re$  fungirá como objetivo para el entrenamiento.

Continuando con la descripción del flujo del fluido alrededor de un cilindro para los problemas planteados, se ha generado una base de datos con las simulaciones utilizando el método de LBM. Siguiendo las condiciones del sistema planteadas en las simulaciones del trabajo [Schäfer & Turek, (1996)], se ha establecido un tubo de dos dimensiones con un largo de  $L_x = 2$  m, ancho de  $L_y = 0.41$  m, y considerando distintas velocidades de flujo entrante tipo Poiseuille con una densidad  $\rho = 10^3$  kg/m<sup>3</sup> y una viscosidad cinemática de  $\nu = 10^{-3}$  m<sup>2</sup>/s. Revisar nuevamente la Figura 4.2, donde dichas simulaciones se hicieron siguiendo las especificaciones del fluido mencionadas. La base de datos de las simulaciones se generó considerando 120 valores de velocidades de flujo entrante, lo que supone a su vez, por lo establecido en la ecuación 4.2, 120 números de Reynolds que van desde  $Re = 1$  hasta  $Re = 120$  con incrementos de  $\Delta Re = 1$ . En cuanto a la simulación, se implementó una malla numérica con 164 nodos para el ancho de la tubería y 820 nodos para representar el largo del dominio, donde el único parámetro libre de entrada para cada simulación es el perfil de velocidad de entrada parametrizada por  $v_c$ . Recordando que para realizar la estimación del número de  $Re$  mediante la RNA, se extrajeron dos diferentes valores físicos: con los perfiles sobre el eje  $y$  de la componente  $x$  del perfil del campo de velocidad ( $v_x$ ) y la componente  $z$  de la vorticidad ( $(\nabla \times \mathbf{v})_z$ ). Ambas variables fueron extraídas de la simulación una vez que el flujo del fluido alcanzó una estabilidad neutral, es decir, un comportamiento regular del flujo del fluido en el tiempo. Dicha estabilidad se alcanza desde los 20 hasta 50 segundos en el tiempo físico, dependiendo del valor del  $Re$ .

### 4.2.1. Procesamiento datos

Los valores de  $v_x$  y  $(\nabla \times \mathbf{v})_z$  se extrajeron en cinco distintas ubicaciones posteriores al obstáculo cilíndrico, esto es a  $x = 0.3, 0.5, 0.7, 1.1$  y  $1.9$  m. Estas ubicaciones se llamarán de ahora en adelante puntos de medición y su posición relativa es tal y como se muestra en el esquema dibujado en la Figura 4.4.

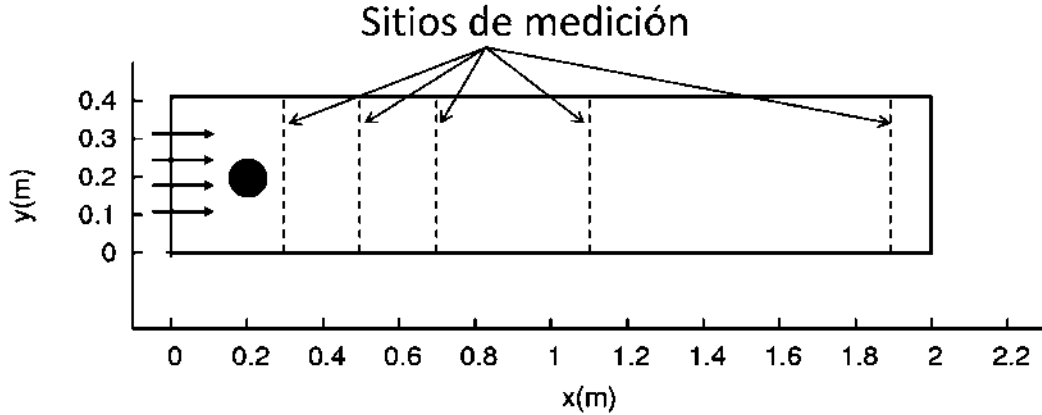


Figura 4.4: Representación esquemática del flujo alrededor de un cilindro situado en  $x = 0.2$  m y  $y = 0.2$  m. El fluido se mueve de la izquierda a la derecha y las mediciones se realizan cuando el flujo alcanza una estabilidad neutral. Los sitios de medición se encuentran en  $x = 0.3, 0.5, 0.7, 1.1$  y  $1.9$  m, donde se han extraído los valores de los perfiles de  $v_x$  y de la componente  $z$  de la vorticidad en 1,4,6,11,21,41 y 82 nodos de la malla numérica.

El primer punto de medición en  $x = 0.3$  m se localiza inmediatamente detrás del obstáculo, de tal forma que se mida la capacidad predictiva de la RNA donde se generan inicialmente los vórtices. Los siguientes tres sitios de medición se encuentran en las regiones donde se observa el fenómeno de vórtices de Karman. Finalmente, el último sitio en  $x = 1.9$  m se encuentra alejado del obstáculo de tal forma que se puede medir la precisión de predicción de la RNA para ubicaciones en las cuales es posible que la vorticidad disminuya en magnitud en comparación a las demás posiciones. Ejemplificación de esto, se puede observar en la Figura 4.5 donde se han graficado diferentes perfiles de  $v_x$  y de vorticidad  $(\nabla \times \mathbf{v})_z$ , para  $Re = 100$  en tres distintas posiciones a  $x = 0.3, 1.1$  y  $1.9$  m. Mientras que en la

Figura 4.6, se observan los valores de  $v_x$  y  $(\nabla \times \mathbf{v})_z$  tal y como se extraen en el primer punto de medición.

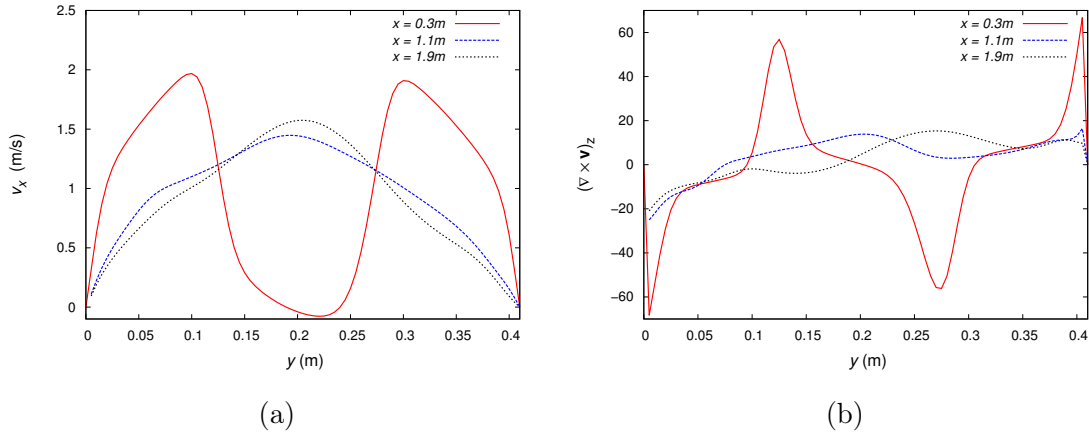


Figura 4.5: Ejemplo de las mediciones obtenidas en diferentes ubicaciones a lo largo del tubo para  $Re = 100$ . Se grafican tanto  $v_x$  (a) como  $(\nabla \times \mathbf{v})_z$  (b) contra la posición en  $y$  en cada sitio de medición en  $x = 0.3\text{m}$ ,  $1.1\text{m}$  and  $1.9\text{m}$ .

Como se refirió en el capítulo 2, una de las cuestiones en utilizar RNAs, es definir las cantidades o propiedades adecuadas que hay que utilizar como valores de entrada. Dado que en este caso se ha decidido utilizar los valores de los perfiles de  $v_x$  y  $(\nabla \times \mathbf{v})_z$ , la cantidad de puntos está restringida por el número de nodos de la malla numérica del modelo de LBM utilizado, en este caso 165. Sin embargo, el considerar tantos puntos de medición parece innecesario, así que en este caso en particular se ha investigado cual es la precisión de la RNA al variar también el número de valores del perfil de  $v_x$  y  $(\nabla \times \mathbf{v})_z$  a lo largo del eje  $y$  en un sitio de medición. Con esto se reducen los valores considerados de los 164 nodos en la malla numérica de LBM para 7 distintos casos considerando únicamente el valor de 4, 6, 11, 21, 41 y 82 nodos equidistantes y por último el valor central de un único nodo en el eje  $y$ .

Ya se que se considere  $v_x$  o  $(\nabla \times \mathbf{v})_z$ , las RNA construídas tienen una sola capa oculta con 10 neuronas dispuestas con una función de activación del tipo tangente hiperbólica. Dado que se está tratando con un problema de regresión, la RNA cuenta con una única neurona de salida con una función lineal que determina el  $Re$ . En concordancia con la ecuación 2.7, la predicción del número de Reynolds

por parte de la RNAs es de la forma:

$$Re = \sum_{j=1}^m w_{jk} \tanh \left( \sum_{i=1}^n w_{ij} (v_{x,y=i}) + b_j \right) + b_0, \quad (4.6)$$

donde  $n = 1, 4, 6, 11, 21, 41, 82$  es el número de entradas,  $m$  el número de neuronas ocultas y  $v_{x,y=i}$  define el valor de  $v_x$  sobre el eje  $y$  en la posición  $i$ -ésima correspondiente.

En términos del entrenamiento, el conjunto de datos extraídos en las simulaciones se han dividido en los tres subconjuntos de entrenamiento, validación y predicción de la siguiente forma:

- Entrenamiento: 80 simulaciones muestreadas en un rango de  $mRe \in [1, 116]$ , exceptuando aquellos del conjunto de predicción que se mencionan a continuación.
- Validación: 20 simulaciones, diferentes al conjunto de entrenamiento con  $Re$  desde 6 a 117.
- Predicción: 20 simulaciones, diferentes al conjunto de entrenamiento y predicción, con  $Re = 12, 17, 22, 27, 32, 52, 64, 70, 76, 92, 99, 102, 107, 112, 116, 117, 118, 119$  y 120. Los primeros 15 son muestreados dentro del rango de entrenamiento, lo que se interpretaría como una interpolación en el rango de la regresión. Mientras que los últimos 5 se encuentran fuera de dicho rango, por lo cual se espera un decremento en la precisión de la predicción.

En este caso el algoritmo de *backpropagation* tiene valores de  $\gamma = 0.05$  y momento  $\alpha = 0.5$  entrenándose a la RNA hasta que el error de validación incrementa, guardando los pesos con el menor error en el conjunto de validación. Con estas consideraciones, los entrenamientos han tenido una duración menor a los 3 minutos con 1000 a 2000 iteraciones.

## 4.2.2. Resultados

En las Tablas 4.1 y 4.2 se muestra el RMSE del promedio de las predicciones del comité de las 10 RNAs entrenadas en cada sitio de medición y para todos los



casos del conjunto de predicción. Por un lado, cuando se considera el perfil de  $v_x$  como vector de entrada el error se incrementa notablemente hasta que el número de sensores de medición disminuye por debajo de 11, tal y como se muestra en la Figura 4.7a. Por otro lado, la misma consideración utilizando  $(\nabla \times \mathbf{v})_z$  los errores son mayores y menos regulares, con un RMSE similar a partir de utilizar 41 sensores tal y como se muestra en 4.7b.

De la Figura 4.8, se puede ver que las RNAs utilizando 82 puntos de muestreo, tienen un error relativo menor al 4% en todos los sitios de medición. De la misma figura se puede decir que Los resultados son más precisos utilizando  $v_x$  para los sensores que se localizan más alejados del obstáculo y para valores altos de  $Re$ . Por el contrario, utilizando  $(\nabla \times \mathbf{v})_z$  da mejores predicciones para valores bajos de  $Re$  y a distancias más cercanas al cilindro, decreciendo para los sitios de medición en  $x = 0.7$  y  $1.1\text{m}$ , tal vez debido a que las estructuras de los vórtices son más difusas a tal distancia. Para los casos de extrapolación con  $Re = 115, 116, 117, 118, 119$  y  $120$ ; la precisión del número de Reynolds decrece pero el error no excede el 2% y 4% para  $v_x$  y  $(\nabla \times \mathbf{v})_z$  respectivamente.

En las Tablas 4.3 y 4.4 se muestran los promedios de las estimaciones hechas por el comité de 10 RNAs considerando 82 detectores en los distintos puntos de medición para todos los  $Re$  en el conjunto de predicción y comparando el rendimiento en cada sitio utilizando la prueba  $\chi^2$ . Cuando se considera  $v_x$  como información de entrada a la RNA, el peor resultado se obtiene en el sitio  $x = 0.5\text{m}$ , mientras que utilizando valores de  $(\nabla \times \mathbf{v})_z$  esto sucede tanto en  $x = 0.7\text{m}$  y  $1.1\text{m}$ , corroborando lo visto en la Figura 4.8.

Hasta ahora, los resultados se obtienen considerando la predicción en un instante de tiempo de extracción igual para todos los casos. Sin embargo, se puede inspeccionar si estos resultados se mantienen con mediciones de  $v_x$  y  $(\nabla \times \mathbf{v})_z$  en distintos instantes de tiempo al que se entrenó la RNA. Este análisis se muestra para los sitios de medición en  $x = 0.3, 0.5$  y  $0.7\text{m}$  en las Figuras 4.9 y 4.10 donde se han empleado las simulaciones con números de Reynolds  $Re = 30$  y  $90$  respectivamente. Del primer conjunto de gráficos se puede ver que al tratarse de un flujo con un  $Re$  bajo, las estimaciones alcanzan un valor constante a partir de cierto tiempo.

Mientras que en el par de gráficos de la Figura 4.10, las predicciones oscilan a lo largo del tiempo, esto es de esperarse dada la frecuencia de la generación de los vórtices en el flujo del fluido. A pesar de ello, las estimaciones se mantienen en el mismo orden de magnitud que los casos similares ya revisados.

### Conclusiones de la sección

De los resultados se puede concluir que utilizar  $v_x$  como información de entrada para la RNA es mejor que  $(\nabla \times \mathbf{v})_z$  en la mayoría de los casos; excepto para  $Re$  pequeños y cercanos al obstáculo. En términos de la estimación del número de Reynolds en el régimen de extrapolación, el error incrementa a medida que se aleja de la simulación con el  $Re$  más alto para el cual fueron entrenadas las RNAs. Las zonas más complicadas para las predicciones parecen ser en la región media del tubo, esto es en  $x = 0.5, 0.7$  y  $1.1\text{m}$ , probablemente se deba al comportamiento más dinámico del fluido. Se sospecha que a medida que se modifica la velocidad del flujo entrante, las magnitudes del campo de velocidad también incrementan. De tal manera que esto puede ayudar a que las RNAs encuentren patrones más claros que relacionan los perfiles tanto de  $v_x$  como  $(\nabla \times \mathbf{v})_z$  con su respectivo número de Reynolds. Otro resultado que hay que hacer notar, es el que número mínimo de sensores que se pueden utilizar para obtener resultados confiables es a partir de 11.

Haciendo un análisis del flujo del fluido en distintos instante de tiempo se obtienen buenas estimaciones del  $Re$  hasta se que llega a la estabilidad neutral. Los resultados indican que para escenarios más generales con distintos tipos de obstáculos la dependencia en el tiempo y espacio puede ser más complicada. Con esta sección del trabajo se espera que el rendimiento pueda ser incrementado si las RNAs son entrenadas con patrones de flujo en el tiempo y espacio, tal y como se inspeccionará en la siguiente sección.

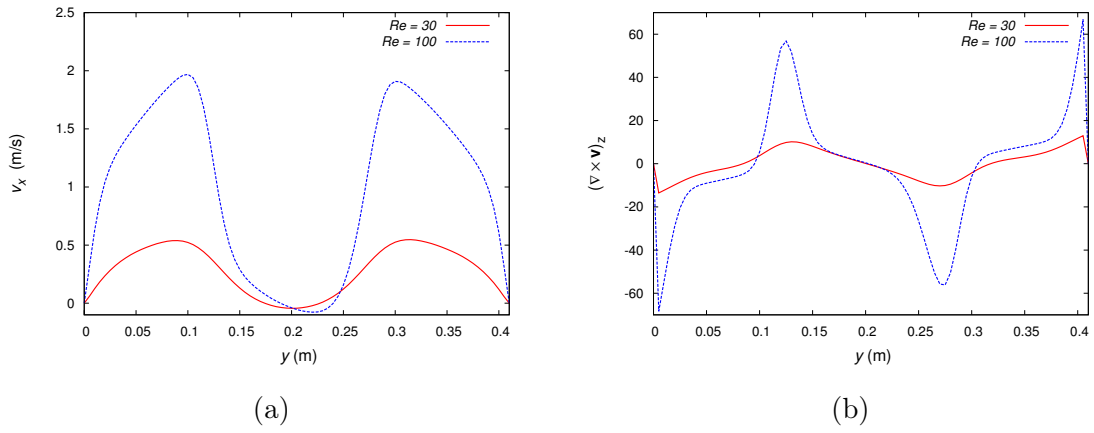
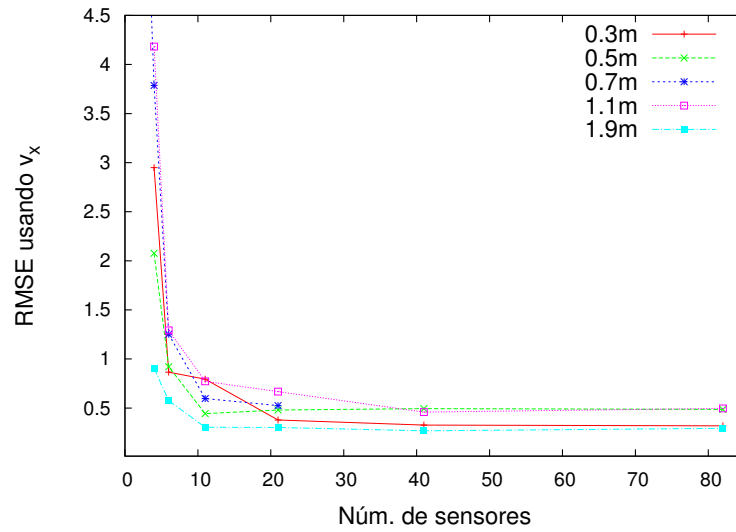


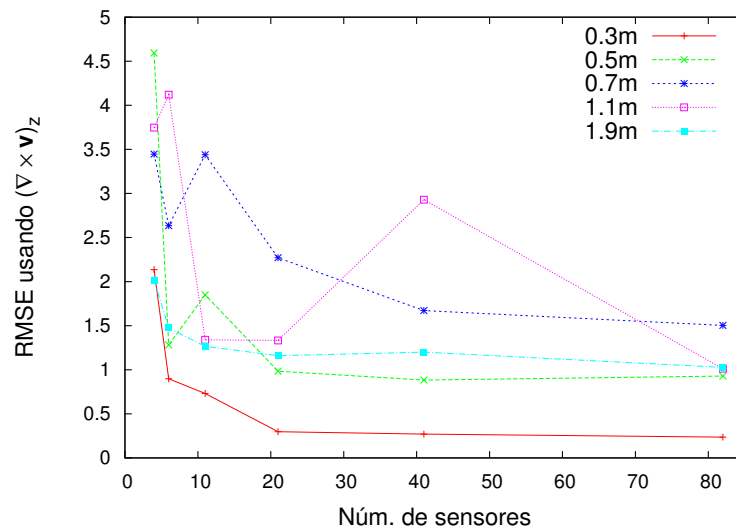
Figura 4.6: Se grafican los perfiles de los valores de  $\mathbf{v}_x$  (a) y  $(\nabla \times \mathbf{v})_z$  (b), y en la derecha para dos distintos números de Reynolds  $Re = 30$  y  $Re = 100$  en  $x = 0.3\text{m}$ . A diferencia de los valores graficados en la Fig. 4.5, es de notarse la gran diferencia en magnitud que existen en ambos casos para los dos distintos números de Reynolds.

RMSE usando $v_x$					
	0.3m	0.5m	0.7m	1.1m	1.9m
1	36.652	43.596	9.5976	3.931	3.446
4	1.950	2.077	3.786	1.880	0.904
6	0.421	0.921	1.252	0.528	0.575
11	0.366	0.443	0.597	0.370	0.304
21	0.141	0.479	0.525	0.228	0.302
41	0.133	0.494	0.488	0.227	0.268
82	0.132	0.486	0.486	0.220	0.294

Tabla 4.1: RMSE del promedio de las predicciones del comité de RNAs en el sitio de medición a lo largo del tubo en  $x = 0.3, 0.5, 0.7, 1.1$  y  $1.9$  m; usando  $v_x$  para 1, 4, 6, 11, 21, 41 y 82 sensores numéricos. Es de notar como el RMSE se incrementa dramáticamente considerando menos de 6 puntos en el muestreo.

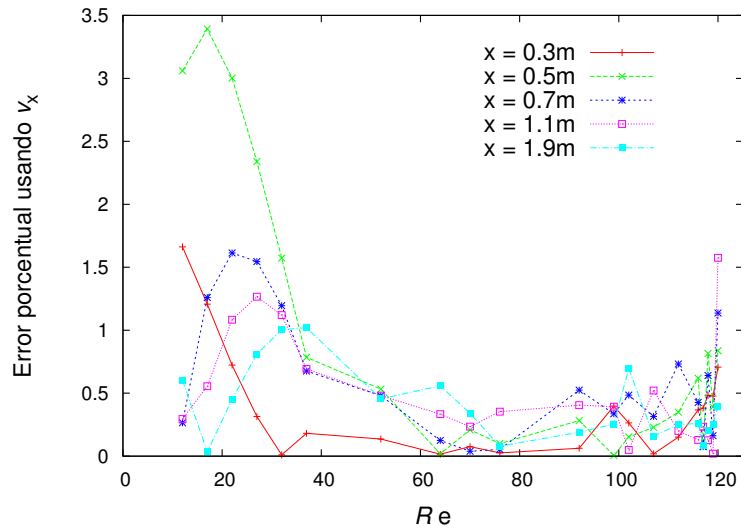


(a)

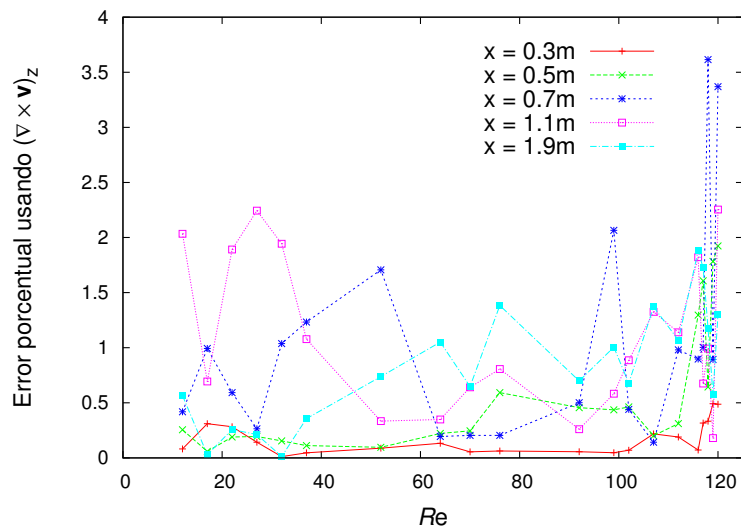


(b)

Figura 4.7: Raíz del error cuadrático medio (RMSE) en el conjunto de predicción para distintos lugares de medición en  $x = 0.3, 0.5, 0.7, 1.1$  y  $1.9$  m; con 4, 6, 11, 21, 41 y 82 sensores de medición usando tanto  $v_x$  o  $(\nabla \times \mathbf{v})_z$ .

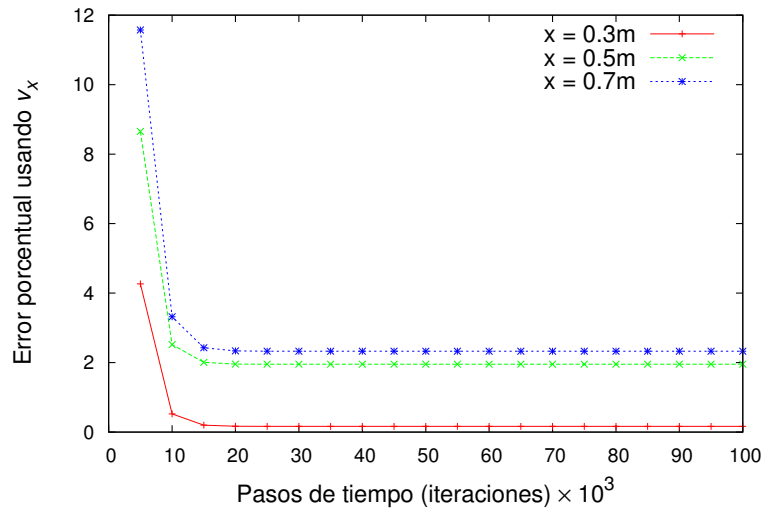


(a)

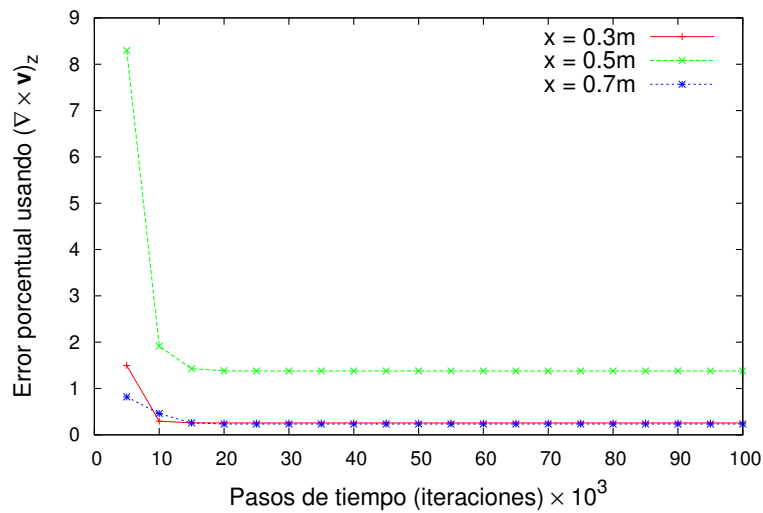


(b)

Figura 4.8: Error porcentual en las estimaciones de  $Re$  en el conjunto de predicción para los distintos lugares de medición en  $x = 0.3, 0.5, 0.7, 1.1$  y  $1.9$  m, usando 4, 6, 11, 21, 41 y 82 sensores de medición usando  $v_x$  (a) y  $(\nabla \times \mathbf{v})_z$  (b).

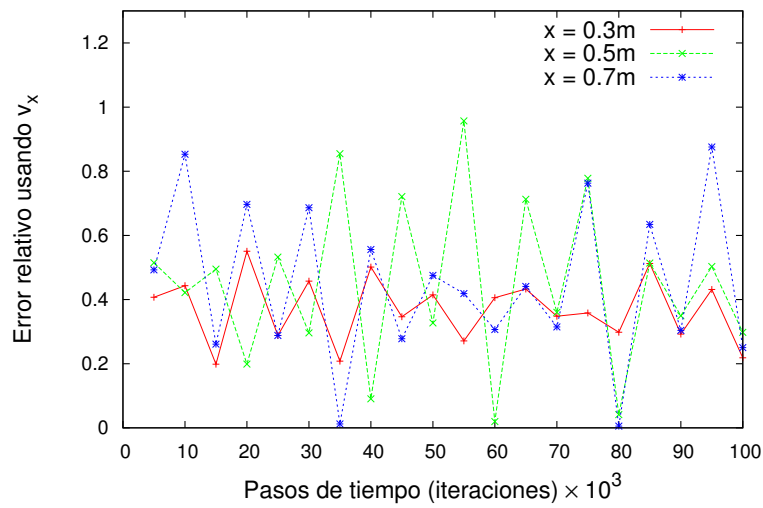


(a)

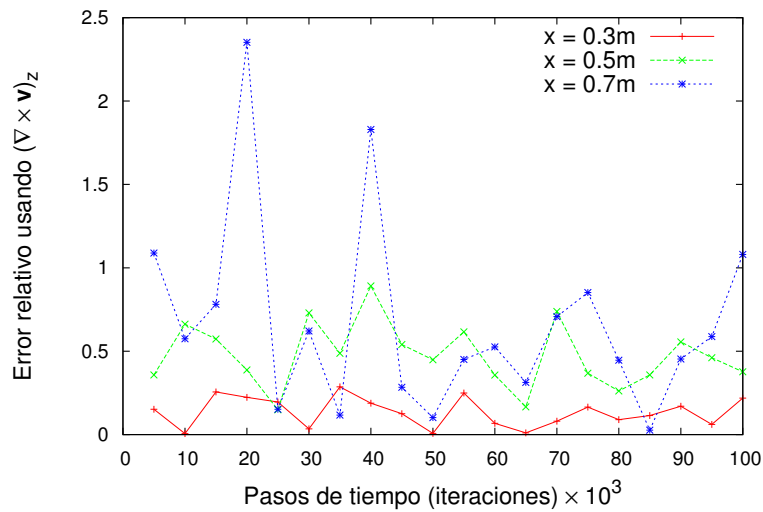


(b)

Figura 4.9: Porcentaje de error en el conjunto de predicción para los distintos lugares de medición:  $x = 0.3$ ,  $0.5$  y  $0.7$  m extrayendo la serie de tiempo generada de  $v_x$  o  $(\nabla \times \mathbf{v})_z$  en los 82 sensores de medición para un flujo con un  $Re = 30$ . Se observa como las fluctuaciones de porcentaje de error cambian en el tiempo. Dado que es un  $Re$  relativamente bajo, el flujo mantiene un mismo perfil al cabo de cierto tiempo, por lo que las predicciones de la RNA también lo hacen.



(a)



(b)

Figura 4.10: Porcentaje de error en el conjunto de predicción para distintos lugares de medición:  $x = 0.3$ ,  $0.5$ , y  $0.7$  m extrayendo la serie de tiempo generada de  $v_x$  o  $(\nabla \times \mathbf{v})_z$  en los 82 sensores de medición para un flujo con un  $Re = 99$ . A diferencia de los resultados mostrados en la Figura 4.9, donde se utilizó un  $Re = 30$ , dado que los vórtices se generan periódicamente por lo que las predicciones de la RNA también lo hacen aparentemente.

### 4.3. Estimación de un obstáculo en el flujo dentro de un conducto de dos dimensiones

Dado que las tuberías son la manera más usual de transportar fluidos en las industrias y vida cotidiana de las personas así como también en los sistemas biológicos, resulta crucial el reconocimiento de la formación de obstrucciones en conductos presentes en las redes de tuberías urbanas [Ugarelli et al., (2010)], ingeniería industrial [Kernan et al., (2014)] o incluso en cuestiones fisiológicas [Davies & Desai, (2008)].

Esto provee la motivación para la detección y caracterización de la forma y tamaño de algún obstáculo dentro de un flujo continuo. En esta sección se intentan dar los primeros pasos en esa dirección y utilizar las RNAs como una herramienta para caracterizar una obstrucción en una tubería a través de los patrones de flujo en condiciones determinadas. Nuevamente se realizó la simulación numérica mediante LBM para el caso del flujo alrededor de un obstáculo cilíndrico en dos dimensiones. De tal forma que a través de la medición de los patrones de flujo se puedan identificar el tamaño y ubicación del obstáculo en una sección transversal del tubo.

De forma análoga a la sección anterior, las variables consideradas como información fundamental extraída de los patrones de flujo son las componentes en  $x$  del campo de velocidades ( $v_x$ ) y en este caso se ha considerado la presión dinámica ( $q$ ) como una variable que suele ser más medido que la vorticidad.

Nuevamente, hay que remarcar que los detalles técnicos sobre la simulación en LBM se encuentran en el apéndice D. Para nuestros casos de estudio, se consideró el obstáculo cilíndrico inmerso en un flujo libre de fuerzas externas, moviéndose en la dirección positiva en el eje  $x$ . Además, se hicieron simulaciones numéricas, considerando como parámetros libres la velocidad entrante de un flujo tipo Poiseuille, el diámetro del obstáculo y la posición de este con respecto al eje  $y$ . Las simulaciones numéricas se detuvieron una vez que el sistema alcanzó la estabilidad neutral, lo cual ocurre en los casos de estudios presentados aproximadamente a los 16 s en el tiempo físico.



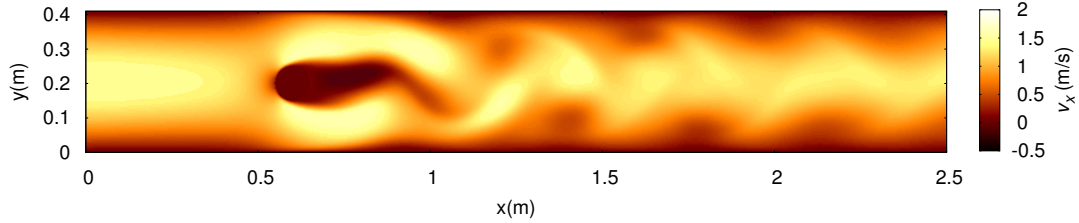


Figura 4.11: Magnitud de  $v_x$  obtenida en la simulación en la malla de LBM con un flujo entrante de Poiseuille con una velocidad característica de  $v_c = 1.5\text{m/s}$  y un obstáculo cilíndrico. Los obstáculos para los casos en esta sección se encuentran en  $x = 0.6\text{m}$  y su posición varía sobre el eje  $y$ .

En las siguientes subsecciones se presentan los parámetros utilizados para elaborar las simulaciones, la metodología en los casos de estudio que se presentan, así como también la selección de las características de entrada para la estructura de las RNAs, y finalmente los resultados concernientes a la caracterización de un obstáculo.

En términos de interpretación, se ha definido  $\beta$  como la razón de los diámetros del cilindro y la tubería. El diámetro se ha variado tal que  $\beta$  cubre un rango de  $1/80$  hasta valores cercanos a 1, siendo esto una obstrucción casi completa de la tubería, la cual tiene dimensiones de  $L_y = 0.41\text{m}$  de ancho (eje  $y$ ) y  $L_x = 2.5\text{m}$  de largo (eje  $x$ ). Para cubrir estas dimensiones en términos del código LBM, se ha definido una malla de  $165 \times 1000$  nodos. Colocando los obstáculos siempre a una distancia  $x = 0.6\text{m}$  considerando como  $x = 0\text{m}$  el inicio del dominio numérico de la tubería. A partir de ahora se refiere a la ubicación del obstáculo a su posición relativa respecto al origen del eje  $y$ .

Siguiendo de nuevo como referencia el problema utilizado en [Schäfer & Turek, (1996)], se ha utilizado un flujo con velocidad entrante estacionaria tipo Poiseuille, con una densidad  $\rho = 10^3\text{ kg/m}^3$  y una viscosidad cinemática de  $\nu = 10^{-3}\text{ m}^2/\text{s}$ . Una ejemplificación de una simulación para este sistema en particular se puede observar en la Figura 4.11, donde se ha colocado un obstáculo de diámetro  $d = 0.1\text{m}$  con un flujo de Poiseuille con una velocidad característica  $v_c = 1.5\text{ m/s}$

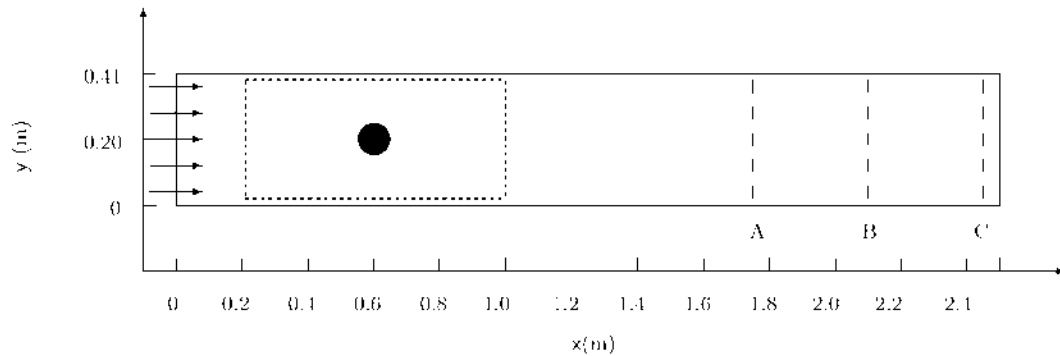


Figura 4.12: Representación esquemática del conducto en dos dimensiones, el obstáculo cilíndrico y los distintos sitios de medición marcados con líneas punteadas y localizados a lo largo del tubo en: A = 1.75 m, B = 2.10 m y C = 2.45 m. El rectángulo ilustrado con puntos muestra el área que es utilizada como objetivo para la RNA. Para todos los casos de estudio, los obstáculos se han ubicado en  $x = 0.6$  m.

### 4.3.1. Metodología

Para realizar la estimación del tamaño y ubicación del obstáculo, primeramente se debe definir cual es la información que se va introducir en la RNA y la que se quiere recibir por parte de ella. Para ello, por un lado se ha definido una *región objetivo* que contiene la información tanto del obstáculo como del fluido en la cual está inmerso, y por el otro lado se establecieron tres ubicaciones donde se colocaron los sensores (numéricos), permitiendo extraer información del flujo en distintas ubicaciones en el conducto. Estos sitios de medición se localizan en  $x = 1.75$ ,  $2.10$  y  $2.45$  m, que serán llamados de ahora en adelante como A, B y C. Todo esto se puede visualizar mejor con ayuda de la Figura 4.12.

En las siguientes subsecciones se describirán los casos de estudio, los datos utilizados como vectores de entrada y objetivo, así como la estructura de la RNA para cada uno de los casos que se detallan.

## Construcción de los casos de estudio

Para comprobar si las RNAs son capaces de dar una estimación sobre un obstáculo en una región de la tubería con un fluido, se realizaron los siguientes casos de estudio:

1. Se hicieron 80 simulaciones para el caso en que el obstáculo se encuentra inmerso justo en el centro del tubo, es decir,  $y = 0.21$  m. En cada una de ellas se ha variado el diámetro del obstáculo desde  $d = 0.005$  m hasta  $d = 0.395$  m en incrementos de  $\Delta d = 0.005$  m. Además de estos, se incluyó un obstáculo minúsculo de  $d = 0.001$  m. Para estos casos, se seleccionó como conjunto de predicción aquellos obstáculos con diámetros que barren de  $d = 0.02$  m a  $d = 0.375$  m con  $\Delta d = 0.025$  m. El resto de ellos fue usado en el conjunto de entrenamiento. Aunado a esto, se dividió en tres subcasos dependientes de distintos datos de entrada:
  - a) Se han considerando los perfiles de  $v_x$  o  $q$  en el tiempo  $t = 16$  s. Es decir, los vectores de entrada para la RNA se conforman solamente por el perfil de los valores de  $v_x$  o  $q$  en un instante de tiempo en un único sitio de medición. De los 165 nodos de la malla numérica que representan el ancho del tubo, se seleccionaron únicamente 83 de estos nodos como sensores numéricos.
  - b) Como segundo subcaso y, a diferencia del anterior, se alimentó a la RNA con la serie de tiempo discreta y conformada por 300 pasos de tiempo, desde  $t = 0$  s, es decir, desde que se inicia el flujo del fluido sobre el tubo, hasta  $t = 16$  s. Sin embargo, no se consideraron los 83 nodos como en el subcaso anterior, sino solamente la información de un nodo en el centro del tubo en  $y = 0.21$  m.
  - c) El mismo procedimiento que el subcaso anterior, pero se incrementó de 1 a 3 nodos equidistantes de extracción, esto es a  $y = 0.105$ ,  $0.21$  y  $0.315$  m, lo que supondría una mejora en la estimación del objetivo. Con estos dos subcasos, se intenta medir qué opción es más efectiva para el objetivo descrito: si una única medición con mucha información sobre el

eje  $y$  o una serie de tiempo con pocos puntos de medición, como puede ser el caso en una aplicación más realista.

2. En contraste con el caso 1, se consideraron únicamente obstáculos con diámetros  $d = 0.05, 0.1$  y  $0.2$  m. Pero en cada caso, se cambió la ubicación del centro del tubo en 43 posiciones distintas sobre el eje  $y$ , de tal forma que este puede estar en el centro o como en las cercanías a las paredes del tubo. Esto significa un total de 129 simulaciones, de las cuales 22 de ellas fueron seleccionadas para el conjunto de entrenamiento y validación, mientras que las 21 restantes de ellas conforman el conjunto de predicción. Aunado a este conjunto de predicción se consideraron, 21 simulaciones en distintas posiciones para un obstáculo de tamaño  $d = 0.15$  m el cual es totalmente desconocido para la RNA en las fases de entrenamiento y validación. Este escenario se dividió en tres subcasos siguiendo los mismos lineamientos del caso 1.
3. En este caso, se analizó el escenario donde el obstáculo se encuentra nuevamente situado en  $y = 0.21$  m y  $x = 0.6$  m. Sin embargo, se ha cambiado la velocidad característica del flujo entrante al conducto, lo que obviamente cambia los perfiles de  $v_x$  o  $q$ . La velocidad característica  $v_c$  se ha variado desde  $v_c = 0.15$  m/s, lo que puede significar poca perturbación en el flujo por parte del obstáculo, hasta  $v_c = 1.8$  m/s en incrementos de  $\Delta v_c = 0.15$  m/s. Este caso, es una extensión del subcaso 1a, considerando doce distintas velocidades de flujo y once tamaños diferentes de obstáculos que van desde  $d = 0.005$  m hasta  $d = 0.395$  m. Esto se resume en la Tabla 4.6. Para este ejemplo, los vectores de entrada se conforman de diez valores tanto del perfil de flujo entrante (antes del obstáculo) como en el sitio de medición B, correspondientes a diez nodos equidistantes a lo largo del eje  $y$ , es decir, un análisis similar al caso 1a. Dadas estas consideraciones, los valores de entrada para la RNA son 20 valores de cada una de las 132 simulaciones hechas. En este caso, el conjunto de predicción consta de 20 simulaciones que fueron seleccionadas de forma aleatoria para garantizar una mayor homogeneidad dado que el número de parámetros libres se ha incrementado.

### 4.3.2. Región objetivo

Dado que se ha utilizado un entrenamiento supervisado, se requiere proveer los patrones objetivo relacionados con el verdadero tamaño y ubicación del obstáculo dentro del tubo. Por esa razón y como se mencionó anteriormente, se definió una región en el área circundante al obstáculo, ver Figura 4.12. Por cuestiones prácticas, dado que en la simulación se tiene un gran número de nodos, la cantidad de información que se podría considerar en esta región sería muy grande para formar un vector de salida, por lo que se propuso considerar dicha región con un mallado de 40 celdas de largo por 20 de ancho, donde cada celda contiene  $8 \times 8$  nodos de la malla numérica. A esta región se le llamará de ahora en adelante *malla objetivo*.

Para darle un sentido físico a la malla objetivo, y a cada una de las celdas dentro de esta, se les ha asignado un valor numérico dependiendo de la cantidad de nodos que representan a los elementos del fluido y a los sólidos. En otras palabras, la proporción de celdas que sirven como vectores objetivo ocupadas por el obstáculo son el número de nodos que representan elementos sólidos entre el número de nodos totales contenidos en la celda. Para simplificar la notación, se ha definido un índice de ocupación para cada celda en la malla objetivo, donde se ha calculado la relación sólido-líquido de la celda, definiendo un valor numérico de 1 a una celda que contiene únicamente nodos que representan elementos sólidos y un valor de 0 a una celda con nodos que representan elementos de fluido únicamente. De ahora en adelante, se llamará a este valor como *índice de razón sólido-líquido* (IRSL). Una ejemplificación sobre la transformación de la simulación de la malla numérica en LBM y la malla objetivo para un obstáculo de 0.2 m de diámetro se muestra en la Figura 4.13. En esta figura, el tono oscuro esquematiza las celdas que son ocupadas en su mayoría por el obstáculo (IRSL alto), mientras que las celdas con tonos más claros significan que hay una proporción mayor de elementos de fluido (IRSL bajo).

### 4.3.3. Estructura de la RNA

Recordando que para el caso 1a, se han reducido el número de nodos a considerar de 165 a 83, donde se extraen los valores de  $v_x$  o  $q$  como vector de entrada para

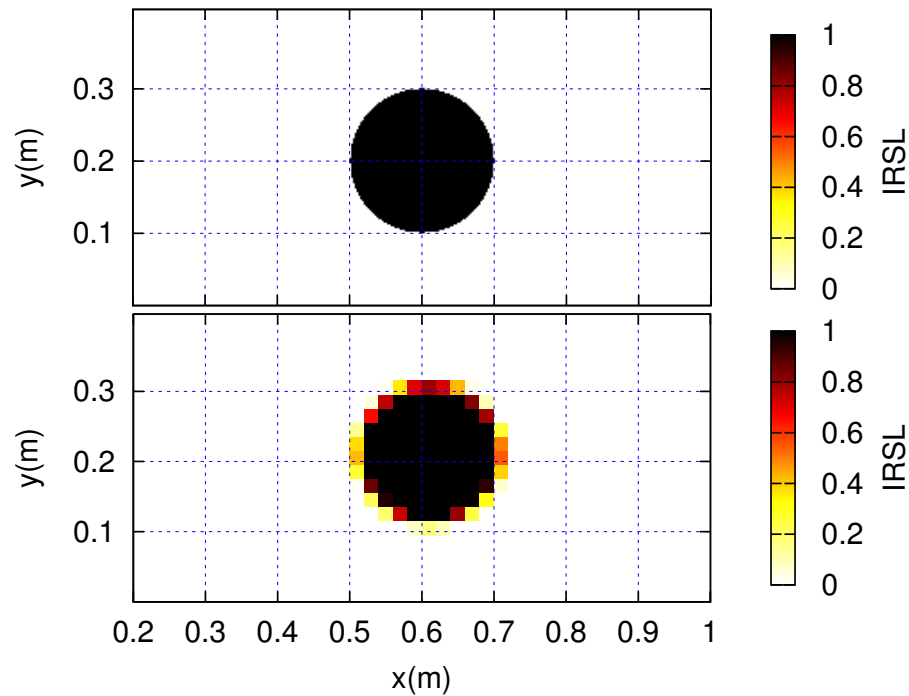


Figura 4.13: En el gráfico superior se representa el índice de razón de elementos sólido-líquido en la malla de numérica de LBM, para un obstáculo cilíndrico con un diámetro  $d = 0.2$  m. Abajo se representa la transformación del mismo obstáculo pero en la malla objetivo que se utiliza como objetivo para el entrenamiento de la RNA y comparativa en los resultados. El menor número de celdas en la región objetivo en comparación con la misma área en la simulación de LBM causa una pérdida de resolución en el mallado. Un valor igual a 0 significa que la celda está compuesta únicamente por elementos de fluido y un valor 1 representa solamente elementos sólidos. Nota: Las líneas en los gráficos son únicamente de referencia y no representan el tamaño real de las celdas utilizadas en ninguno de ambos casos.

RMSE usando $(\nabla \times \mathbf{v})_z$					
	0.3m	0.5	0.7m	1.1m	1.9m
1	42.625	26.614	30.177	21.029	22.663
4	2.135	4.593	3.447	3.748	2.016
6	0.897	1.282	2.634	4.120	1.474
11	0.731	1.850	3.439	1.340	1.267
21	0.298	0.983	2.271	1.333	1.159
41	0.271	0.883	1.671	2.930	1.200
82	0.236	0.928	1.503	1.012	1.026

Tabla 4.2: Misma descripción que en la Tabla 4.1 pero usando en este caso  $(\nabla \times \mathbf{v})_z$  en lugar de  $v_x$ .

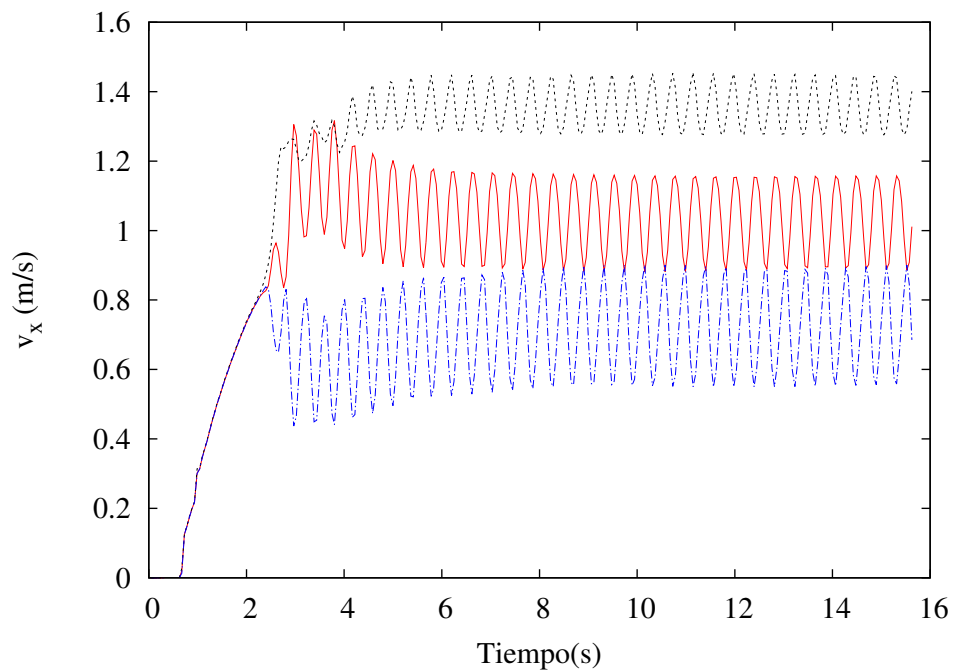


Figura 4.14: Series de tiempo formadas por los valores de  $v_x$  en  $y = 0.105$  m (línea roja continua),  $0.210$  m (línea negra punteada) y  $0.315$  m (línea azul punteada y con puntos). Los valores fueron medidos en el sitio de medición B para un obstáculo con  $d = 0.2$  m y centrado en  $y = 0.21$  m. En esta gráfica se aprecia como el flujo ha alcanzado una estabilidad neutral a partir del segundo 9 aproximadamente. La concatenación de 300 valores cada una de las series de tiempo funciona como el vector de entrada para la RNA.

Predicción de $Re$ usando $v_x$					
$Re$	0.3m	0.5m	0.7m	1.1m	1.9m
12	11.801	11.633	11.968	12.035	12.072
17	16.795	16.433	16.786	16.906	17.006
22	21.841	21.340	21.645	21.762	21.901
27	26.915	26.368	26.583	26.658	26.781
32	31.997	31.497	31.617	31.641	31.679
37	37.067	36.709	36.750	36.743	36.623
52	52.071	52.277	52.250	52.250	52.237
64	63.990	63.988	64.080	64.214	64.357
70	69.947	69.853	70.027	70.164	70.238
76	76.020	76.076	75.959	76.268	76.061
92	92.057	92.260	91.518	91.627	91.826
99	99.393	98.994	99.333	98.611	98.747
102	102.269	102.156	102.495	101.950	101.291
107	107.020	106.755	106.662	106.443	107.171
112	111.831	111.609	111.180	111.776	112.279
116	115.574	115.2822	115.506	116.150	116.302
117	116.553	116.912	116.912	116.726	117.095
118	117.430	117.037	117.244	118.149	118.238
119	118.432	118.536	119.193	119.024	119.303
120	119.152	118.996	118.636	118.109	120.471
$\chi^2$	<b>0.025</b>	<b>0.103</b>	<b>0.058</b>	<b>0.054</b>	<b>0.024</b>

Tabla 4.3: Valores predichos de  $Re$  para los distintos sitios de medición muestreando 82 valores de  $v_x$ . La comparación del desempeño entre los distintos casos se hace bajo la prueba  $\chi^2$ . Bajo esta métrica se considera un modelo mejor que otro cuando el valor de  $\chi^2$  es más pequeño. El peor de los casos se obtiene en el sitio de medición  $x = 0.5m$ .



Predicción de $Re$ usando $(\nabla \times \mathbf{v})_z$					
$Re$	0.3m	0.5m	0.7m	1.1m	1.9m
12	11.990	12.031	12.050	12.244	12.068
17	17.053	16.990	17.169	16.882	16.995
22	22.062	21.958	22.130	21.584	21.943
27	27.039	26.947	26.929	26.394	26.943
32	32.004	31.951	31.668	31.378	32.004
37	36.983	36.959	36.544	36.601	37.132
52	51.954	51.950	52.887	51.826	52.384
64	63.915	63.858	64.125	63.777	64.673
70	69.961	70.172	70.143	69.552	70.454
76	75.952	75.549	75.845	75.387	77.051
92	92.051	92.420	92.461	91.760	92.642
99	99.047	99.432	101.045	99.576	99.995
102	101.928	102.472	101.551	101.096	101.308
107	107.234	107.215	106.848	105.582	108.467
112	112.212	111.651	110.902	110.724	113.197
116	115.917	114.496	114.959	118.113	113.819
117	116.630	115.119	115.826	117.790	119.018
118	117.607	117.241	113.734	116.829	116.612
119	118.414	116.878	117.937	119.214	119.683
120	119.415	117.692	115.957	117.295	118.436
$\chi^2$	<b>0.001</b>	<b>0.148</b>	<b>0.406</b>	<b>0.214</b>	<b>0.197</b>

Tabla 4.4: La descripción es similar a la Tabla 4.3, pero en este caso se ha utilizado  $(\nabla \times \mathbf{v})_z$  en lugar de  $v_x$  como dato de entrada para la RNA.

Parámetros de los casos de estudio					
Caso	Diámetros	Posiciones en $y$	Vel. de flujo $v_c$	Núm. sensores	Tiempo
1a	80	1	1	83	1
1b	80	1	1	1	300
1c	80	1	1	3	300
2a	4	43	1	83	1
2b	4	43	1	1	300
2c	4	43	1	3	300
3	11	1	12	10	1

Tabla 4.5: Detalle de los valores de los parámetros usados en las simulaciones para los diferentes casos de estudio. La columna llamada “Diámetro” se refiere al número de distintos tamaños de los obstáculos. La columna “Posiciones en  $y$ ” es el número de diferentes ubicaciones de los obstáculos a lo largo del eje  $y$ , es decir, una sección transversal del tubo en  $x = 0.6$  m. La columna llamada “Velocidades de flujo  $v_c$ ” se refiere a las variaciones de velocidades de flujo entrante usadas para el caso 3. Mientras que la columna “Núm. Sensores” define el número de sensores igualmente espaciados en cada uno de los sitios de medición. Finalmente la última columna “Tiempo” corresponde al número de mediciones tomadas en el tiempo en todos los sensores en el sitio de medición indicado.

Parámetros para el caso 3		
Diámetro (m)	$v_c$ (m/s)	Ubicación de sensor (m)
0.005	0.15	0.025
0.04	0.30	0.065
0.08	0.45	0.105
0.12	0.60	0.145
0.16	0.75	0.185
0.20	0.90	0.225
0.24	1.05	0.265
0.28	1.20	0.305
0.32	1.35	0.365
0.36	1.5	0.405
0.395	1.65	
	1.8	

Tabla 4.6: Parámetros para los obstáculos con diferentes diámetros, velocidad de flujo entrante y ubicaciones de los sensores usados para el caso 3. Los sensores se ubican en posiciones equidistantes tanto en  $x = 0$ m como en  $x = 2.10$ m.

la RNA, lo que ha simplificado la estructura de la RNA y acelera el cómputo en el entrenamiento. Con esto, el vector de entrada consiste de 83 neuronas relacionadas con el perfil de flujo en el sitio de medición. Por ejemplo, para el caso 1a si se consideran los valores de  $v_x$  en el sitio B, el vector de entrada  $\mathbf{I}$  es:

$$\mathbf{I} = v_{x_1}, v_{x_3}, \dots, v_{x_i}, \dots, v_{x_{165}}, \quad (4.7)$$

donde los subíndices enumeran el sensor numérico en el nodo en la malla numérica. Para el caso 1b, el vector de entrada esta definido por la serie de tiempo de  $v_x$  o  $q$  en  $y = 0.21$  m durante 300 pasos de tiempo equidistantes. Por ejemplo para  $v_x$  el vector  $\mathbf{I}$  es:

$$\mathbf{I} = v_{x_{85},t_1}, v_{x_{85},t_2}, \dots, v_{x_{85},t_{300}}; \quad (4.8)$$

donde  $v_{x_{85},t}$  etiqueta el valor de  $v_x$  en el nodo 85 en la malla de la simulación en  $y = 0.210$  m en el instante de tiempo  $t$ . Para el caso 1c el vector de entrada para la RNA consiste en los mismos 300 pasos de tiempo para  $v_x$  o  $q$  en las posiciones  $y = 0.105, 0.21$  y  $0.315$  m en el sitio de medición que se haya considerado. En la Figura 4.14 se presenta la evolución temporal de  $v_x$  en el sitio B. En este escenario el vector de entradas para  $v_x$  se representa como:

$$\mathbf{I} = v_{x_{43},t_1}, v_{x_{85},t_1}, v_{x_{127},t_1}, v_{x_{43},t_2}, v_{x_{85},t_2}, v_{x_{127},t_2}, \dots, v_{x_{43},t_{300}}, v_{x_{85},t_{300}}, v_{x_{127},t_{300}}. \quad (4.9)$$

En el segundo escenario, los vectores de entrada son similares a los descritos en las ecuaciones 4.7-4.9. Para el caso 3, considerando los datos descritos en la Tabla 4.5, el vector de entrada conformado por los valores de  $v_x$  es:

$$\mathbf{I} = v_{x_{11}}(x = 0), v_{x_{27}}(x = 0), \dots, v_{x_{163}}(x = 0) \\ , \dots, v_{x_{11}}(x = 2.1), v_{x_{27}}(x = 2.1), \dots, v_{x_{163}}(x = 2.1). \quad (4.10)$$

El mismo conjunto de ecuaciones (4.8-4.10), se aplican cuando se considera el perfil de  $q$  en lugar de  $v_x$  como vector de entrada. Siguiendo dichas ecuaciones, la estructura interna de la RNA, puede diferir internamente para cada caso, es decir, tanto el número de valores de entrada como del número de neuronas ocultas. Sin embargo, el objetivo es el mismo para todos los casos: estimar la morfología del obstáculo que bloquea el flujo dentro de la tubería.

Como se mencionó, la región donde se encuentra localizado el obstáculo se ha definido por una malla de 40x20 celdas, por tanto los vectores de salida constan de 800 elementos. Describiendo esto en términos matemáticos se tiene que para cualesquiera de los vectores de entrada  $\mathbf{I}$  como los definidos en las ecuaciones 4.7-4.10, el vector resultante arrojado la RNA es de la forma  $\mathbf{O} = [O_1, O_2, \dots, O_{800}]$ , donde el  $k$ -ésimo elemento de este vector se define como:

$$O_k = F_2 \left( \sum_{j=1}^m w_{jk} F_1 \left( \sum_{i=1}^n w_{ij} I_i + b_j \right) + b_k \right) \quad (4.11)$$

con  $1 \leq k \leq 800$ ,  $I_i$  el  $i$ -ésimo elemento del vector de entrada,  $n$  el número de entradas para los distintos casos y  $m$  en este caso define el número de neuronas ocultas.  $F_1$  y  $F_2$  son las funciones de activación para las capas oculta y de salida respectivamente;  $w_{ij}$  y  $b_j$  indizan los pesos y bias entre la  $i$ -ésima neurona de entrada y la  $j$ -ésima neurona oculta; mientras que  $w_{jk}$  y  $b_k$  indizan los pesos y bias entre los elementos de la capa oculta y de salida respectivamente. Para todos los escenarios analizados 1-3, las RNAs construidas utilizan 20 neuronas ocultas y una capa de salida con 800 elementos. En todos los casos se utilizó una función de activación tipo sigmoide tanto en la capa oculta como en la de salida, y entrenando las RNAs por un máximo de 15,000 iteraciones usando nuevamente la técnica de validación como criterio de evaluación, con una constante de aprendizaje de  $\gamma = 0.001$ .

#### 4.3.4. Resultados

Para estimar el desempeño de la RNA en todos los casos a analizar, se ha empleado el coeficiente  $R^2$ . El cómputo de este coeficiente se ha hecho tomando en cuenta tanto la malla objetivo como la malla predicha por la RNA, considerando los valores reales y predichos del IRSL:

$$R^2 = 1 - \frac{\sum_{i=1}^{800} (O_i - \langle T \rangle)^2}{\sum_{i=1}^{800} (T_i - \langle T \rangle)^2}, \quad (4.12)$$

donde  $T_i$  y  $O_i$  son los  $i$ -ésimos elementos de los vectores objetivo y de salida de la RNA respectivamente, mientras que  $\langle T \rangle$  es el promedio del IRSL para el vector

objetivo definido por:

$$\langle T \rangle = \frac{1}{800} \sum_{i=1}^{800} T_i. \quad (4.13)$$

Esto significa que el coeficiente  $R^2$  esta delimitado entre los valores  $(-\infty, 1]$ , donde un valor de  $R^2 = 1$  implica que los vectores objetivo y de predicción son iguales término a término. A medida que  $R^2 \rightarrow 0$  significa que la predicción se asemeja más a  $\langle T \rangle$ .

Para el caso 1a, se presentan resultados midiendo en los detectores localizados no solamente en el sitio B, sino también en A y C, tal y como se muestra en la Figura 4.15. En este caso, se estudia el desempeño de la RNA cuando solamente es entrenada en B y evaluada en A, B y C para los mismos obstáculos. Las mediciones en el sitio B, ambas variables  $q$  y  $v_x$ , muestran valores muy cercanos a 1 si se consideran obstáculos con diámetros más grandes a 0.05 m o equivalentemente un  $\beta > 0.25$ . Ver por ejemplo la Figura 4.16, donde se grafican tanto la malla objetivo y de predicción con un obstáculo de 0.195 m y se alcanza un  $R^2 = 0.979$ . Sin embargo, para obstáculos más pequeños la precisión disminuyó a coeficientes de  $R^2 = 0.654$  para  $v_x$  y  $R^2 = 0.802$  para  $q$ , ambos con un obstáculo de diámetro  $d = 0.02$  m. Las estimaciones de tamaño de las mediciones en A y C muestran resultados similares. Por ejemplo la estimación para el obstáculo con diámetro  $d = 0.05$  m, el coeficiente  $R^2$  disminuye un 16 %. Esto implica que para mediciones mucho más alejadas de este punto de medición, las predicciones de la RNA pueden ser aún buenas. A partir de estos resultados, no es posible establecer cual de las variables físicas ( $v_x$  o  $q$ ) sea más relevante, y dar estimaciones más precisas.

Para los casos 1b y 1c, el vector de entrada de la RNA consiste en la serie de tiempo, de una única medición en un nodo o en tres nodos en el sitio de medición B. Los resultados para estos subcasos se muestran en la Figura 4.17. Particularmente para el caso 1b, los valores de  $R^2$  son superiores a 0.8 considerando  $v_x$ , y decremantan para  $q$  hasta un valor de  $R^2 = 0.231$ . Mientras que para el caso 1c con tres sensores los resultados mejoran considerablemente, manteniendo resultados similares para ambas cantidades físicas.

Para el caso 2a, donde las RNAs deben estimar no sólo diámetro del obstáculo

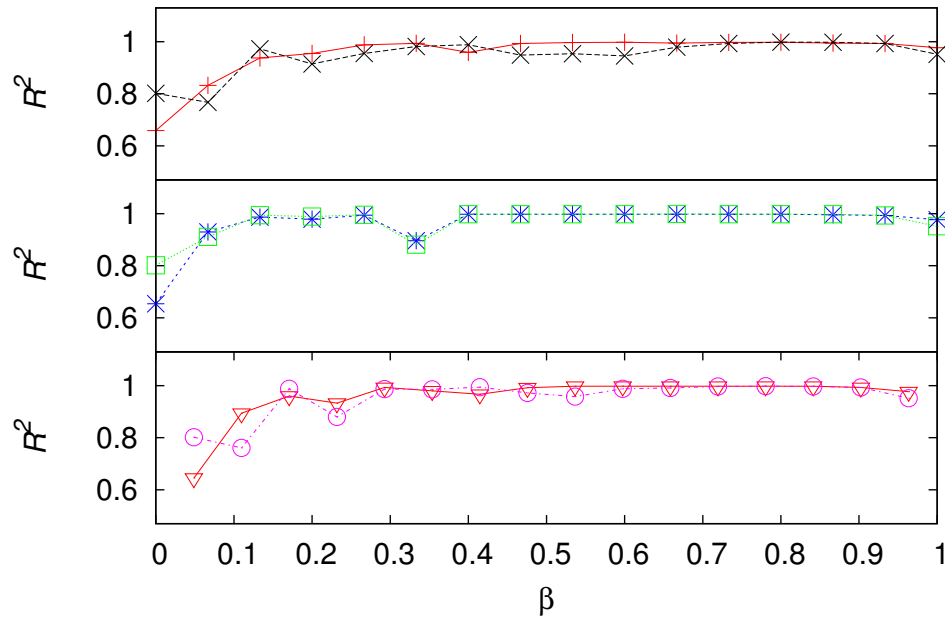


Figura 4.15: Valores de  $R^2$  en la predicción de la RNA producidos para obstáculos de diferentes diámetros y localizados en el centro de la tubería. Por un lado, la RNA se entrenó con valores de  $v_x$  o  $q$  en un único sensor numérico ubicado en el nodo central de la malla de LBM en el sitio de medición B. Las predicciones corresponden a los sitios de medición A (arriba), B (medio) o C (abajo). Por otro lado, se probó el desempeño de la red para obstáculos de tamaños desconocidos tanto en el sitio de medición A, B y C. En el gráfico de arriba, la curva roja (+) muestra los resultados utilizando  $v_x$  en el sitio A, mientras que la curva negra (x) son aquellos utilizando  $q$ . Las líneas azul (\*) y verde (□) representan los ajustes de  $R^2$  usando  $v_x$  y  $q$  respectivamente. Abajo, la curva roja (∇) y magenta (○) corresponden a los resultados para  $v_x$  y  $q$  en C respectivamente. Todos los valores de  $R^2$  son mayores a 0.6, lo que se puede considerar como una buena correlación entre la predicción y el obstáculo objetivo. Nótese como los resultados son independientes del lugar de extracción de los datos, aun cuando la RNA se entrenó en un único sitio de medición.

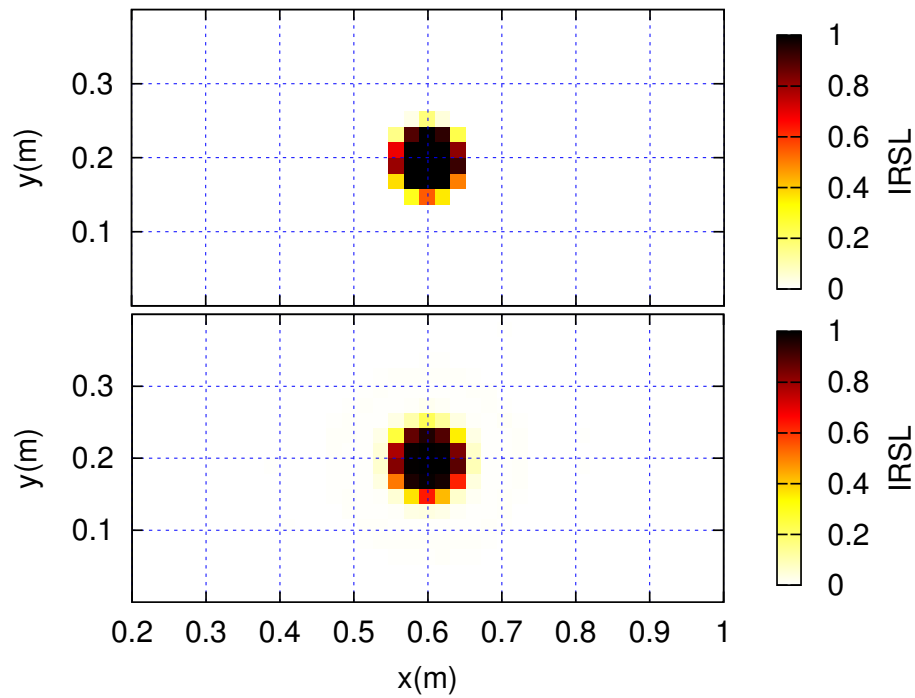


Figura 4.16: En el gráfico superior se muestra la malla objetivo, mientras que abajo se presenta la predicción sobre la misma por parte de la RNA, la precisión se basa en la diferencia entre los valores de ambos. En este caso se consideró la componente  $v_x$  del perfil del flujo como valores de entrada, con un obstáculo en el centro del tubo con diámetro  $d = 0.195$ . El color del recuadro representa el IRSL. La diferencia entre ambos gráficos es prácticamente imperceptible a simple vista, el valor de  $R^2$  para este caso en particular es de 0.979.

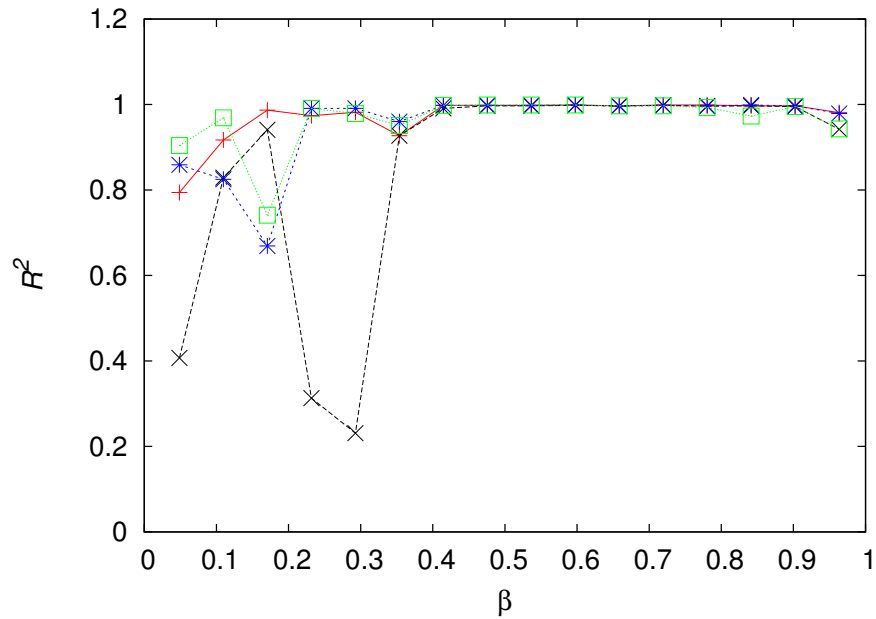


Figura 4.17: Análisis de las predicciones de la RNA considerando los casos  $1b$  y  $1c$  para diferentes diámetro de obstáculos centrados en  $x = 0.6\text{m}$ . Arriba, los resultados para la serie del tiempo de un solo sensor en  $y = 0.205\text{m}$  ( $1b$ ) considerando  $v_x$  (+) y  $q$  (x). En el caso  $1c$  los tres sensores se localizan en  $y = 0.105\text{m}$ ,  $0.210\text{m}$  y  $0.315\text{m}$ , y los resultados para  $v_x$  (\*) y  $q$  (□) tienen valores de  $R^2$  cercanos a 1 para diámetros mayores a  $d = 0.85\text{m}$ , mientras que estos valores se alcanzan para diámetros mayores a  $d = 0.15\text{m}$  en el caso de usar un único sensor.



sino también su ubicación sobre el eje  $y$ , se ha alcanzado un valor de  $R^2$  cercano a 1, para bloqueos al menos mayores a  $\beta = 0.25$ , como se puede ver en la Figura 4.18. Para obstáculos con un diámetro  $d = 0.15$  m para lo cuales la RNA no fue entrenada, el valor de  $R^2$  disminuyó hasta 0.356 para  $v_x$  y 0.243 para  $q$ .

El peor caso se presentó para un obstáculo con  $d = 0.1$  m y ubicado en el centro del tubo, obteniendo un valor de  $R^2 < 0$  utilizando la variable  $q$  y  $R^2 = 0.66$  para  $v_x$ . Un ejemplo sobre este caso se representa en la Figura 4.19, donde el diámetro objetivo es de  $d = 0.05$ m y se obtiene un coeficiente de  $R^2 = 0.09$ , el cual indica que la estimación tanto del tamaño como la ubicación no es muy precisa.

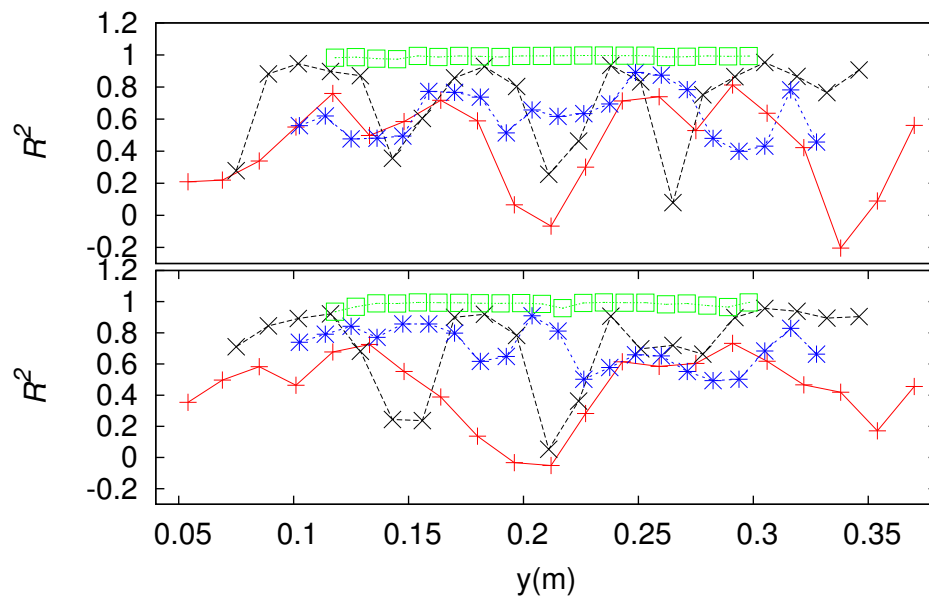


Figura 4.18: Análisis de las predicciones para obstáculos en distintas posiciones sobre el eje  $y$  y en  $x = 0.6$ m. La RNA ha sido entrenada y probada con los perfiles de  $v_x$  (arriba) y  $q$  (abajo) en el sitio de medición B. La curva roja (+) muestra los resultados para  $d = 0.05$ m, mientras que las negra ( $\times$ ), azul ( $\ast$ ) y verde ( $\square$ ) indican aquellos para  $d = 0.1$ m,  $0.15$  m y  $0.2$ m respectivamente. Los valores de ajuste de  $R^2$  son similares en ambas propuestas. La peor predicción es en general para el obstáculo más pequeño, tanto en el centro del tubo como en los extremos cercanos a las paredes de este.

Respecto al subcaso 2b, donde se empleó la serie de tiempo de un único nodo en

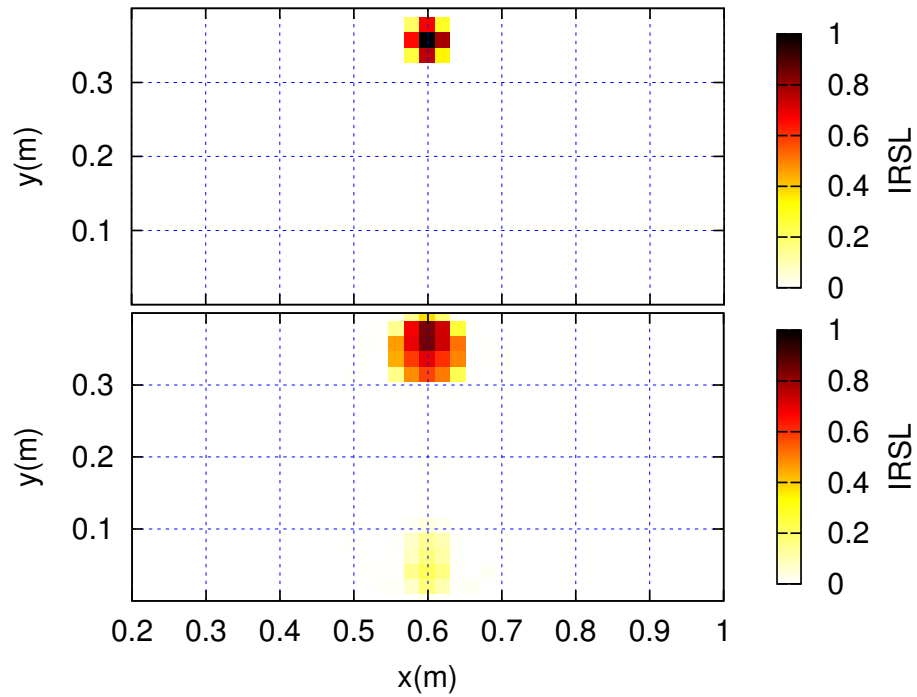


Figura 4.19: IRSL para la región objetivo (arriba) y la predicción hecha por la RNA (abajo) considerando 83 valores de la componente  $x$  de la velocidad del fluido en el sitio B. En este caso el obstáculo se encuentra en  $y = 0.35\text{m}$  y tiene un diámetro  $d = 0.05\text{m}$ . Contrastando con la Fig. 4.16, se observa como la RNA se confundió en este caso, indicando un bajo IRSL en la parte inferior del tubo, por lo cual se obtuvo un coeficiente  $R^2 = 0.09$ .

el centro de la tubería, la RNA fue incapaz de aprender el comportamiento del flujo sobre todo cuando el obstáculo es pequeño y se encuentra cerca de las paredes. Esto es evidente de los resultados mostrados en la Figura 4.20, donde incluso se obtienen valores negativos para estas obstrucciones, alcanzando un mínimo de  $R^2 = -0.312$  cuando estos se encuentran en el centro de la tubería. Aunque para obstáculos más grandes la RNA obtuvo mejores resultados, dado que las perturbaciones en el flujo son mayores, también tiene dificultades cuando se encuentran cerca de las paredes.

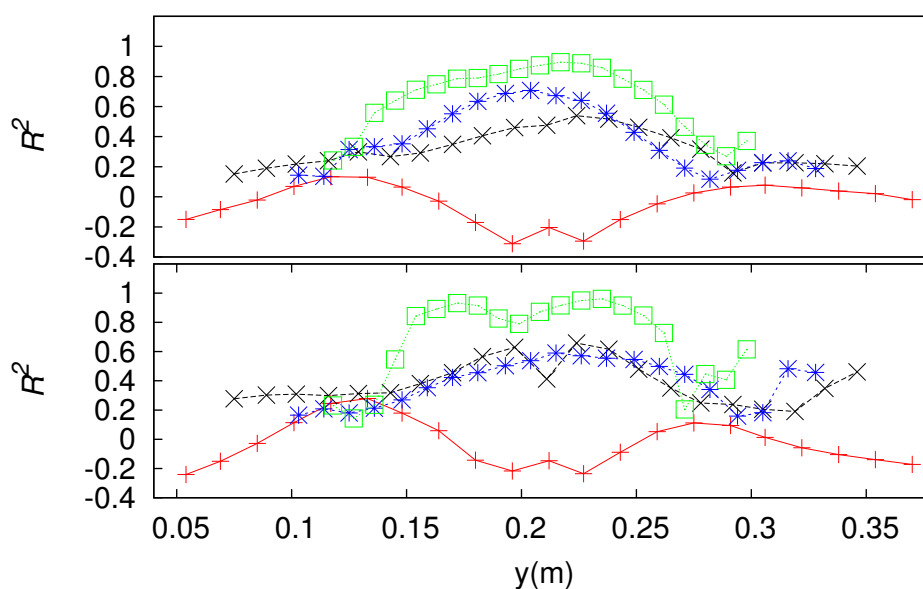


Figura 4.20: Análisis de predicción para los obstáculos en distintas posiciones sobre el eje  $y$  (caso 2b). Las RNAs utilizadas fueron entrenadas con las series de tiempo de  $v_x$  (arriba) o  $q$  (abajo) tomadas del centro de la tubería y en el sitio de medición B. Las curvas roja (+), negra ( $\times$ ), azul ( $\ast$ ) y verde ( $\square$ ) muestran los valores de  $R^2$  para diámetros  $d = 0.05\text{m}$ ,  $0.15\text{m}$ , y  $0.2\text{m}$  respectivamente. Es evidente la dificultad de las RNA para dar una aproximación precisa acerca de la localización del obstáculo cuando los obstáculos más pequeños y cercanos a las paredes del conducto.

Para el caso de estudio 2c, donde se ha considerado la serie de tiempo de tres valores de nodos equidistantes como vector de entrada en  $y = 0.105\text{m}$ ,  $0.210\text{m}$  y  $0.314\text{m}$  en el sitio B, los resultados que se muestran en la Figura 4.21 indican

valores semejantes de  $R^2$  para ambas variables  $q$  y  $v_x$ . En este escenario, las predicciones para los obstáculos pequeños indican una mejoría sobre el caso 2b. Para un obstáculo con  $d = 0.15\text{m}$  todos los valores de  $R^2$  están sobre 0.4 usando tanto  $v_x$  como  $q$ . Mientras que los resultados para diámetros  $d = 0.1\text{m}$  y  $0.2\text{m}$  en todas las ubicaciones, las predicciones se ajustan mucho mejor al objetivo, con  $R^2$  arriba de 0.8. Sin embargo, para el obstáculo con  $d = 0.05\text{m}$  ubicado en el centro del tubo usando  $q$  tiene un valor de  $R^2 = -0.102$ .

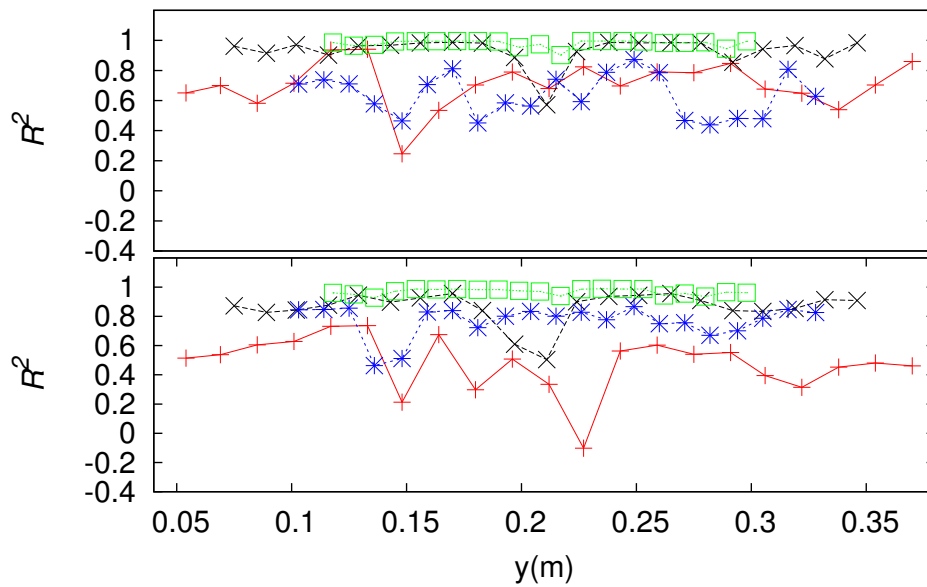


Figura 4.21: Valores del coeficiente  $R^2$  resultantes de la comparación entre obstáculo predicho para el caso 2c. Las RNAs fueron entrenadas con las series de tiempo de  $v_x$  (arriba) y  $q$  (abajo) en tres diferentes lugares  $y = 0.105\text{m}$ ,  $0.210\text{m}$  y  $0.315\text{m}$  en el sitio de medición B. Las curvas roja (+), negra ( $\times$ ), azul ( $\times$ ) y verde ( $\square$ ) muestran los resultados para obstáculos con  $d = 0.05\text{m}$ ,  $0.1\text{m}$ ,  $0.15$  y  $0.2\text{m}$  respectivamente. De nuevo el peor ajuste se obtuvo con el cuerpo más pequeño con un  $R^2 = -0.102$  y utilizando  $q$  como valor de entrada. Sin embargo, hay una mejora apreciable considerando tres sensores en lugar de un único sensor como en el caso 2b.

En general, para el caso 2, las peores predicciones ocurren para obstáculos pequeños y/o se encuentran cerca de las frontera. No es de sorprender esta confusión

en la RNA, dado que las fronteras afectan de manera más importante el flujo después del obstáculo. Bajo este marco, las capas límite se ven afectadas por la naturaleza viscosa del flujo, es decir, que si el obstáculo es muy pequeño, el efecto que este provoca sobre el flujo se vea neutralizado por las fuerzas viscosas del mismo. Esto implica que cuando se mide lejos de la obstrucción, el flujo se comporta prácticamente como si no hubiera obstáculo alguno. Más aún, para obstáculos más pequeños, sin importar su distancia respecto a las paredes del tubo, estos no afectan el flujo de tal manera que los patrones que se obtienen detrás de ellos es difícilmente distinguible.

De los resultados de los casos 1 y 2, muestran que mientras se incrementa el número de puntos en tiempo y espacio, la predicción sobre el tamaño del obstáculo mejora considerablemente. Por ejemplo, comparando el subcaso 1a, donde se utilizaron 83 diferentes valores de  $v_x$  o  $q$  sobre el eje  $y$  en un único instante de tiempo, con el subcaso 1c, donde se tomaron en cuenta valores 3 sobre el eje  $y$  en 300 pasos de tiempo. Lo mismo se puede concluir al comparar los subcasos 2a y 2c. Sin embargo, en una posible aplicación puede ser más factible y deseable, hacer múltiples mediciones a lo largo del tiempo con pocos sensores en un único sitio de medición, que realizar una sola medición con un 83 o más sensores.

En referencia a los casos 3, los resultados, en términos del coeficiente  $R^2$  para cada escenario, se muestran en la Tabla 4.7. El problema de una predicción deficiente con el obstáculo más pequeño de  $d = 0.005$  m persiste, con predicciones considerablemente malas de  $R^2 = -23.75$  utilizando  $v_x$  y un valor de  $R^2 = -246.95$  usando  $q$ . Sin embargo, hay que mencionar que estos resultados se han obtenido considerando la segunda velocidad característica entrante al tubo más baja con  $v_c = 0.3$  m/s. Para poner esto en perspectiva, se puede comparar con el caso 1a, donde la velocidad característica asociada es  $v_c = 1.5$  m/s. Con lo que el decremento de precisión también se puede asociar con la velocidad de flujo entrante. Es de notar, como al incrementar  $v_c$  hasta 1.35 m/s se han incrementando los valores de  $R^2$  hasta -0.645 para  $v_x$  y -0.680 para  $q$ . Algo similar ocurre para diámetros  $d = 0.004$  m mejorando de un  $R^2 = 0.25$  hasta  $R^2 = 0.975$  para  $v_x$  y de  $R^2 = -4.518$  hasta  $R^2 = 0.975$  para  $q$ . Para el resto de los tamaños de obstáculos,

los resultados son mucho mejores, llegando a valores muy cercanos a 1. Además de esto, se puede notar que los resultados son mejores utilizando  $v_x$  sobre  $q$ , y que la habilidad de las RNA se extiende a tener buenos resultados con hasta 10 sensores en el espacio en un único sitio de medición.

Hasta ahora, se han hecho predicciones sobre los tamaños y ubicaciones de obstáculos circulares sobre la sección transversal de la tubería. Sin embargo, es fácil explorar cual sería el desempeño de una RNA de las ya entrenadas, para estimar también la forma de otro obstáculo: en este caso se ejemplifica con un cuadrado de 0.1 m de lado bajo las mismas condiciones que en el caso 1, este escenario se muestra en la Figura 4.22. Por simplicidad, solamente se presenta el caso de la serie de tiempo con tres sensores en el sitio B. A pesar de que la forma predicha es circular (algo que es comprensible dado que la RNA únicamente fue entrenada con este tipo de morfologías), el tamaño del obstáculo cuadrado es semejante alcanzando un valor de  $R^2 = 0.93$  tal y como se visualiza en la Figura 4.23. Esto significa que el patrón del flujo medido en B son muy similares entre ambos objetos.

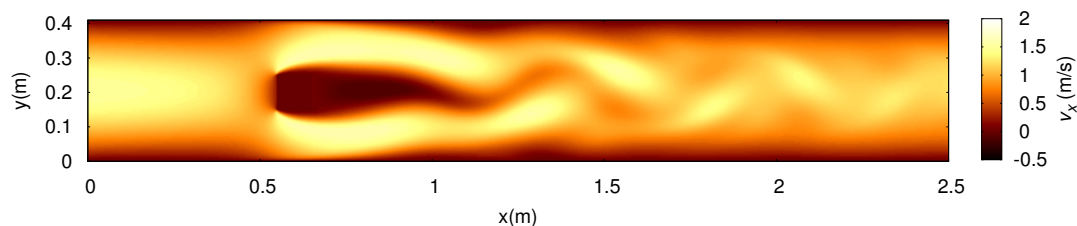


Figura 4.22: Mapa de calor que representa las magnitudes de  $v_x$  para una simulación de LBM para un flujo alrededor de un obstáculo cuadrado y con una velocidad característica de flujo de Poiseuille entrante con  $v_c = 1.5$  m/s. El centro del obstáculo se localiza en  $x = 0.6$  m y  $y = 0.205$  m. De la misma forma que para los obstáculos cilíndricos, los vórtices formados pasado el obstáculo son llevados a lo largo de la dirección del flujo, generando también la calle de vórtices de Karman.

Diámetro (m)	$v_c$ (m/s)	$R^2$ para $v_x$	$R^2$ para $q$
0.005	0.3	-23.753	-246.954
0.005	1.35	-0.645	-0.680
0.04	0.15	0.250	-4.518
0.04	0.75	0.966	0.919
0.04	1.8	0.975	0.975
0.08	1.5	0.961	-0.016
0.12	0.45	0.989	0.936
0.12	1.65	0.989	0.944
0.16	0.45	0.998	0.978
0.16	1.2	0.991	0.997
0.2	0.9	0.999	0.988
0.2	1.05	0.998	0.994
0.2	1.5	0.999	0.986
0.24	1.05	0.998	0.997
0.24	1.2	0.997	0.997
0.32	0.75	0.999	0.996
0.32	1.2	0.999	0.999
0.36	1.35	0.999	0.992
0.395	0.6	0.999	0.966
0.395	0.9	0.999	0.992

Tabla 4.7: Valores de  $R^2$  para el conjunto de predicción, considerando diez valores del perfil de  $v_x$  o  $q$  del flujo entrante antes del obstáculo, así como también diez valores en el sitio B. Excluyendo los tres obstáculos más pequeños, los resultados son prominentes con valores cercanos a 1.

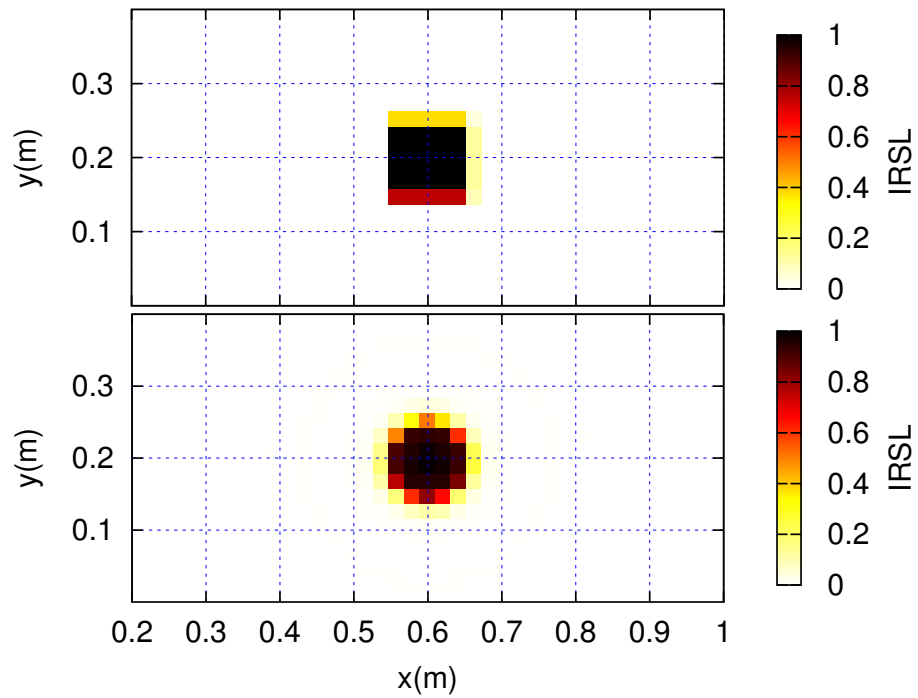


Figura 4.23: Predicción en términos del IRSL para un obstáculo cuadrado centrado en el ancho del tubo. En este caso se consideró como valores de entrada la serie de tiempo en tres distintos sensores en el sitio de medición B, tal y como en los casos 1c y 2c. Como era de esperarse, la RNA consideró este obstáculo de forma circular, dado que solamente fue entrenada para este tipo de formas. Sin embargo, sí muestra una ubicación y tamaño equiparable al objetivo real.



### 4.3.5. Conclusiones de la sección

Una serie de RNAs ha sido construidas siguiendo los lineamientos en distintos casos de estudio, y han sido entrenadas con la capacidad de estimar el tamaño y ubicación de un obstáculo. Las RNAs usan como vectores de entrada los perfiles de  $v_x$  o  $q$  a determinadas distancias. Los casos de estudio consideran la variación del diámetro del obstáculo, posición de este sobre una sección transversal del tubo y distintas velocidades de flujo entrante.

Se consideraron dos variables físicas ( $v_x$  y  $q$ ) para investigar si alguna es dominante como característica de entrada para eficacia de la RNA, sin embargo esto no fue claramente observado. Se inspeccionó si la cantidad de sensores sobre una sección transversal es mucho más relevante que una cantidad mínima de ellas. Y aunque sí fue el caso, en general se espera que sea más práctico hacer mediciones a lo largo del tiempo con pocos sensores.

De los resultados obtenidos, las RNA son altamente capaces de dar estimaciones que se consideran buenas, considerando un coeficiente de tipo  $R^2$  con valores mayores cercanos a 1 para obstáculos relativamente grandes. Aunque esto no siempre sucede para obstáculos pequeños, es decir, aquellos que tienen una relación aproximada entre el diámetro del tubo y el obstáculo menor a  $\beta = 1/4$ . Esto no es de sorprender, ya que las perturbaciones en el flujo para dichos tamaños se pueden disipar bastante rápido. Sin embargo, se observó que al incrementar la velocidad de flujo entrante, la precisión de la RNA mejora, como se constató en el caso 3, donde se usaron diferentes flujos de velocidad entrante, obteniendo en general valores superiores de  $R^2 > 0.9$  para diámetros mayores a 0.1 m.

## 4.4. Discusión de capítulo

Los métodos presentados en este capítulo son bastante sólidos para dar una idea de la resolución de la estimación de parámetros en una simulación de flujo alrededor de obstáculos cilíndricos en dos dimensiones. El primer caso se estudió la estimación del número de Reynolds para distintos perfiles de velocidad y vorticidad variando el número de sensores numéricos en la sección transversal del tubo en

distintos sitios de medición. Lo que llevó a que este enfoque podía ser capaz de obtener más parámetros utilizados para las simulaciones, como es la morfología y posición del obstáculo. Esto se desarrolló en la siguiente parte del capítulo, donde se propuso una metodología que fuese capaz de estimar regiones que están compuestas mayormente ya sea de elementos sólidos o del fluido. Esto último se ha hecho con la finalidad de posteriormente estudiar obstrucciones con obstáculos regulares, sino con morfologías más complicadas que sean más difíciles de caracterizar con un parámetro, como en este caso lo es el diámetro.

Con estos dos problemas se ha intentado investigar y proponer a los algoritmos de aprendizaje automático como una herramienta de resolución de problemas inversos en la dinámica de fluidos. Aunque se ha utilizado una herramienta sencilla como las RNAs multicapa, los problemas a resolver con algoritmos más complejos se pueden utilizar ya sea para dar estimaciones más precisas, obtener parámetros iniciales de este o algún otro problema donde exista la incertidumbre sobre alguna variable que determina la evolución del sistema como es observada.



# Capítulo 5

## Conclusiones

Este trabajo de tesis se ha enfocado en utilizar redes neuronales artificiales para la resolución de problemas inversos, dado que su implementación permite asociar una gran cantidad de características con las observaciones directas de algún fenómeno. En este caso fueron empleadas para determinar los parámetros implicados en dos fenómenos físicos: el análisis de señales provocadas por ondas gravitacionales y reconocimiento de perfiles en flujos alrededor de un cilindro en un tubo de dos dimensiones.

En relación al análisis de las formas de onda de dos distintos catálogos conformados por las soluciones numéricas de ondas gravitacionales causadas por el colapso de sistemas binarios de agujeros negros. Para ello, se utilizaron dos catálogos independientes, el primero de ellos realizado en el Instituto de Física y Matemáticas [Gracia, (2017)], se caracteriza por tener una fase orbital de media órbita, de tal manera que se inspecciona con las RNAs la sección de la forma de onda correspondiente a la colisión del sistema binario en el estudio de la razón de masas  $q = m_1/m_2$ , esto es, la fusión de los agujeros negros y decaimiento a una fase más estable como un único agujero negro. El segundo, realizado por el grupo de investigación del Georgia Tech [Jani et al., (2016)], cuenta con una fase orbital mucho más extendida, y por ello, se examinó si la inclusión brindaba una mayor información sobre el parámetro  $q$ . En ambos enfoques se hizo el análisis de las formas de onda simuladas considerando la serie de tiempo, tanto libres de ruido como con la inclusión de ruido Gaussiano desde un  $SNR = 0.5 - 25$  en el primer

caso y  $SNR = 5 - 25$  para el segundo. Además, dado que en el segundo enfoque se cuenta con una fase orbital bastante larga, se permitió hacer un análisis en el espectro de frecuencias mediante la transformada de Fourier. En general, cuando las señales están libres de ruido, las RNAs obtuvieron una predicción en promedio incluso menor al 1% en el mejor de los casos. Sin embargo, la inclusión del ruido afectaba enormemente la precisión en la estimación llegando a niveles del 10% o incluso mayores, dependiendo de si las formas de onda corresponden a valores de  $q$  que se encuentran lejos del espacio de parámetros de interpolación. Sin embargo, también en el análisis de frecuencias las predicciones se vieron ligeramente afectadas, amortiguando la presencia del ruido Gaussiano.

En este aspecto, es de suma importancia continuar con la aplicación de métodos de aprendizaje automático, dado que este tipo de sistemas son muy adecuados a la problemática que representan, esto es, el reconocimiento de una señal correspondiente a un evento físico que se encuentra embebida dentro de un ruido considerable, corrección de fallas en la instrumentación y finalmente la resolución del problema inverso en la estimación de los parámetros responsables de la fenomenología que se observa. Además, de ser posible realizar una extensión de este método hacia la gran cantidad de parámetros que involucran los distintos eventos generadores de ondas gravitacionales.

En términos de los problemas en flujos en un tubo dos dimensiones con algún obstáculo inmerso dentro de él se obtuvieron resultados mixtos dependientes de las características tanto del flujo como del obstáculo. Por un lado, en el primer caso las RNAs fueron altamente capaces de hacer la estimación del número de Reynolds con una buena estimación, menor al 4% de error, en general para los casos en que el flujo no es laminar, esto es para  $Re$  mayores a 60 y hasta 120. Por otro lado, la estimación de la forma y posición relativa a una sección transversal del tubo, tuvo resultados mixtos. Se encontró que en la presencia de cuerpos pequeños dentro del tubo, menores al 20%-30% entre la relación del grosor interno del conducto y el tamaño del obstáculo, las estimaciones de las RNAs no eran muy precisas ni en el área relativa que representan en su ubicación, esto debido a que la perturbación provocada por ellos es relativamente pequeña haciéndolos prácticamente imper-

ceptibles a una cierta distancia de medición. Por otro lado, para los obstáculos más grandes, los distintos enfoques fueron satisfactorios donde el ajuste realizado presentó casi una predicción perfecta. Dado que puede ser descomunal la variedad de fenómenos y configuraciones que se pueden presentar en sistemas de este tipo, es necesario desarrollar metodologías e implementar herramientas que puedan ser implementadas en el uso real.

Es inminente el incremento del uso de métodos de aprendizaje automático, conviviendo con las técnicas estadísticas tradicionales, en problemas donde la cantidad de información es inmensa o donde la estimación de parámetros requiere la resolución computacional e iterativa de un problema en un tiempo razonable. Siendo que estos métodos resultan muy prácticos en la resolución de problemas inversos donde la velocidad de respuesta puede ser crítica. En particular en el ramo de la experimentación y simulación de fenómenos físicos esta metodología es útil para realizar una rápida respuesta por parte de la instrumentación y/o dar aproximación sobre el valor de un parámetro involucrado en una observación en específico. Este trabajo ha dejado una gran enseñanza sobre la dirección a enfocar los esfuerzos en realizar investigación y herramientas de gran impacto.



# Apéndice A

## Implementación Algoritmo de Backpropagation

Para sintetizar como codificar una RNA e implementar una manera de implementar algoritmo, se presenta un pseudo-código considerando una RNA con las especificaciones utilizadas en el capítulo 2, con  $n$  neuronas de entrada, una sola capa oculta con  $m$  elementos y  $l$  neuronas de salida. Siguiendo las ecuaciones (2.5)-(2.7) y (2.15), con los resultados de las capas ocultas y de salida etiquetadas por  $h$  y  $y$  respectivamente, y los vectores tanto de entrada como los de objetivo escritos de la forma  $\mathbf{x}^p = (x_1^p, x_2^p, \dots, x_n^p)$  y  $\mathbf{t}^p = (t_1^p, t_2^p, \dots, t_l^p)$  respectivamente, con  $p$  indizando cada uno de los patrones de entrenamiento del par  $\{\mathbf{x}, \mathbf{t}\}$ .

Continuando con la descripción del algoritmo, este se puede dividir en dos partes: la propagación hacia adelante y la modificación de los pesos propagando el error hacia atrás, a las capas anteriores. Además se supone que estos dos procesos se continúan repetidamente hasta alcanzar un máximo de iteraciones  $M_T$ .

**mientras** iter  $\leq M_T$

**% Propagación hacia adelante %**

**%** Introducir cada vector de entrada a la RNA, obteniendo el vector de salida correspondiente  $\mathbf{y}^p$  **%**

**para cada** vector de entrada  $\mathbf{x}^p$  &  $\mathbf{t}^p$



```

% Introducir  $\mathbf{x}^p$  en la RNA. %
para cada neurona de salida  $y_k$  en la capa de salida  $O$ 
  para cada neurona de oculta  $y_j$  en la capa oculta  $H$ 
    para cada neurona en la capa entrada  $I$ 
       $\sigma_j = \sigma_j + w_{ij}x_i$ 
      fin para cada valor de salida  $y_i$  en la capa entrada  $I$ 
       $h_j = 1/(1 + \exp(-\sigma_j))$ 
       $\sigma_k = \sigma_k + w_{jk}h_j$ 
      fin para cada neurona oculta  $y_j$  en la capa oculta  $H$ 
       $y_k = 1/(1 + \exp(-\sigma_k))$ 
    fin para cada neurona de salida  $y_k$  en la capa de salida  $O$ 

% Modificación de los pesos de la RNA %
% Calculado el vector de salida  $\mathbf{y}^p$ , se obtiene el error  $E^p = \frac{1}{2l}(\mathbf{t}^p - \mathbf{y}^p)^2$  %
% modificando los pesos de la RNA de acuerdo a la ecuación (2.15) %
para cada neurona salida  $y_k$ 
  % Propagar error a las neuronas ocultas y de entrada %
  para cada peso  $w_{jk}$  de la neurona oculta  $j$  a la de salida  $k$ 
    para cada peso  $w_{ij}$  de la neurona de entrada  $i$  a la neurona oculta  $j$ 
       $w_{ij} = w_{ij} - \gamma [(y_k - t_k) (1 - y_k)y_k] w_{jk}h_j(1 - h_j)x_i$ 
    fin para cada peso  $w_{ij}$  de la neurona de entrada  $i$  a la neurona oculta  $j$ 
       $w_{jk} = w_{jk} - \gamma [(y_k - t_k) (1 - y_k)y_k] h_j$ 
    fin para cada peso  $w_{jk}$  de la neurona oculta  $j$  a la de salida  $k$ 
       $E^p = E^p + \frac{1}{2l}(t_k - y_k)^2$ 
  fin para cada neurona de salida
  % Calcular Error Global  $E_G$  %
   $E^G = E^G + E^p$ 
  Reiniciar  $E^p$ 
fin para cada vector de entrada y salida  $\mathbf{x}^p$  &  $\mathbf{y}^p$ 
  Reiniciar  $E^G$ 
  iter = iter + 1

```

**fin mientras**  $\text{iter} \leq M_T$

Hay que remarcar que el pseudo-código es una ejemplificación para implementar el algoritmo de *backpropagation*, hay varias formas de mejorarlo lo que varia la rapidez de ejecución, como puede ser la representación matricial tanto del conjunto de ejemplos, como los pesos de las conexiones de las neuronas entre capas. Además, por simplicidad de lectura, no se han insertado cada uno de los vectores de entrada de forma aleatoria así como tampoco utilizar el método de validación cruzada como criterio de parada, tal y como se utilizó a lo largo del trabajo.



# Apéndice B

## Prueba de Código de la RNA

Para mostrar la capacidad de la RNA en el código que se desarrolló, se mostrará una prueba de clasificación con un ejemplo clásico utilizando una base de datos de UCI Machine Learning Repository [Lichman,(2013)]. Este repositorio es una colección de bases de datos y generador de información que es utilizada por la comunidad de aprendizaje automático para el análisis empírico de algoritmos. Uno de estos conjuntos de datos es el reconocimiento óptico de dígitos escrito a mano, también conocido como *Optdigits Dataset*.

### B.1. Reconocimiento de dígitos

El conjunto muestral de los dígitos escrito a mano recolectado en esta base de datos, ya ha sido preparado para su implementación, de tal manera que la representación de los números del 0 al 9 son representados por matrices binarias de tamaño de  $32 \times 32$ . Ver la Figura B.1 donde se representan los ejemplos de estas matrices para 0 y 7.

Dado que el conjunto de datos se conforma por matrices de  $32 \times 32$ , la RNA tiene 1024 neuronas de entrada. Mientras que los vectores de salida se han dispuesto de tal forma que una neurona de salida represente cada dígito. Por ejemplo, si el número 0 es el patrón de entrada, el vector objetivo es  $\mathbf{t} = (1, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ ; mientras que para el número 7 significa que  $\mathbf{t} = (0, 0, 0, 0, 0, 0, 1, 0, 0, 0)$ .

Como se ha revisado a lo largo del trabajo, existen diferentes funciones de error

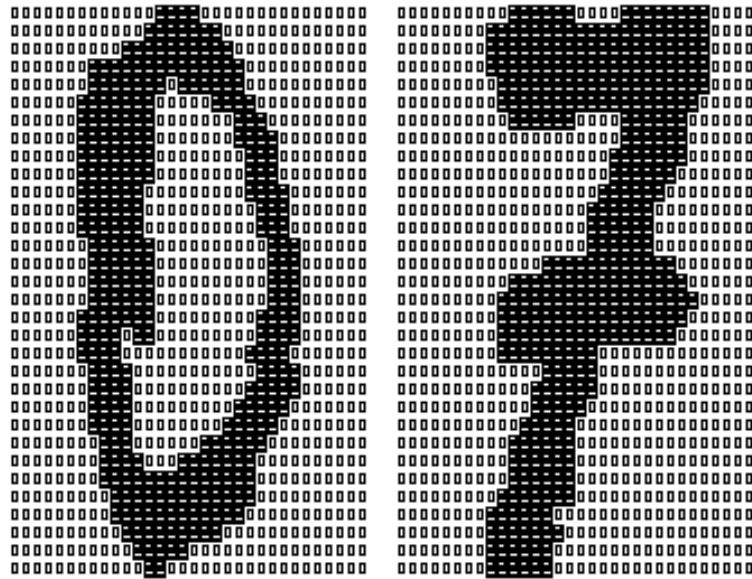


Figura B.1: Muestra de la representación de los dígitos extraídos del catálogo del UCI Machine Learning Repository. Los dígitos son originalmente escritos a mano y mapeados a una representación en píxeles para posteriormente reemplazados por valores binarios de 0 y 1. Fuente [Lichman,(2013)].

o de costo que se pueden minimizar para llegar a la resolución del problema. La elección de dicha función es decisión del usuario y/o programador, basándose si es un problema de clasificación o regresión. Y las funciones de activación se eligen el tipo de respuesta a obtener en la capa de salida como puede ser una probabilidad, una votación o un valor numérico real. Para demostrar, el poder y flexibilidad de la RNA desarrollada, se comparan el desempeño del ejemplo mediante dos enfoques diferentes:

- El reconocimiento de los dígitos utilizando una RNA con una función de activación de tipo sigmoide logística en la capa de salida, optimizando la función de costo de error cuadrático medio que ya ha sido revisada anteriormente.
- La clasificación de los dígitos utilizando la función de activación de normalización exponencial o *softmax* en la capa de salida, minimizando la función de entropía cruzada.

Primero se revisará la estructura de RNA utilizada tanto el primer enfoque

como en el segundo, así como las condiciones para la resolución del ejemplo. Posteriormente, se presentarán y contrastan los resultados de ambos procedimientos.

### B.1.1. Clasificación con función de costo de error cuadrático medio

Para la resolución del problema bajo este enfoque, se aprovecha el hecho que tanto los dígitos de entrada como los valores esperados están representados de forma binaria, y es natural pensar en la función de activación logística como la más adecuada, tanto en la capa oculta como en la de salida. Posteriormente, para determinar si la RNA ha acertado en su predicción del dígito representado, se ha hecho la normalización de los valores de todas las neuronas de salida. Esto es, el vector resultante es dividido por la suma de todos los valores de salida:

$$d_k = \frac{y_k}{\sum_{k=1}^{10} y_k}, \quad (\text{B.1})$$

donde  $y_k$  es el  $k$ -ésimo valor de salida, mientras que  $d_k$  significa la probabilidad asociada al dígito que representa. Por ejemplo, si el vector de entrada es el número 1, y la RNA arrojó el vector con valores:

$$\mathbf{y} = (0.2, 0.4, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2, 0.2). \quad (\text{B.2})$$

Haciendo la transformación de la ecuación (B.1) se traduce en el vector:

$$\mathbf{D} = (0.091, 0.182, 0.091, 0.091, 0.091, 0.091, 0.091, 0.091, 0.091, 0.091). \quad (\text{B.3})$$

Aunque en esta ejemplificación, el valor más alto no sea cercano al 1 esperado en el segundo elemento, la ubicación de dicho valor indica la posición más probable del dígito verdadero, en este caso con una probabilidad de 0.182%.

Recordemos que para el caso en la clasificación de costo de error cuadrático medio se ha utilizado la ecuación (2.12), que en este caso en particular se escribe como:

$$E = \frac{1}{10 \cdot N} \sum_{p=1}^N \sum_{k=1}^{10} \frac{1}{2} (t_k^p - y_k^p)^2 \quad (\text{B.4})$$

donde  $p$  indiza los  $N$  elementos que existen, ya sea en el conjunto de entrenamiento, validación o predicción.

### B.1.2. Clasificación con función de costo de entropía cruzada

A diferencia del método descrito anteriormente, la función de activación de normalización exponencial, no depende únicamente del valor de una sola neurona sino que considera el valor de los demás neuronas en la capa de salida. Esto se hace con la finalidad de representar la respuesta de la RNA como una probabilidad a cada vector de entrada, y no hacer una normalización a posteriori, como en el método utilizado antes. En términos del problema que se está abordando, la función de normalización exponencial se define matemáticamente por:

$$d_k = \frac{\exp(\sigma_k)}{\sum_k^{10} \exp(\sigma_k)}, \quad (\text{B.5})$$

donde  $d_k$  representa la probabilidad del dígito predicho por la RNA,  $k$  es el índice sobre cada una de las neuronas de salida, tal como en la sección anterior. Mientras que  $\sigma$  es la combinación lineal de los vectores de pesos  $\mathbf{w}_k$  y respuestas  $\mathbf{h}$  provenientes de la capa oculta hacia la  $k$ -ésima neurona de salida, tal como se ha revisado en 2.3.1:

$$\sigma_k = \mathbf{w}_k^T \cdot \mathbf{h}. \quad (\text{B.6})$$

De la ecuación (B.5), se ve que al igual que la ecuación (B.1), la suma de todos los elementos del vector resultante es igual a 1, y del valor más alto se puede inferir la probabilidad del dígito que considera la RNA que se le ha presentado. Para el entrenamiento de una RNA con esta función de activación en la capa de salida, la función de error no suele ser la MSE, sino la llamada función de entropía cruzada definida en este caso como:

$$E = - \sum_k^{10} t_k \ln(d_k) \quad (\text{B.7})$$

El cálculo del algoritmo de *backpropagation* de esta función de error, se sigue tal y como se mostró en 2.4.2. La iteración del algoritmo de aprendizaje con esta

función es tal que si un término en el vector de salida predomina, i.e., el resultado en la neurona de salida se acercará rápidamente a 1, mientras que los demás a tienden a 0.

Estas funciones de costo y de activación así como las normalizaciones son únicamente una ejemplificación para el entrenamiento de una RNA con el problema que se aborda. Y aunque existen otras formas de resolverlo, son suficientes para mostrar la capacidad del tipo de redes neuronales construidas con el código que se elaboró para este trabajo.

### B.1.3. Resultados

Continuando con el ejemplo, es hora de ver cual es el desempeño de las RNAs para el reconocimiento de dígitos, comparando los aciertos porcentuales entre ambos enfoques. En el entrenamiento de la RNA siguieron los procedimientos presentados tanto en B.1.1 como B.1.2.

La estructura de todas las RNAs construidas consiste en una capa de entrada con 1024 neuronas y 10 neuronas como salida, mientras que se ha explorado usar dos capas ocultas con distintos números de elementos. En ambos procedimientos se utilizaron 1000 elementos de la base de datos como conjunto de entrenamiento, 500 de validación y los restantes 433 como el conjunto de prueba o de predicción. También se ha utilizado la técnica de validación cruzada, en la cual se guardan los pesos de las RNAs, cuando en la fase de entrenamiento se ha alcanzado el mínimo en el error de validación. Una ejemplificación de esto se presenta en la Figura B.2 donde se ha utilizado la función de costo de error cuadrático medio.

De los resultados mostrados en las tablas B.2 y B.1, se puede ver que con una única neurona oculta, la RNA es incapaz siquiera de realizar un entrenamiento apropiado, quedándose estancada en un 20% de efectividad en ambos casos. Al incrementar a 5 neuronas ocultas en una sola capa, la razón de aciertos incrementa hasta el 97% usando la entropía cruzada, mientras que para el error cuadrático medio en un 90%. Sin embargo, para el conjunto de predicción esta razón baja hasta el 88% y 77% respectivamente. Incrementando un poco la efectividad en este caso cuando se agrega una capa oculta adicional con el mismo número de



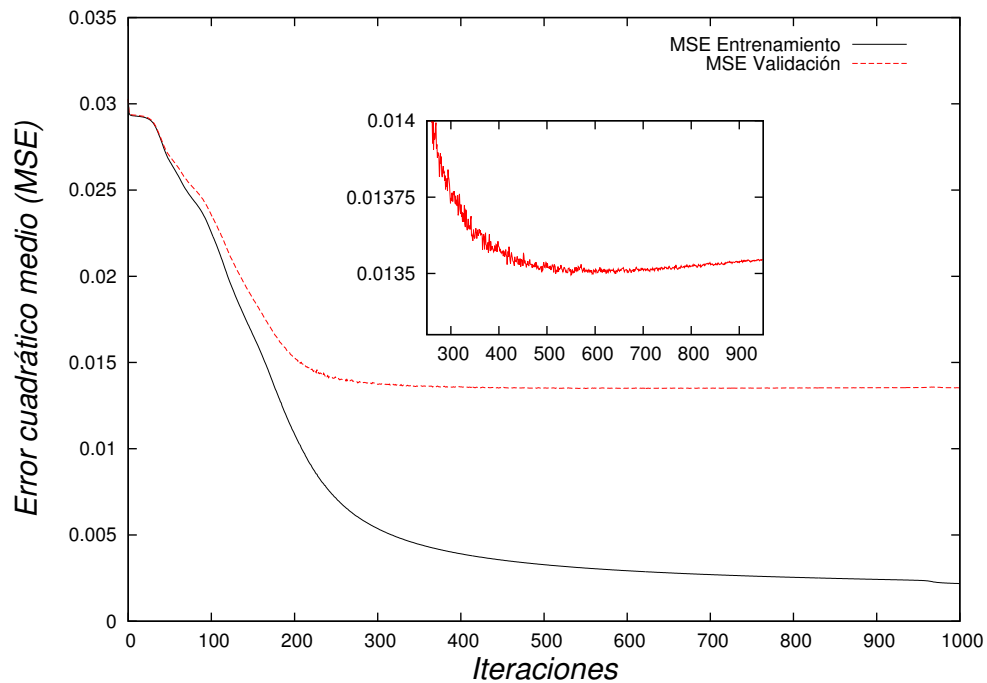


Figura B.2: Gráficas de error que se obtienen tanto el conjunto de entrenamiento (línea negra) y validación (línea roja) en función del número de iteraciones en la fase de aprendizaje del reconocimiento de dígitos por una RNA entrenada con la función de costo del error cuadrático medio definida en la ecuación (B.4). El recuadro dentro de la gráfica es un acercamiento para el error sobre el conjunto de validación, el mínimo aquí se encuentra en la iteración 594. En este caso la RNA cuenta con 2 capas ocultas con 10 elementos cada una, una constante de aprendizaje  $\gamma = 0.1$ , momento  $\alpha = 0$ .

neuronas. A partir 10 neuronas ocultas en una sola capa, no es perceptible un incremento en el rendimiento de la RNA ni aumentando el número de neuronas o capas ocultas en ambos casos. Aunque, en algunos casos se observa una razón de aciertos en el entrenamiento casi del 100 %, el error para valores desconocidos disminuye en alrededor del 5 %.

De las mismas tablas, también se puede inferir que en este caso, no hay una diferencia clara entre usar la función sigmoide o una normalización exponencial; una función de costo de error cuadrático o una entropía cruzada respectivamente. Sin embargo, algo que se puede agregar, es que el segundo caso llegaba en general en un número de iteraciones menor a la tasa de éxito que el primero. Hay que remarcar

Estructura RNA		Parámetros	Precisión (%)		
Capas ocultas	Neuronas ocultas	Constante de Aprendizaje	Entrenamiento	Validación	Predicción
1	1	0.01	20.3	19.185	18.244
1	5	0.01	90.2	77.170	77.136
1	10	0.01	99.8	92.39	93.53
1	20	0.01	99.7	93.462	95.15
1	50	0.01	99.8	93.569	94.457
1	20	0.035	99.8	94.319	96.074
2	5	0.01	98.6	83.065	84.526
2	20	0.01	99.8	93.248	94.457
2	20	0.035	99.8	92.283	93.995

Tabla B.1: Resultados para el reconocimiento de dígitos, utilizando una RNA con una capa de salida con función de activación logística, minimizando la función de error cuadrático medio.

que en la literatura, se menciona que el segundo enfoque es más recomendable, dado que el primero puede sobreestimar datos que sean considerados como puntos aislados (*outliers*, por su nombre en inglés). En general la RNA, es capaz de tener margen de error menor al 10% para casi todas las estructuras investigadas.

Estructura RNA		Parámetros	Precisión (%)		
Capas ocultas	Neuronas ocultas	Constante de Aprendizaje	Entrenamiento	Validación	Predicción
1	1	0.01	20.3	19.185	18.244
1	5	0.01	99.7	88.96	88.914
1	10	0.01	100.0	91.747	93.071
1	20	0.01	99.9	91.961	93.303
1	50	0.01	100.0	93.783	95.150
1	20	0.035	99.9	92.497	93.764
2	5	0.01	99.5	82.744	82.217
2	20	0.01	100.0	91.318	91.224
2	20	0.035	99.700	92.069	91.455

Tabla B.2: Resultados para el reconocimiento de dígitos, utilizando la función de activación de normalización exponencial en la capa de salida, y minimizando la función de costo de entropía cruzada.

# Apéndice C

## Ondas Gravitacionales

Este apéndice está fundamentalmente basado en las referencias [Saulson, (1994), Caltech: Educational resources, (2017)] y relata la base teórica de la generación de ondas gravitacionales, en particular de las producidas por el colapso de un sistema binario de agujeros negros, dando así un antecedente del capítulo 3.

En la teoría de Relatividad General, la gravedad se describe como una propiedad geométrica del espacio-tiempo, el cual se deforma en la presencia de masa y energía. Esto significa que el movimiento acelerado de cuerpos masivos produce deformaciones significativas en el espacio-tiempo, las cuales se propagan desde la fuente a la velocidad de la luz. Estas deformaciones son las llamadas ondas gravitacionales (GWs por sus siglas en inglés). Desde los años 90's se han dedicado esfuerzos para detectar dichas perturbaciones, a través de distintos proyectos experimentales como el Laser Interferometer Gravitational Wave Observatory (LIGO), Advance Virgo en Italia, KAGRA en Japón, entre otros [Caltech: Educational resources, (2017)]. Estos experimentos están continuamente en busca de señales de GWs como evidencia directa de fuentes astrofísicas. Sin embargo, la perturbación provocada por las GWs puede ser tan pequeña, que se pensaba jamás podrían ser detectadas. Con el avance de la tecnología hacia la detección de las GWs, ha sido necesario desarrollar herramientas apropiadas para el análisis de las señales y determinar las propiedades astrofísicas de las fuentes. Algunas de estas fuentes son la coalescencia de cuerpos binarios, tales como dos agujeros negros o estrellas de neutrones.

Las ondas gravitacionales son una solución directa de la teoría de Relatividad General desarrollada por Einstein. Al igual que en las ecuaciones de Maxwell que describen la interacción electromagnética, en dicha teoría existen soluciones tipo onda, donde las perturbaciones generadas por la masa y energía se propagan a la velocidad de la luz, a diferencia de la teoría de gravitación de Newton, donde la propagación es instantánea. La teoría de Relatividad General se describe de forma simple y elegante por las ecuaciones:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{8\pi G}{c^4}T_{\mu\nu}, \quad (\text{C.1})$$

donde  $T_{\mu\nu}$  es el tensor de energía-momento, el cual representa la masa y energía en una región del espacio-tiempo determinado;  $g_{\mu\nu}$  y  $R_{\mu\nu}$  son los tensores métricos que describen la configuración de masa y energía. En una aproximación de campo débil, es decir, donde métrica es casi plana, se pueden ignorar las contribuciones no-lineales de la métrica de espacio-tiempo  $g_{\mu\nu}$ , lo que es llamada como la teoría linealizada de la gravitación.

Como en la relatividad especial, donde el espacio-tiempo están unidos de tal manera que se genera un sistema coordinado tetra-dimensional, introduciendo una nueva métrica con la propiedad de ser invariante para todos los observadores. Este intervalo espacio-tiempo  $ds$  entre dos puntos se define como:

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2 = \sum_{\mu,\nu=0}^3 \eta_{\mu\nu}dx^\mu dx^\nu \quad (\text{C.2})$$

donde  $\eta_{\mu\nu}$  es llamada la métrica de Minkowski y los índices  $\mu$  y  $\nu$  representan el sistema de coordenadas de tiempo  $t$  y espacio  $x, y$  y  $z$  respectivamente. En el sistema de coordenadas cartesianas con un espacio-tiempo plano, es decir, sin perturbaciones el tensor de Minkowski se escribe como:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{C.3})$$

Sin embargo, está métrica no es cierta en relatividad general donde se considera que el espacio-tiempo puede ser curvo. Tomando en cuenta que las GWs curvan el espacio-tiempo ligeramente, se puede considerar a la métrica  $g_{\mu\nu}$  como una suma de pequeñas perturbaciones  $h_{\mu\nu}$  en este espacio-tiempo plano tal que en un aproximación de campo débil se tiene:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \quad h_{\mu\nu} \ll 1 \quad (\text{C.4})$$

El asumir que  $h_{\mu\nu}$  es muy pequeña se permite ignorar cualquier perturbación mayor al primer orden de esta cantidad, tal que se mantenga la parte lineal de la Relatividad General. Bajo estas suposiciones la ecuación de Einstein toma la forma de una ecuación de onda:

$$\left( \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) h_{\mu\nu} = 0, \quad (\text{C.5})$$

con soluciones de la forma:

$$h = (2\pi ft - \mathbf{k} \cdot \mathbf{x}), \quad (\text{C.6})$$

donde  $f = |\mathbf{k}|/2\pi c$  representa una onda plana viajando en la dirección  $\hat{k}$  a la velocidad de la luz  $c$ . Proponiendo una onda propagándose en la dirección  $\hat{k}$ , con  $h_{\mu\nu}$  de la forma:

$$h_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & a & b & 0 \\ 0 & b & -a & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{C.7})$$

El cual se puede escribir como la combinación lineal de dos tensores, definidos como  $h_+$  y  $h_\times$ , tal que:

$$h_+ = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{C.8})$$

y

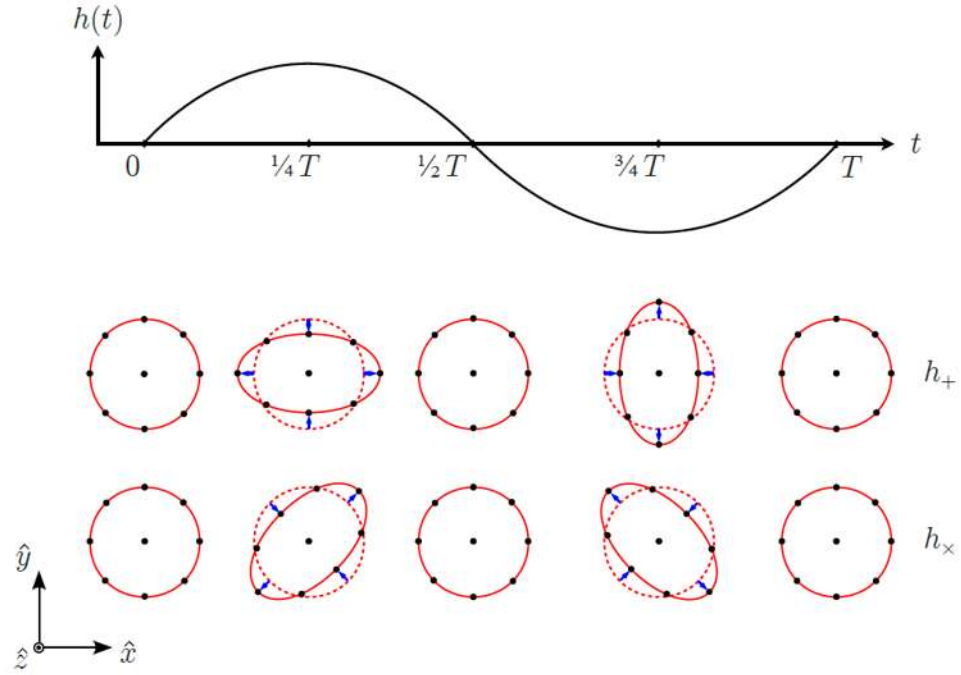


Figura C.1: En la figura se muestran los efectos de una onda gravitacional de periodo  $T$  que se propaga en la dirección del eje  $z$  sobre un anillo de partículas libres en reposo. Se distinguen los efectos de  $h_+$  y  $h_\times$  por separado. Imagen tomada de [Le Tiec & Novak, (2017)].

$$h_\times = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (\text{C.9})$$

Las GW son transversales y cuadrupolares por naturaleza, esto significa que al propagarse estiran y comprimen alternadamente el espacio en direcciones transversales en dirección ortogonal al mismo plano, y una polarización se vuelve la otra cuando su eje principal se rota  $45^\circ$ . Este efecto se muestra en la Figura C.1, donde se simula la interacción de una onda gravitacional perpendicular al plano sobre un anillo de partículas libres.

Dado que las ondas gravitacionales estudiadas en este trabajo son producto del colapso de sistemas binarios de agujeros negros, a continuación se da una breve

descripción sobre este tipo de GWs.

## C.1. Ondas gravitacionales de sistemas binarios de agujeros negros

En la señal ocasionada por la coalescencia de sistemas binarios de agujeros negros se pueden distinguir tres distintas fases: orbital, fusión y decaimiento, más conocidas por sus nombres en inglés como fases *inspiral*, *merger* y *ringdown* respectivamente.

En la fase *inspiral* u orbital, los dos agujeros negros se encuentran muy separados, y la órbita producida entre ellos puede ser descrita por una aproximación Post-Newtoniana [Blanchet, (2014)]. Al ir rotando y emitiendo radiación gravitacional, pierden energía y se acercan el uno al otro debido a la conservación del momento angular. Esto provoca que la emisión de radiación gravitacional aumente su frecuencia hasta alcanzar su máxima proximidad, en lo que se conoce como el “chirp” debido al paso de una baja a una alta frecuencia. En esta etapa las ecuaciones de Relatividad General se vuelven no-lineales y no existe un modelo analítico que lo describa, y es necesario el uso de métodos de relatividad numérica para obtener soluciones que satisfagan las ecuaciones de Einstein. El instante en que los agujeros negros se han unido en un único agujero negro se le conoce como fusión o *merger* por su nombre en inglés. Finalmente, el agujero negro resultante se encuentra en un estado excitado y continúa emitiendo radiación en la llamada fase de decaimiento o *ringdown*. En la Figura C.2, se muestra la simulación numérica que ha ajustado mejor a la GW150914, mostrando claramente las fases del colapso. En este caso en particular se estimó que la onda gravitacional fue provocada por dos agujeros negros a  $410_{-180}^{+160}$  Mpc de la Tierra, con masas iniciales aproximadas de  $m_1 = 36_{-4}^{+5} M_\odot$  y  $m_2 = 29_{-4}^{+4} M_\odot$ , resultando en un solo agujero negro con una masa estimada en  $62_{-4}^{+4} M_\odot$ , y una energía gravitacional radiada de  $3.0_{-0.5}^{+0.5} M_\odot c^2$ , donde  $M_\odot$  es la masa solar.



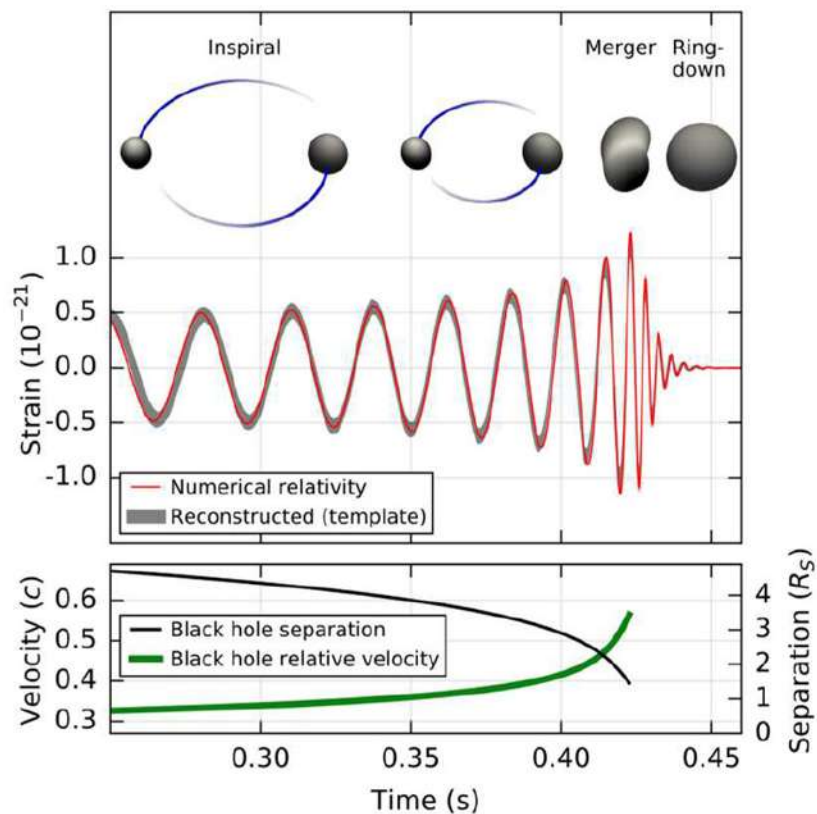


Figura C.2: Fases orbital, fusión y decaimiento (*inspiral*, *merger* y *ringdown* en inglés) de una onda gravitacional producida por la colisión de dos agujeros negros (arriba), esta en específico corresponde a la simulación numérica realizada para comparación con la GW151904 detectada por LIGO. El *strain* obtenido de los detectores fue filtrado y posteriormente es comparado con plantillas creadas por medio de simulaciones de relatividad numérica, representado en esta imagen por la línea roja. (Abajo) Estimación del cambio de la posición y velocidad relativa de los agujeros negros en su evolución al colisionar. Figura tomada de [Caltech: Educational resources, (2017)].

# Apéndice D

## Método de Lattice Boltzman

Este apéndice se basa principalmente en el libro [Mohamad, (2011)]. Por razones de auto-contenido del trabajo de tesis se resumen los aspectos importantes a comprender sobre el método de LBM que fue utilizado en la simulación de flujo alrededor de un cilindro en el capítulo 4.

Las ecuaciones de transporte para caracterizar variables físicas como el calor, masa o momento, pueden ser simuladas en diferentes escalas. Por un lado, en la escala macroscópica el uso de ecuaciones diferenciales parciales (EDP) como las de Navier-Stokes son utilizadas. Sin embargo, este tipo de ecuaciones pueden ser difíciles de resolver analíticamente debido a las no-linealidades presentes en las mismas ecuaciones, geometría y/o condiciones de frontera del problema. Por esa misma razón la mayoría de los problemas se resuelven mediante métodos numéricos computacionales como los de diferencias finitas (MDF), volúmenes finitos (MVF) o de elementos finitos (MEF).

Por otro lado, en la escala microscópica, se trata de caracterizar el sistema intentando determinar tanto la posición como la velocidad de cada una de las partículas. Esto conlleva al problema de tener que simular la dinámica molecular de una gran cantidad de elementos, incluso para el más sencillo de los problemas, cuando en la mayoría de los casos lo que se quiere conocer el efecto global hacia y desde el sistema.

La forma de integrar ambos enfoques ha sido a través del método de Boltzmann, en el cual no se intenta simular cada una de las partículas como en el caso

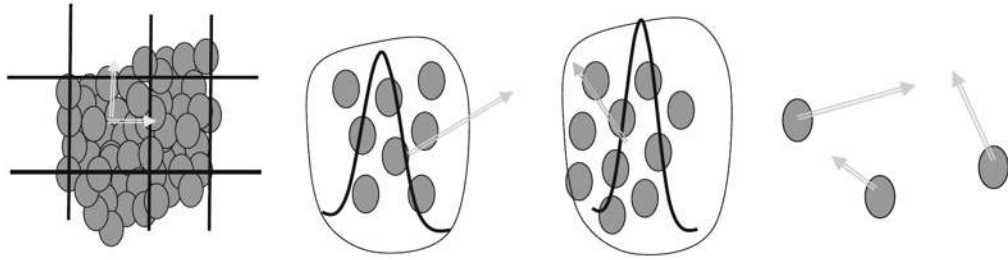


Figura D.1: En esta figura se esquematizan los distintos enfoques de modelación en la dinámica de fluidos. A la izquierda, se considera todo el sistema como un continuo, donde se utilizan los métodos de MDF, MVF y MEF. En medio se representa una colección de partículas en el espacio, las cuales son caracterizadas por medio de distribuciones estadísticas, en el cual se utiliza LBM para determinar sus direcciones de translación preferencial. Finalmente a la derecha en la escala microscópica, la dinámica de las partículas determinan las propiedades físicas de todo el sistema. Figura modificada de [Mohamad, (2011)].

de la dinámica molecular, sino por la caracterización estadística de una colección de partículas en un espacio determinado. Este modelo se basa construyendo modelos cinéticos que incluyan las variables físicas, de tal manera que el promedio de los valores obtenidos en todos los conjuntos sean los dominantes observados globalmente. La forma de caracterizar estos conjuntos de partículas es a través de funciones de distribución, haciendo más eficiente la simulación de la dinámica de fluidos, donde la cantidad de elementos es variable y con escalas de partículas muy grande. Los tres enfoques se pueden visualizar con ayuda de la Figura D.1.

En el método de *lattice* Boltzmann (LBM) los fluidos se consideran como una gran colección de pequeños elementos contenidos en un conjunto de celdas que definen el dominio físico en una malla numérica o *lattice*, en cuyos nodos las cantidades físicas son calculadas.

Los elementos en cada celda se representan como un ensamble de partículas con movimientos aleatorios descritos por una función de densidad de distribución  $f$ . A través de la ecuación de difusión de Boltzmann, dicha función de distribución  $f$  se puede interpretar como el intercambio de energía y momento debido al flujo

y colisión de las partículas:

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f = \Omega(f) \quad (\text{D.1})$$

donde  $\mathbf{v}$  es la velocidad del ensamble y  $\Omega$  es el llamado operador de colisión, que define el cambio entre el estado final e inicial de la distribución  $f$ . Una discretización de esta ecuación es la llamada aproximación BGK [Bhatnagar, Gross & Krook, (1954)]:

$$f_i(\mathbf{x} + \mathbf{c}_i \delta t, t + \delta t) = f_i(\mathbf{x}, t) + \frac{1}{\tau} (f_i^{eq} - f_i(\mathbf{x}, t)), \quad (\text{D.2})$$

donde  $f_i$  es la función de densidad de distribución en la red discreta en la dirección  $i$ ,  $\mathbf{c}_i$  es la velocidad de flujo también en la dirección  $i$ ,  $f_i^{eq}$  es la función de distribución en equilibrio y  $\tau$  el tiempo de relajación en el que  $f_i \rightarrow f_i^{eq}$ . El segundo término de la derecha es el operador de colisión que describe las interacciones microscópicas entre partículas y la variación del número de partículas que se mueven en cada dirección en la malla numérica.

La función de distribución de equilibrio, es usada para determinar la velocidad local del fluido debido a las colisiones y es calculada a través de las variables macroscópicas, tal que se preserven la masa, el momento y la energía de cada celda.

Aunque existen varios modelos de mallado para las posibles direcciones de velocidad, el más usual en dos dimensiones y que fue utilizado en este trabajo es el llamado modelo D2Q9, ver Figura D.2.

En dicho modelo los elementos de fluido siguen una dirección determinada a través de la red a lo largo de 9 diferentes direcciones en cada celda:

$$\mathbf{c}_i = \begin{cases} (0, 0) & i = 0 \\ (c, 0), (0, c), (-c, 0), (0, -c) & i = 1, 2, 3, 4 \\ (c, c), (-c, c), (-c, -c), (c, -c) & i = 5, 6, 7, 8 \end{cases}, \quad (\text{D.3})$$

con la función de distribución de equilibrio:

$$f_i^{eq}(\rho, \mathbf{v}) = \rho w_i \left( 1 + \frac{3}{c^2} \mathbf{c}_i \cdot \mathbf{v} + \frac{9}{c^4} (\mathbf{c}_i \cdot \mathbf{v}^2)^2 - \frac{3}{2c^2} \mathbf{v}^2 \right), \quad (\text{D.4})$$

donde  $\rho$  es la densidad,  $\mathbf{v}$  la velocidad,  $\mathbf{c}_i$  la velocidad en la malla definida por:

$$\mathbf{c}_i = \frac{\delta x}{\delta t}, \quad (\text{D.5})$$

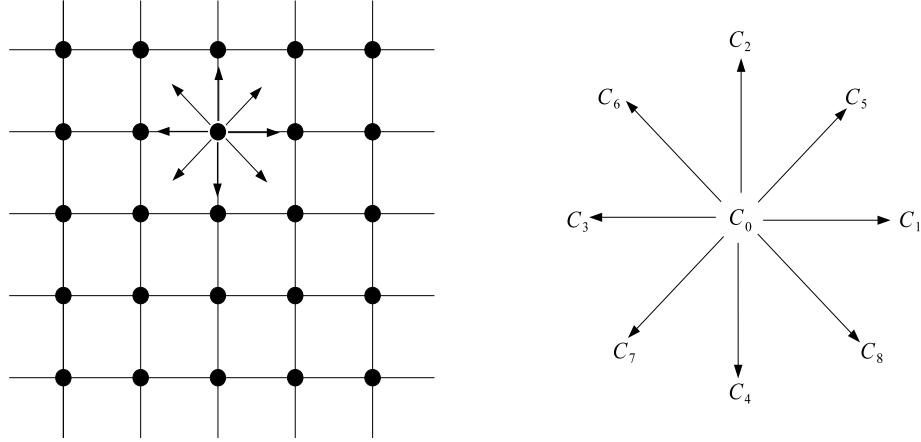


Figura D.2: Representación esquemática de la malla numérica utilizada en LBM y el modelo D2Q9 con posibles direcciones para  $f$  en dos dimensiones espaciales.

con  $\delta x$  el espacio entre nodos y  $\delta t$  el paso de tiempo. Y finalmente los  $w_i$  son un factor de peso para cada dirección:

$$w_i = \begin{cases} 4/9 & i = 0 \\ 1/9 & i = 1, 3, 5, 7 \\ 1/36 & i = 2, 4, 6, 8, \end{cases} \quad (\text{D.6})$$

Mientras que las variable macroscópicas como la densidad y la velocidad se calculan a través de las funciones de distribución  $f_i$  en el espacio de velocidades de las partículas mediante:

$$\rho(x, t) = \sum_i f_i^{eq}, \quad (\text{D.7})$$

$$\mathbf{v}(x, t) = \frac{1}{\rho} \sum_i f_i^{eq} \mathbf{c}_i. \quad (\text{D.8})$$

Desde un punto de vista computacional, se deben imponer condiciones de frontera tales que a una distancia finita suficientemente lejos, los parámetros de flujo no se vean afectados por los cálculos internos. La frontera interna de las paredes físicas internas son definidas de distintas formas, sin embargo en el trabajo se consideró las realizadas por Zou and He conocida como *Bounce Back* [Zou & He, (1997)]. Este tipo de condición asegura conservación de masas y momentos en la frontera y

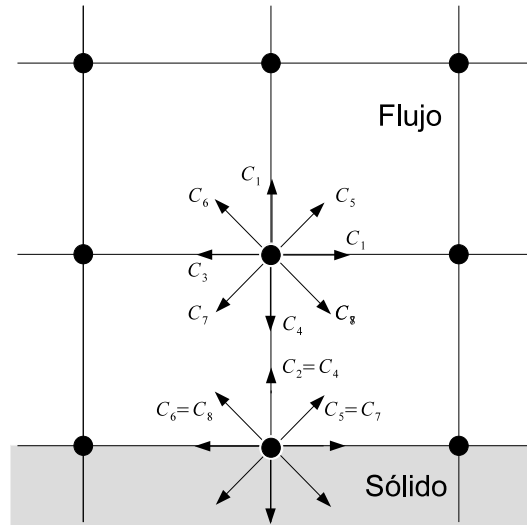


Figura D.3: Modelo de condición de frontera tipo *bounceback*. En el esquema se especifica el “rebote” que soporta la distribución  $f$ , específicamente en este modelo resulta que  $c_2 = c_4$ ,  $c_5 = c_7$  y  $c_6 = c_8$ .

se utiliza para modelar sólidos estacionarios sin deslizamiento. En este arreglo que se muestra en la Figura D.3 el “rebote” de los elementos en la distribución es tal que  $c_2 = c_4$ ,  $c_5 = c_7$  y  $c_6 = c_8$ .

Con estos antecedentes se han revisado los conceptos físicos básicos y consideraciones apropiadas para realizar la simulación del flujo alrededor de un obstáculo cilíndrico en dos dimensiones, para el trabajo descrito en el capítulo 4. En otros problemas pueda ser necesario tener mayores restricciones sobre las velocidades máximas que se permiten en el código así como en las condiciones de frontera.



# Glosario

**Algoritmo de *backpropagation*** Algoritmo de aprendizaje supervisado utilizado frecuentemente en redes neuronales artificiales basado en un método de gradiente descendente. 21, 27

**Aprendizaje automático** Programación de algoritmos capaces de manejar una gran cantidad de datos y encontrar una solución a partir únicamente de la información proporcionada para realizar o no alguna tarea específica con futuros datos desconocidos. 2, 5, 16, 21, 42, 85

**Capa límite** Región donde una capa de fluido es adyacente a una frontera sólida. 45, 46

**Escalar de Weyl** Elemento del tensor de Weyl que proporciona una medida sobre la curvatura del espacio-tiempo y el cual gobierna la propagación de las ondas gravitacionales en regiones del espacio vacío. 25, 109

**Estabilidad neutral** Término en la dinámica de fluidos para describir regímenes estables o inestables en el flujo de un fluido. 50

**Flujo de Poiseuille** Flujo de un fluido Newtoniano incompresible con un flujo laminar a través de un tubo cilíndrico en una sección transversal constante, en el cual se describe la caída de presión. XII, 45–47, 62, 63, 88, 91

**Punto de estancamiento** En la dinámica de fluidos, un punto de estancamiento es un punto en un campo de flujo donde la velocidad local del fluido es cero. Estos existen en la superficie de los objetos inmersos en el campo de flujo, donde el fluido es puesto en reposo debido a la interacción con el objeto. 46



**Razón señal-ruido** Razón entre la potencia de una señal con un significado físico y el ruido de la señal medida. 25, 31

**Strain** Deformación ocurrida en los detectores debido al paso de una onda gravitacional a través del espacio. La deformación del interferómetro provoca la señal observada en los detectores. VIII, XIII, 23, 25, 26, 36, 37, 109, 110

# Bibliografía

- [Abbot et al., (2016a)] Abbott, B. P. et al., (2016), *Observation of Gravitational Waves from a Binary Black Hole Merger*. Phys. Rev. Lett., 116 (6), 061102. doi: 10.1103/PhysRevLett.116061102
- [Abbot et al., (2016b)] Abbott, B. P. et al., (2016), *GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence*. Phys. Rev. Lett., 116 (24), 241103. doi: 10.1103/PhysRevLett.116.241103
- [Abbot et al., (2017)] Abbott, B.P. et at., (2017), *GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2*. Phys. Rev. Lett., 118 (22), 221101, doi: 10.1103/PhysRevLett.118.221101
- [Ajith et at., (2012)] Ajith, P. et al., (2012). *The NINJA-2 catalog of hybrid post-Newtonian/numerical-relativity waveforms for non-precessing black-hole binaries*. Class. Quantum Grav. 29(12):27, 124001. doi: 10.1088/0264-9381/29/12/124001
- [Alpaydin, (2010)] Alpaydin, E., (2010). *Introduction to Machine Learning*. 2da. edición, Cambridge, EE. UU., MIT Press.
- [Arnowitt, Deser & Misner, (1959)] Arnowitt, R., Deser, S. & Misner, C.W., (1959). *Dynamical Structure and Definition of Energy in General Relativity*. Phys. Rev., 116(5). doi: 10.1103/PhysRev.116.1322
- [Aster, (2003)] Aster, R.C., Borchers, B. & Thurber, C.H., (2003). *Parameter Estimation and Inverse Problems*. 2da. ed., Waltham, EE. UU., Academic Press. doi: 10.1016/B978-0-12-385048-5.00030-6
- [Baker et al., (2015)] Baker, P. T. et al., (2015). *Multivariate classification with random forests for gravitational wave searches of black hole binary coalescence*. Phys. Rev. D, 91 (6), 062004. doi: 10.1103/PhysRevD.91.062004
- [Baumgarte & Shapiro, (1999)] Baumgarte, T.W. & Shapiro S.L., (1999). *On the Numerical Integration of Einstein's Field Equations*. Phys. Rev. D, 59 (2), 24007. doi: 10.1103/PhysRevD.59.024007

- [Begg, Kamruzzaman & Sarker, (2006)] Begg, R., Kamruzzaman, J. & Sarker, R., (2006). *Neural Networks in Healthcare: Potential and Challenges*. IGI Global, Hershey, PA, EE. UU., 1-19. doi: 10.4018/978-1-59140-848-2
- [Bhatnagar, Gross & Krook, (1954)] Bhatnagar, P.L., Gross, E.P. & Krook, M., *A model for Collision Processes in Gases I. Small Amplitude Processes in Charged and Neutral One-Component Systems*, Phys. Rev., 94, 511-525. doi: 10.1103/PhysRev.94.511
- [Bishop, (1995)] Bishop, C., (1995). *Neural Networks for Pattern Recognition*. Nueva York, EE. UU., Oxford University Press.
- [Bishop, (2006)] Bishop, C., (2006). *Pattern Recognition and Machine Learning*. Nueva York, EE. UU., Springer.
- [Biswas et al., (2013)] Biswas, R. et al., (2013). *Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data*. Phys. Rev. D, 88 (6), 062003. doi: 10.1103/PhysRevD.88.062003
- [Blanchet, (2014)] Blanchet, L., (2014). *Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries*. Living Rev. Relativ. 17, 2. doi:10.12942/lrr-2014-2
- [Bogdan, (2009)] Bogdan, M.W., (2009). *Neural Networks Architectures and Learning*. IEEE Ind. Elec. Mag., 3(4). doi: 10.1109/MIE.2009.934790
- [Bortolan, Degani & Williams, (1991)] Bortolan, G., Degani, R. & Williams J.L, (1991). *ECG classification with neural networks and cluster analysis*. En: *Proceedings in Computers in Cardiology*, Venecia, Italia, IEEE, 177-180.
- [Caltech: Educational resources, (2017)] Caltech: Educational resources, (Junio, 2017). <https://www.ligo.caltech.edu/page/detection-education-resources>.
- [Carrillo, González, Gracia & Guzmán, (2016)] Carrillo, M., González, J.A., Gracia, M. & Guzmán F.S., (2016). *One parameter estimates in binary black hole collisions using neural networks*. Gen. Rel. Gravit., 48, 141. doi: 10.1007/s10714-016-2136-0
- [Carrillo, Que & González, (2016)] Carrillo, M., Que, U. & González, J.A., (2016). *Estimation of Reynolds number for flows around cylinders with lattice Boltzmann methods and artificial neural networks*. Phys. Rev. E, 94(6), 063304. doi: 10.1103/PhysRevE.94.063304
- [Carrillo, Que, González & López, (2017)] Carrillo, M., Que, U., González, J.A., & López, C., (2017). *Recognition of an obstacle in a flow using artificial neural networks*. Phys. Rev. E, 96(2), 023306. doi: 10.1103/PhysRevE.96.023306

- [Catalano, et al., (2003)] Catalano, P. et al., (2003). *Numerical simulation of the flow around a circular cylinder at high Reynolds numbers*. Int. J. Heat Fluid Flow, 24(4), 463. doi: 10.1016/s0142-727x(03)00061-4
- [Chen, Chen, Min, Fisher & Wu, (2003)] Chen, H.Y., Chen, T.C., Min, D.I., Fischer, G.W. & Wu, Y.M., *Therapeutic Drug Monitoring*, **21(1)**, (1999), 50.
- [Hon & Wei, (2003)] Hon, Y.C. & Wei, T. (2003). *A fundamental solution method for inverse heat conduction problem*, Engineering Analysis with Boundary Elements, 28 (5), 489-495. doi: 10.1016/S0955-7997(03)00102-4
- [Cirovic & Dragan, (1997)] Cirovic & A. Dragan, (1997). *Feed-forward artificial neural networks: applications to spectroscopy*. Trends Anal. Chem., 16(3), 148. doi:10.1016/S0165-9936(97)00007-1
- [Cybenko, (1989)] G. Cybenko, (1989) *Approximation by superpositions of a sigmoidal function*. Mathematics of Control, Signals and Systems, 2(4), 303. doi: 10.1007/BF02551274
- [Davies & Desai, (2008)] Davies, C. E. & Desai, (2008). *Blockage in vertical slots: Experimental measurement of minimum slot width for a variety of granular materials*. Pow. Tech. **183**, 436(3), 436. doi: 10.1016/j.powtec.2008.01.009
- [De Vito, E., et al.,(2005)] De Vito, E., Rosasco, L., Caponnetto, A., De Giovannini, U. & Odone, F., (2005). *Learning from examples as an inverse problem*. J. Mach. Learn. Res. 6, 883. doi:
- [Dollet, Durth & Graner, (2006)] Dollet, B., Durth, M. & Graner, F., (2006). *Flow of foam past an elliptical obstacle*. Phys. Rev. E, 73(6), 061404. doi:10.1103/PhysRevE.73.061404
- [Dorsey, et al., (1997)] Dorsey, S. G., Waltz, C. F., Brosch, L., Connerney, I., Schweitzer, Eugene, J. & Bartlett, S. T., (1997), *A Neural Network Model for Predicting Pancreas Transplant Graft Outcome*. Dia. Care, 20(7), 1128. doi:10.2337/diacare.20.7.1128
- [Du, Ke-Lin & Swamy, M.N.S., (2013)] Ke-Lin Du & Swamy, M.N.S., (2013). *Neural Networks and Statistical Learning*. Londres, Reino Unido, Springer-Verlag. doi: 10.1007/978-1-4471-5571-3
- [Eldon, (1994)] Eldon, Y. L., (1994). *Artificial neural networks and their business applications*. Information & Management, 27(5), 303. doi:10.1016/0378-7206(94)90024-8
- [Field, Galley, Hesthaven, Kaye & Tiglio, (2014)] Field, S. E., Galley, C. R., Hesthaven, J. S., Kaye, J. & Tiglio, M., (2014). *Fast Prediction and Evaluation of*

- Gravitational Waveforms Using Surrogate Models*. Phys. Rev. X, 4 (3), 031006. doi: 10.1103/PhysRevX.4.031006
- [Garliauskas, (1998)] Garliauskas, A., (1998). *Neural network chaos analysis*. Non-linear Analysis: Modelling and Control, 3(43).
- [Georgia Tech, (2017)] Georgia Tech, (2017). *Binary Black Hole Simulations*. <http://www.einstein.gatech.edu/catalog>
- [Gernoth, (1999)] Gernoth, K.A, (1999). *Nuclear physics with neural networks*. En: *Proc. Scientific Applications of Neural Nets*, Berlin, Alemania, Springer. doi: 10.1007/BFb0104279
- [Gracia, (2017)] Gracia, M., (2017), *Simulación de Acreción en Hoyos Negros y Choque de Hoyos Negros*. Tesis doctoral, Instituto de Física y Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo.
- [Graff, (2012)] Graff, P., (2012). *Bayesian Methods for Gravitational Waves and Neural Networks*. Tesis doctoral, Cambridge.
- [Grucelski & Pozorski, (2013)] Grucelski, A. & Pozorski, J., (2013). *Lattice Boltzmann simulations of flow past a circular cylinder and in simple porous media*. Comp. & Fluids, 71, 406. doi: 10.1016/j.compfluid.2012.11.006
- [Hebb, (1949)] Hebb, D.O., (1949). *The organization of behavior: A neuropsychological theory*. New York. EE. UU. Wiley.
- [Heskes, (1997)] Heskes, T., (1997). *Practical confidence and prediction intervals*, En: *Proc. of the 9th International Conference on Neural Information Processing Systems*, MIT press, 176.
- [Hulse & Taylor, (1975)] Hulse, R.A. & Taylor, J.H., (1975). *Discovery of a pulsar in a binary system*. Astro. Phys. J., 195, L51. doi: 10.1086/181708
- [Jani et al., (2016)] Jani, K., Healy, J., Clark, J. A., London, L., Laguna, P. & Shoemaker, D., (2016). *Georgia tech catalog of gravitational waveforms*. Class. Quantum. Grav., 33(20). doi: 10.1088/0264-9381/33/20/204001
- [Kaipio & Somersalo, (2005)] Kaipio, J. y Somersalo, E. (2005). *Statistical and computational inverse problems*. New York, EE. UU., Springer-Verlag. doi:10.1007/b138659
- [Karabelas et al., (2012)] Karabelas, S. J., Koumroglou, B. C., Argyropoulos, C. D. & Markatos, N. C., (2012). *High Reynolds number turbulent flow past a rotating cylinder*, Applied Mathematical Modelling, 36 (1), 379. doi: 10.1016/j.apm.2011.07.032

- [Kernan et al., (2014)] Kernan, W. N. et al., (2014). *Guidelines for the Prevention of Stroke in Patients With Stroke and Transient Ischemic Attack*. A Guideline for Healthcare Professionals. 45(2), E54. doi: 10.1161/str.0000000000000059
- [Kerr & Bartlett, (1995)] Kerr, J. P. & Bartlett, E. B., (1995). *A statistically tailored neural network approach to tomographic image reconstruction*. Med. Phys., 22(5), 601. doi:10.1118/1.597586
- [Khan et al., (2016)] Khan, S. et al.,(2016). *Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era*. Phys. Rev. D, 93 (4), 044007. doi: 10.1103/PhysRevD.93.044007
- [Khanam et al., (2016)] Khanam, P.N. et al.(2016), *Optimization and Prediction of Mechanical and Thermal Properties of Graphene/LLDPE Nanocomposites by Using Artificial Neural Networks*, Int. J. Pol. Sc., 5340252. doi:10.1155/2016/5340252
- [Knoll et al., (1999)] Knoll, P., Mirzaei, S., Muellner, A., Leitha, T., Koriska, K., Koehn, H. & Neumann M., (1999). *An artificial neural net and error backpropagation to reconstruct single photon emission computerized tomography data*. Med Phys., 26, 244. doi: 10.1118/1.598511
- [Kohonen, (1995)] Kohonen, T.,(1995). *The self-organizing maps*. Berlin, Alemania, Springer. doi: 10.1007/978-3-642-56927-2
- [Lacaze et al.,(2013)] Lacaze, L., et al., (2013). *Wave patterns generated by an axisymmetric obstacle in a two-layer flow*. Experiments in Fluids, 54(12), 10. doi:10.1007/s00348-013-1618-z
- [Larsen & Walther, (1998)] Larsen, A. & Walther, J. H., (1998). *Discrete vortex simulation of flow around five generic bridge deck sections*. J. Wind Eng. Ind. Aerodyn., 77(8), 591. doi: 10.1016/s0167-6105(98)00175-5
- [LeCun, Bottou, Orr, Muller,(2012)] LeCun, Y. A., Bottou, L., Orr, G. B. & Muller, K.R., (2012). *Efficient BackProp*. En: *Neural Networks: Tricks of the Trade*, 2da. edición, Berlin, Alemania, 9. doi: 10.1007/978-3-642-35289-8\_3
- [Le Tiec & Novak, (2017)] Le Tiec, A., & Novak, J., (2017). *Theory of Gravitational Waves*. En: *An Overview of Gravitational Waves*, Singapur, Singapur, World Scientific, 1. doi: 10.1142/97898131417660001
- [Lichman,(2013)] Lichman, M., (2013). *UCI Machine Learning Repository*, University of California, Irvine, School of Information and Computer Sciences. url: <http://archive.ics.uci.edu/ml>

- [Löffler, (2012)] Löffler, F. et al.,(2012). *The Einstein Toolkit: a community computational infrastructure for relativistic astrophysics*. *Class. Quantum Grav.* 29, 115001:2012. doi: 10.1088/0264-9381/29/11/115001  
url: <https://einsteintoolkit.org>
- [MacPherson, Conway & Brown, (1995)] Macpherson, K. P., Conway, A. J. & Brown, J. C., (1995). *Prediction of solar and geomagnetic activity data using neural networks*. *J. Geophys. Res.: Space Phys.*, 100(A11), 21735. doi: 10.1029/95JA02283
- [Martín et al., (2004)] Martín, J.D. et al, (2004). *Neural networks as effective techniques in clinical management of patients: some case studies*. *Trans. Inst. Measurement Control*, 26(3), 169. doi: 10.1191/0142331204tm118oa
- [McCulloch & Pitts, (1943)] McCulloch, W. S. & Pitts, W., (1943). *A logical calculus of the ideas immanent in nervous activity*. *The Bull. Math. Biophys.*, 5(4), 115. doi: 10.1007/BF02478259
- [McNelis, (2005)] McNelis, P. D., (2005). *Neural Networks in Finance: Gaining Predictive Edge in the Market*. Boston, EE. UU. Elsevier Academic Press.
- [Minsky & Papert, (1988)] Minsky, Marvin L. & Papert, S., (1988). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, EE. UU., MIT Press.
- [Mohamad, (2011)] Mohamad, A. A., (2011). *Lattice Boltzmann Method Fundamentals and Engineering Applications with Computer Codes*. Londres, Reino Unido, Springer.
- [Mroue, (2013)] Mroue, A. H. et al., (2013). *Catalog of 174 Binary Black Hole Simulations for Gravitational Wave Astronomy*. *Phys. Rev. Lett.*, 111 (24), 241104. doi: 10.1103/PhysRevLett.111.241104
- [Ramón y Cajal, (1899)] Ramón y Cajal, S., (1899). *Textura del sistema nervioso del hombre y los vertebrados*. Madrid, España, Nicolás Moya.
- [Rojas, (1996)] Rojas, R., (1996). *Neural Networks: A Systematic Introduction*. Berlín, Alemania, Springer-Verlag. doi: 10.1007/978-3-642-61068-4
- [Rosenblatt, (1962)] Rosenblatt, F., (1962). *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*, Washington, EE. UU., Spartan Books.
- [Rumelhart, Hinton & Williams, (1986)] Rumelhart, D. E., Hinton, G. E. & Williams, R. J., (1986). *Learning representations by back-propagating errors*, *Nature*, 323, 533. doi: 10.1038/323533a0

- [Russell & Norvig, (1995)] Russell, S. J. & Norvig, P., (1995). *Artificial Intelligence: A Modern Approach*. New Jersey, EE. UU., Prentice-Hall.
- [Saulson, (1994)] Saulson, P., (1994). *Fundamentals of Interferometric Gravitational Wave Detectors*. Singapur, Singapur, World Scientific.
- [Sathyaprakash & Schutz, (2009)] Sathyaprakash, B. S. & Schutz, B. F., (2009). *Physics, Astrophysics and Cosmology with Gravitational Waves*. Living Rev. Rel., 12. doi: 10.12942/lrr-2009-2
- [Schäfer & Turek, (1996)] Schäfer, M. & Turek, S., (1996). *Benchmark computations of laminar flow around a cylinder*. En: *Flow Simulation with High-Performance Computers II*, 48, 547. doi: 10.1007/978-3-322-89849-4\_39
- [Shibata & Nakamura, (1995)] Shibata, M. & Nakamura, T., (1995). *Evolution of three-dimensional gravitational waves: Harmonic slicing case*. Phys. Rev. D, 52(10), 5428, doi: 10.1103/PhysRevD.52.5428
- [Smith et al., (2013)] Smith, R. J. E., Cannon, K., Hanna, C., Keppel, D. & Mandel, I., (2013). *Towards rapid parameter estimation on gravitational waves from compact binaries using interpolated waveforms*. Phys. Rev. D, 87 (12), 122002. doi:10.1103/PhysRevD.87.122002
- [Sukop & Thorne, (2010)] Sukop, M. C. & Thorne, D. T., (2010). *Lattice Boltzmann Modeling: An Introduction for Geoscientists and Engineers*. Berlin, Alemania, Springer.
- [Tesauro, (1995)] Tesauro, G., (1995). *Temporal difference learning and TD-Gammon*. En: *Communications of the ACM*, 38(3).
- [Tesauro, Kephart & Sorkin, (1996)] Tesauro G. J., Kephart, J. O. & Sorkin, G. B., (1996). *Neural networks for computer virus recognition*. IEEE Expert, 11(4), 5-6.
- [Tobiyama et al., (2016)] Tobiyama, S., Yamaguchi, Y., Shimada, H., Ikuse, T. & Yagi, T., (2016). *Malware Detection with Deep Neural Network Using Process Behavior*. En: *IEEE 40th Annual Computer Software and Applications Conference*, 2, 577.
- [Ugarelli et al., (2010)] Ugarelli, R., Venkatesh, G., Brattebø, H., Di Federico, V. & Sægrov, S., (2010). *Historical analysis of blockages in wastewater pipelines in Oslo and diagnosis of causative pipeline characteristics*. Urb. Water J., 7(6), 335. doi: 10.1080/1573062X.2010.526229
- [Van der Sluys et al., (2009)] Van der Sluys, M., Mandel, I., Raymond, V., Kalogera, V., Rover, C. & Christensen, N., (2009). *Parameter estimation for signals*



- from compact binary inspirals injected into LIGO data.* *Class. Quantum Grav.*, 26(20), 10. doi: 10.1088/0264-9381/26/20/204010
- [Veitch et al., (2015)] Veitch, J. et al., (2015). *Parameter estimation for compact binaries with ground-based gravitational-wave observations using the lalInference software library.* *Phys. Rev. D*, 91(4), 042003. doi: 10.1103/PhysRevD.91.042003
- [Widrow, (1962)] Widrow, B., (1962). *Generalization and Information Storage in Networks of ADALINE "Neurons"*. En: Yovits, M. C., Jacobi, G. T. y Goldstein, G. D. (editores), *Self-Organizing Systems 1962*, Washington, EE. UU., Spartan.
- [Zeiler & Ferguson, (2013)] Zeiler, M.D. & Fergus, R., (2013). *Visualizing and Understanding Convolutional Networks.* En línea: <http://arxiv.org/abs/1311.2901>
- [Zhao & Li, (2017)] Zhao, N. & Li, Z., (2017). *Experiment and Artificial Neural Network Prediction of Thermal Conductivity and Viscosity for Alumina-Water Nanofluids.* *Materials*, 10(5), 552. doi: 10.3390/ma10050552
- [Zou & He, (1997)] Zou, Q. S. & He, X. Y., (1997). *On pressure and velocity boundary conditions for the lattice Boltzmann BGK model.* *Phys. Fluids*, 9 (6), 1591. doi: 10.1063/1.869307