

UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA

FACULTAD DE INGENIERÍA

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO



**DETECCIÓN, RECONSTRUCCIÓN DE SEÑAL Y ESTIMACIÓN DE
PARÁMETROS PARA ONDAS GRAVITACIONALES.**

POR:

ING. CARLOS PATRICIO CASTAÑEDA BUSTAMANTE

**TESIS PRESENTADA COMO REQUISITO PARA OBTENER EL
GRADO DE**

MAESTRO EN INGENIERÍA EN COMPUTACIÓN

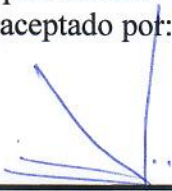
Director de Tesis: Dr. Alain Manzo Martínez

CHIHUAHUA, CHIH., MÉXICO

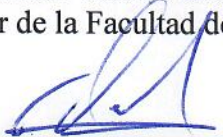
5 de junio de 2026



Detección, Reconstrucción de señal y Estimación de parámetros para ondas gravitacionales. Tesis presentado por Carlos Patricio Castañeda Bustamante como requisito parcial para obtener el grado de Maestro en Ingeniería en Computación, ha sido aprobado y aceptado por:



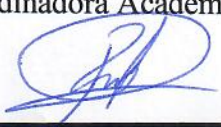
M.I. Fabián Vinicio Hernández Martínez
Director de la Facultad de Ingeniería



M.I. Rodrigo de la Garza Aguilar
Secretario de Investigación y Posgrado



M.S.I Karina Rocío Requena Yáñez
Coordinadora Académica



Dr. Alain Manzo Martínez
Director de Tesis

Agosto 2026

COMITÉ

Dr. Alain Manzo Martínez

Dra. Olanda Prieto Ordaz

Dr. Fernando Martínez Reyes

Dra. Graciela María de Jesús Ramírez Alonso



UNIVERSIDAD AUTÓNOMA DE
CHIHUAHUA

13 de agosto de 2026.

ING. CARLOS PATRICIO CASTAÑEDA BUSTAMANTE
Presente. -

En atención a su solicitud relativa al trabajo de tesis para obtener el grado de Maestro en Ingeniería en Computación, nos es grato transcribirle el tema aprobado por esta Dirección, propuesto y dirigido por el director **Dr. Alain Manzo Martínez** para que lo desarrolle como tesis, con el título **"DETECCIÓN, RECONSTRUCCIÓN DE SEÑAL Y ESTIMACIÓN DE PARÁMETROS PARA ONDAS GRAVITACIONALES"**.

Índice de Contenido

Resumen

Introducción

Hipótesis

Fundamentación teórica

Ondas Gravitacionales

Sistemas binarios coalescentes compactos

Simulación de ondas gravitacionales (relatividad numérica)

Transformada de Wavelet y Transformada de Wavelet Discreta (DWT)

Filtros Digitales

Energía

Redes Neuronales

Optimización por nube de partículas

Metodología

Simulación de GWs y creación de la base de datos

Configuración final de la base de datos

Análisis de la señal

Optimización conjunta del tamaño de ventana mediante PSO



UNIVERSIDAD AUTÓNOMA DE
CHIHUAHUA

Modelos de machine learning

Resultados

- Simulación de las ondas
- Análisis preliminar para la creación de la base de datos
- Análisis de la señal
- Optimización conjunta del tamaño de ventana mediante PSO
- Clasificador LSTM
- Autoencoder
- Modelo CNN de regresión

Discusión

- Comparación con el estado del arte
- Base de datos
- Análisis preliminar
- Análisis de la señal
- Masa del chirp
- PSO
- Modelos de machine learning

Conclusión

ATENTAMENTE
"naturam subiecit aliis"

EL DIRECTOR

**M.I. FABIÁN VINICIO HERNÁNDEZ
MARTÍNEZ**

FACULTAD DE
INGENIERÍA
U.A.CH.



**SECRETARIO DE INVESTIGACIÓN
Y POSGRADO**

M.I. RODRIGO DE LA GARZA AGUILAR

DIRECCIÓN



Índice

Resumen	5
Introducción	6
Hipótesis	10
Fundamentación teórica	10
Ondas gravitacionales	12
Sistemas binarios coalescentes compactos	14
Masa del chirp	15
Simulación de ondas gravitacionales (relatividad numérica)	15
Generación de la onda: relatividad numérica	16
Simulación del ruido	18
PSD	18
Relación señal-ruido óptimo	19
Inyección de la onda	19
Parámetros	20
Transformada Wavelet y Transformada de Wavelet Discreta (DWT)	21
Filtros Digitales	25
Filtros FIR (Finite Impulse Response)	26
Energía	27
Redes Neuronales	27
Redes Neuronales Convolucionales (CNN)	29
Long Short Term Memory (LSTM)	29
Autoencoders	31
Optimización por nube de partículas (PSO)	32
Metodología	36
Simulación de GWs y creación de la base de datos	36
Simulación	36

Base de datos	40
Configuración final de la base de datos	44
Transformada de Wavelet	44
Diferencia de energía	47
Análisis de la señal	51
Cadena de filtrado FIR y extracción de características para detección	51
Optimización conjunta del tamaño de ventana mediante PSO	53
Modelos de machine learning	57
Red LSTM para clasificación binaria de ondas gravitacionales	57
Entrenamiento del clasificador LSTM	58
Autoencoder	60
Entrenamiento del autoencoder	63
Modelo CNN para la regresión de parámetros	65
Entrenamiento del modelo CNN	67
Resultados	68
Simulación de las ondas	68
Análisis preliminar para la creación de la base de datos	72
Transformada de Wavelet	72
Diferencia de energía	74
Análisis de la señal	75
Cadena de filtrado FIR	75
Extracción de características para detección	79
Optimización conjunta del tamaño de ventana mediante PSO	84
Clasificador LSTM	86
Autoencoder	92
Modelo CNN de regresión	103
Discusión	107
Comparación con el estado del arte	107
Base de datos	108

Análisis preliminar	109
Análisis de la señal	110
Masa del chirp	112
PSO	112
Modelos de machine learning	113
Conclusión	114
Curriculum Vitae	123



Resumen

En este estudio se exploró una metodología capaz de detectar una onda gravitacional saturada por ruido gaussiano, mediante el uso de algoritmos computacionales. Para este fin se construyó una base de datos con señales sintéticas simulando el comportamiento de la coalescencia de sistemas de cuerpos compactos binarios mediante el aproximante *IMRPhenomXPHM*, implementado a través de la librería *Bilby*, con una frecuencia de muestreo de 4096 Hz , duraciones de 8 segundos y razones de señal a ruido (SNR) en el rango de 10 a 50 dB , empleando el modelo de ruido *aLIGOZeroDetHighPower* como fondo realista.

El procesamiento de la señal consiste en una cadena de filtros FIR de ventana Hann, de 1500 y 1000 coeficientes con frecuencias de corte de 1000 Hz y 550 Hz , respectivamente, y una descomposición wavelet discreta de 10 niveles con una wavelet madre de *Haar*, de la cual se extraen los coeficientes de detalle del nivel 9. A esta señal resultante se le calcula la energía con una ventana deslizante w_s . El perfil energético resultante se segmenta en W ventanas, a las cuales se les calcula la desviación estándar. El valor óptimo de w_s y W se determina mediante optimización por enjambre de partículas (PSO, 20 partículas, 500 iteraciones), maximizando la desviación estándar de la energía por ventana, resaltando la región donde se concentra el evento gravitacional. Este vector de W desviaciones estándar alimenta un clasificador LSTM de dos capas apiladas (64 unidades ocultas) que determina si la señal contiene una onda gravitacional o corresponde únicamente a ruido.

Las señales clasificadas como positivas se procesan mediante un modelo de autoencoder, de tipo U-Net con conexiones de salto entre cuatro niveles de codificación y decodificación, entrenado mediante aprendizaje por currículo basado en el SNR, para reconstruir la forma de onda limpia correspondiente a la fuente con ruido gaussiano. A partir de esta señal reconstruida, un regresor CNN de cuatro capas convolucionales estima la masa del chirp de la fuente gravitacional.



Introducción

Las ondas gravitacionales han revolucionado la astronomía al permitir el estudio de eventos cósmicos invisibles para los telescopios tradicionales, como fusiones de agujeros negros o estrellas de neutrones. A diferencia de la luz, estas ondas atraviesan el universo sin distorsión, transportando información única sobre fenómenos extremos. Sin embargo, su detección y análisis representan un desafío tecnológico y computacional sin precedentes, donde técnicas de Machine Learning (ML) y Deep Learning (DL) emergen como herramientas clave.

Aunque ya se había confirmado indirectamente la existencia de las ondas gravitacionales, estas nunca habían sido detectadas de manera directa. A partir de 1984, los físicos Kip Thorne, Reiner Weiss y Barry Barish, con el apoyo del MIT, Caltech y la Fundación Nacional de Ciencia (NSF), comenzaron a sentar las bases y a desarrollar el Observatorio de Ondas Gravitacionales con Interferómetro Láser, conocido como LIGO (Laser Interferometer Gravitational-Wave Observatory). LIGO utiliza dos interferómetros gigantes ubicados en Hanford, Washington, y Livingston, Louisiana, que miden la diferencia entre dos haces de láser para generar un patrón de interferencia. Weiss planteó que, si una onda gravitacional distorsionaba el espacio-tiempo, esta distorsión podría detectarse en el patrón de interferencia, lo que permitiría obtener una detección directa de las ondas gravitacionales (Weiss , 2022).

LIGO llevó a cabo varias actualizaciones para aumentar su sensibilidad en la detección de ondas gravitacionales y reducir los efectos del ruido sísmico, térmico e incluso cuántico. Gracias a estas mejoras, el 14 de septiembre de 2015, LIGO logró detectar de manera directa ondas gravitacionales generadas por la fusión de dos agujeros negros con masas de aproximadamente 30 veces la del Sol ($M_{\odot}$), ubicados a una distancia de 1.300 millones de años luz de la Tierra (Abbott et al., 2016). Así, con la publicación de sus resultados en 2016 (Abbott et al., 2016), se confirmó de manera contundente, y un siglo después de su predicción, la existencia de las ondas gravitacionales. Este hito marcó un antes y un después en la física y la astronomía, validando una de las predicciones más importantes de la teoría de la relatividad general de Einstein. En 2017, Kip Thorne, Reiner Weiss y Barry Barish recibieron el Premio Nobel de Física por esta detección.



Desde 2015 hasta 2021, LIGO ha reportado la detección de 90 eventos diferentes que generan ondas gravitacionales (LIGO Lab , n.d.), y se espera que, a medida que se mejore la sensibilidad de sus detectores, se puedan registrar aproximadamente un evento por semana. El proceso de detección de estas ondas implica identificar señales extremadamente débiles dentro de los datos recopilados por los instrumentos. Tradicionalmente, este proceso ha dependido de análisis manuales y algoritmos especializados diseñados para filtrar el ruido y extraer las señales de interés (Abbott et al., 2016).

El principal obstáculo en este campo radica en la naturaleza misma de las señales: las ondas gravitacionales presentan amplitudes extraordinariamente bajas, frecuentemente enmascaradas por diversas fuentes de ruido, tanto instrumentales como ambientales. Esta característica impone requisitos sin precedentes en los métodos de procesamiento y análisis de datos.

La mejora continua en la sensibilidad de los detectores genera un crecimiento exponencial en la complejidad computacional del problema. Este escenario plantea una necesidad urgente: el desarrollo de algoritmos innovadores que puedan mantener, y superar, los actuales estándares de precisión mientras optimizan la velocidad de procesamiento. El futuro de la astronomía de ondas gravitacionales dependerá críticamente de nuestra capacidad para implementar modelos de análisis más eficientes, posiblemente mediante técnicas de machine learning.

Este desafío ha sido abordado mediante diversos enfoques innovadores. Un avance significativo se presentó en 2020 con el trabajo de Vajente et al. (Vajente et al. , 2020), quienes en su artículo '*Machine-learning nonstationary noise out of gravitational-wave detectors*' propusieron un modelo de aprendizaje automático diseñado específicamente para caracterizar y eliminar el ruido de fondo en los detectores. Su enfoque demostró resultados prometedores, logrando una mejora sustancial en la sensibilidad del detector al filtrar eficazmente las interferencias no estacionarias que dificultan la identificación de señales gravitacionales genuinas. Este desarrollo representa un paso crucial hacia la optimización de los sistemas de detección, combinando técnicas avanzadas de inteligencia artificial con la física de ondas gravitacionales.



En 2017, George y Huerta publicaron un artículo titulado '*Deep Neural Networks to Enable Real-time Multimessenger Astrophysics*' (George & Huerta , 2018b). En este estudio, los autores desarrollaron un modelo basado en Deep Learning con dos objetivos principales: clasificar eventos astrofísicos cataclísmicos y estimar sus parámetros fundamentales, especialmente las masas de los objetos implicados. El modelo estaba diseñado para extraer esta información incluso en señales con alto nivel de ruido. Para su entrenamiento y validación, emplearon conjuntos de datos simulados que reproducían fielmente las condiciones reales de observación.

Posteriormente, en su trabajo '*Deep Learning for Real-time Gravitational Wave Detection and Parameter Estimation: Results with Advanced LIGO Data*' (George & Huerta , 2018a), expandieron esta idea al aplicar, por primera vez, redes neuronales convolucionales (CNNs) a datos reales de ondas gravitacionales. Demostraron que su enfoque era capaz de detectar señales y estimar parámetros incluso en condiciones de baja calidad de datos, sentando así un precedente para el desarrollo de algoritmos futuros que pudieran extraer características más complejas.

El avance en los algoritmos y el creciente papel del machine learning en este campo ha alcanzado tal relevancia que, en 2024, Marx et al. presentaron en su artículo '*A Machine-Learning Pipeline for Real-Time Detection of Gravitational Waves from Compact Binary Coalescences*' (Marx et al. , 2024) un modelo innovador para la detección en tiempo real de ondas gravitacionales producidas por coalescencia de agujeros negros binarios. Aunque la integración de redes neuronales en este ámbito no está exenta de desafíos, el equipo logró desarrollar un modelo capaz de identificar estas señales de manera eficiente. Este avance ha llevado a la colaboración LIGO a adoptar activamente estos nuevos algoritmos, consolidando así el machine learning como una herramienta fundamental en la astronomía de ondas gravitacionales.

Estos avances delinean un camino claro para el campo: el desarrollo de algoritmos de machine learning capaces de detectar ondas gravitacionales con eficiencia computacional y precisión se ha vuelto indispensable. Para abordar este desafío, proponemos un enfoque integral que combina simulaciones realistas de formas de onda gravitacionales, técnicas de procesamiento de señales y



arquitecturas de redes neuronales optimizadas. El objetivo es desarrollar un modelo capaz de detectar ondas gravitacionales de manera eficiente, buscando así, contribuir al campo de estudio.

Es importante considerar que este trabajo se enfoca en un campo de estudio emergente, donde el acceso a datos observacionales públicos es aún escaso. Esta restricción obliga a depender de simulaciones computacionales para la construcción de bases de datos, las cuales, implican una demanda computacional significativa, especialmente en términos de procesamiento y almacenamiento. Dada la indisponibilidad de infraestructura computacional especializada para este proyecto, será necesario reducir los conjuntos de datos simulados para adaptarlos a los recursos disponibles.

La detección directa de ondas gravitacionales ha abierto una nueva ventana, hasta ahora inexplorada, para estudiar el comportamiento de nuestro universo. Este avance nos permite observar eventos cataclísmicos, como la coalescencia de agujeros negros o de estrellas de neutrones, fenómenos que, de otro modo, serían invisibles para nosotros. Sin embargo, su estudio presenta numerosos desafíos que aún deben superarse.

Como predijo Einstein en su formulación del cuadrupolo gravitacional (Einstein , 1916b), las ondas gravitacionales serían extremadamente débiles, hasta el punto de que él mismo dudó de que pudieran detectarse alguna vez o de que su estudio tuviera utilidad práctica. Este hecho plantea uno de los mayores desafíos en su investigación: las señales detectadas son tan débiles que el control del ruido instrumental (cuántico, térmico, sísmico, entre otros) (Vajente et al. , 2020) se convierte en un aspecto fundamental del proceso.

Si bien, en la actualidad, laboratorios como LIGO y Virgo emplean métodos avanzados de procesamiento de señales para detectar ondas gravitacionales (Abbott et al., 2020), estos algoritmos son computacionalmente costosos y no garantizan la ausencia de falsos negativos (Abbott et al., 2023).

Inspirados por los más recientes avances, este trabajo tiene como objetivo explorar la aplicación de diversas técnicas de procesamiento de señales y arquitecturas de redes neuronales en el análisis de ondas gravitacionales en sistemas binarios coalescentes, con el propósito de complementar los



métodos tradicionales, y aportar una perspectiva innovadora en el estudio de estos fenómenos para su detección y clasificación, en entornos de alta relación señal ruido (SNR).

Hipótesis

La selección adecuada de métodos de procesamiento de señales aplicados a datos simulados de ondas gravitacionales, proporcionará las características necesarias para que un algoritmo pueda alcanzar una precisión del 90 % en las tareas de detección, reconstrucción y regresión de ondas gravitacionales.

Fundamentación teórica

En 1905, Henri Poincaré publicó un artículo titulado '*Sobre la dinámica del electrón*' (*Sur la dynamique de l'électron*), en el que buscaba describir el comportamiento de un electrón bajo la teoría de la relatividad y el electromagnetismo utilizando las transformaciones de Lorentz (Poincaré , 1905). En este trabajo, entre otras deducciones importantes, Poincaré sugirió que todas las fuerzas, independientemente de su origen, se ven afectadas por una traslación, al igual que las fuerzas electromagnéticas. Es decir, propuso que toda la dinámica podría ser relativa y que todos los componentes de estas fuerzas estarían sujetos a las transformaciones de Lorentz.

Poincaré también planteó que estas ideas podían extenderse a la dinámica de la gravedad, asumiendo, que la gravedad se comportaba de manera similar al electromagnetismo. De esta forma, Poincaré argumentó que la velocidad de propagación de la gravedad no era instantánea, sino que ocurría a la velocidad de la luz. Además, basándose en su suposición de que toda la dinámica podría ser relativa, Poincaré planteó que la posición y velocidad de un cuerpo atrayente determinarían la emisión de ondas gravitacionales (*ondes gravifiques*), mientras que la posición y velocidad del cuerpo atraído corresponderían al momento en que la onda gravitacional lo alcanzara. Aunque Poincaré no profundizó en detalles teóricos o matemáticos, su trabajo constituye la primera sugerencia formal de la posible existencia de ondas gravitacionales.



En ese mismo año, Einstein publicó su artículo '*Sobre la electrodinámica de los cuerpos en movimiento*' (*Zur Elektrodynamik bewegter Körper*), en el cual, utilizando las transformaciones de Lorentz, generalizó la dinámica de los objetos, especialmente a velocidades cercanas a la de la luz (Einstein , 1905), y exploró la relación entre el espacio y el tiempo, bajo la suposición de que esta dimensión es plana. Diez años más tarde, Einstein publicó '*La base de la teoría general de la relatividad*' (*Die Grundlage der allgemeinen Relativitätstheorie*), donde propuso que la gravedad es una propiedad geométrica del espacio-tiempo (Einstein , 1916a). En esta teoría, el espacio-tiempo no es plano (caso que pasó a ser conocido como relatividad especial), sino que presenta curvaturas causadas por la presencia de masas, y la gravedad es un fenómeno resultante de estas deformaciones.

Posteriormente, Einstein dedicó esfuerzos a calcular la radiación gravitacional (Einstein , 1916b), un proceso que requirió múltiples intentos y lo llevó incluso a dudar de la existencia de las ondas gravitacionales. No obstante, finalmente logró derivar la fórmula cuadrupolar (Einstein & Rosen , 1937), una expresión que se construyó como una analogía directa con la radiación electromagnética (Flanagan & Hughes , 2005). Esta fórmula demostró que las ondas gravitacionales se originan a partir de masas en aceleración, de forma análoga a cómo las ondas electromagnéticas son generadas por cargas aceleradas. De esta manera, se sentó una base teórica robusta que predice la existencia de las ondas gravitacionales. La fórmula cuadrupolar de Einstein también anticipó que estas ondas serían extremadamente débiles y, por lo tanto, muy difíciles de detectar.

De este modo, Einstein estableció los fundamentos teóricos que respaldan la existencia de las ondas gravitacionales. El siguiente reto consistiría en demostrarlas experimentalmente. Desde que Einstein predijo su existencia, los físicos han explorado diversas formas de verificar este fenómeno.

No fue hasta 1974 cuando Russell Hulse y Joseph Taylor realizaron un descubrimiento trascendental: una estrella de neutrones que rotaba sobre su eje a una velocidad de 17 veces por segundo (Hulse & Taylor , 1975). Este hallazgo indicaba que habían detectado un púlsar, un tipo de estrella de neutrones que emite radiación electromagnética a altas velocidades de rotación (Lovelace & Tyler , 2012). Hulse y Taylor observaron una variación repetitiva en la detección de los pulsos,



lo que les llevó a concluir que el púlsar formaba parte de un sistema binario, es decir, orbitaba alrededor de otra estrella, que resultó ser también una estrella de neutrones.

Seis años después, Hulse y Taylor sostuvieron que este sistema binario representaba una herramienta esencial para el estudio de la relatividad general, gracias a sus propiedades excepcionales. Las dos estrellas de neutrones, cada una con una masa superior a la de nuestro Sol, se aceleraban mutuamente a una velocidad de 400 km/s (Weisberg et al. , 1981). Estas características las convertían en candidatas perfectas para generar la energía necesaria que permitiría detectar el fenómeno de las ondas gravitacionales, ofreciendo así un laboratorio natural para validar las predicciones de la relatividad general.

Hulse y Taylor lograron medir la pérdida de energía en este sistema binario, y los resultados obtenidos coincidían con las predicciones de la relatividad general. Esto les permitió argumentar que tenían la primera evidencia sólida de la detección indirecta de ondas gravitacionales. En 1993 Hulse y Taylor fueron galardonados con el Premio Nobel de Física por este descubrimiento.

Ondas gravitacionales

Las ondas gravitacionales (GW) son perturbaciones en el tejido del espacio-tiempo, generadas por los fenómenos más violentos y energéticos del universo, estas ondas traen consigo información acerca de sus eventos de origen, así como pistas sobre la naturaleza misma de la gravedad (LIGO Lab , n.d.). En 1916, Albert Einstein predijo su existencia en el marco de su teoría de la relatividad general, argumentando que, debido a la naturaleza no instantánea de la radiación gravitacional, masas aceleradas a velocidades extremas podrían deformar el espacio-tiempo, produciendo oscilaciones que se propagarían en todas direcciones como ondas (Einstein , 1916a).

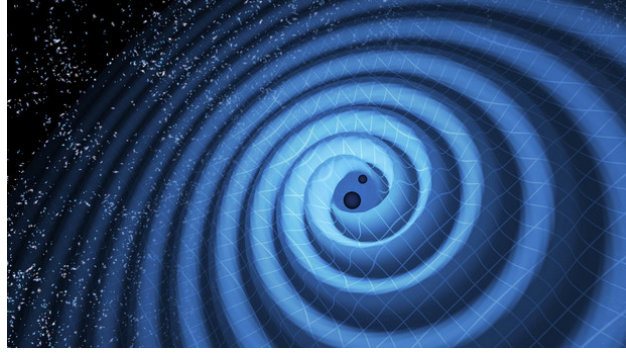


Figura 1: Ilustración de la fusión de dos agujeros negros, generando deformaciones en el espacio-tiempo. (LIGO Lab , n.d.)

Las ondas gravitacionales pueden representarse como perturbaciones $h_{\mu\nu}$ en un espacio-tiempo descrito por la métrica $g_{\mu\nu}$, que es solución de las ecuaciones de campo de Einstein (Antonelli , 2022):

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

En el límite de campo gravitacional débil, la métrica puede escribirse como una perturbación alrededor del espacio-tiempo plano de Minkowski $\eta_{\mu\nu}$:

$$g_{\mu\nu} \approx \eta_{\mu\nu} + h_{\mu\nu} + \mathcal{O}(h_{\mu\nu}^2) \quad (2)$$

En el vacío ($T_{\mu\nu} = 0$), las ecuaciones de Einstein se simplifican a una ecuación de onda:

$$\square h_{\mu\nu} = \left(\frac{\partial^2}{\partial t^2} - \nabla^2 \right) h_{\mu\nu} = 0 \quad (3)$$

Donde $\square$ es el operador d'Alembertiano. De esta manera, $h_{\mu\nu}$ representa la siguiente propagación:

$$h_{\mu\nu}(t, z) = h_+(t - \frac{z}{c})e_{\mu\nu}^+ + h_\times(t - \frac{z}{c})e_{\mu\nu}^\times \quad (4)$$

Donde z es la dirección de propagación y $h_+, h_\times$ son las amplitudes de los dos modos de polarización (plus y cross).

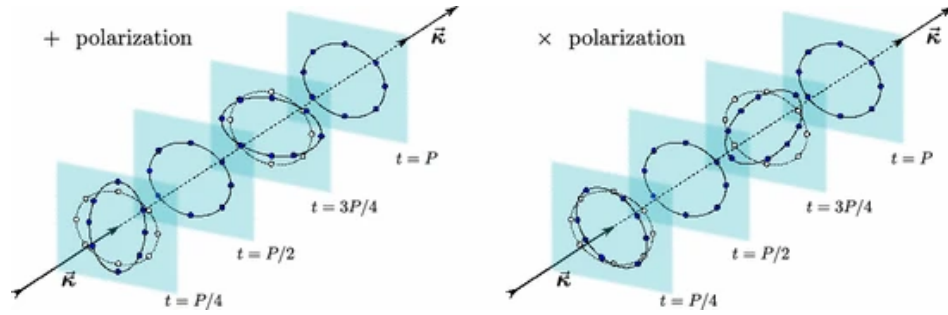


Figura 2: Propagación de perturbaciones $h_{\mu\nu}$ de onda en z con sus polarizaciones h_+ , $h_\times$. (Bishop & Rezzolla , 2016)

Para un análisis riguroso de las soluciones a la ecuación de onda en el contexto de la relatividad general, se recomienda consultar *Introducing Einstein's Relativity* (d'Inverno & Vickers , 2022) (Cap. 21 en la 2ª edición) y *A First Course in General Relativity* (Schutz , 2018) (Cap. 9 en la 2ª edición).

Sistemas binarios coalescentes compactos

Los sistemas binarios coalescentes compactos (CBC, por sus siglas en inglés) están compuestos por un par de objetos compactos que orbitan mutuamente. Estos sistemas pueden estar formados por dos agujeros negros (BBH), dos estrellas de neutrones (BNS) o una combinación de ambos (NSBH). Durante su movimiento orbital, emiten ondas gravitacionales débiles que provocan una pérdida gradual de energía en el sistema. Como consecuencia, la órbita se reduce progresivamente (Spera et al. , 2022).

A medida que este proceso avanza, la disminución de la órbita hace que los objetos incrementen su velocidad, generando ondas gravitacionales más intensas. Esto, a su vez, acelera aún más la contracción de la órbita en un proceso conocido como fase de *inspiral*. Posteriormente, tiene lugar la fusión (*merger*), momento en el que los objetos colisionan y se produce la emisión de las ondas gravitacionales más potentes. Finalmente, el sistema entra en la etapa de *ringdown*, donde se liberan los remanentes de la distorsión espacio-temporal antes de estabilizarse.

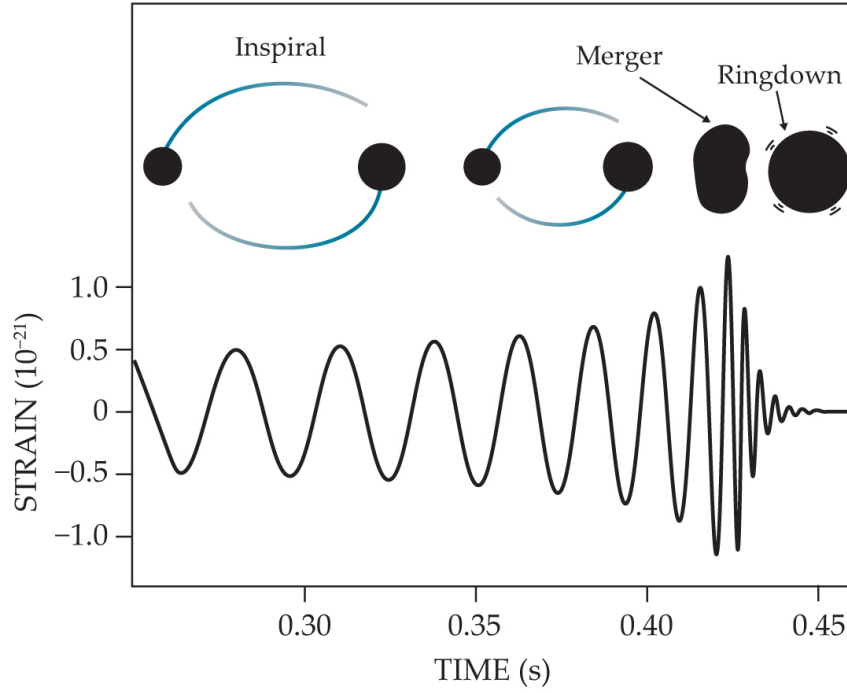


Figura 3: Esquema básico sobre los CBC. (Berti et al. , 2025)

Masa del chirp

A orden Newtoniano, la onda gravitacional generada depende únicamente de la combinación de las masas de los objetos que la generan. Esta masa se denomina masa del chirp $\mathcal{M}_c$ (Cutler & Flanagan , 1994).

$$\mathcal{M}_c = \frac{(m_1 m_2)^{3/5}}{(m_1 + m_2)^{1/5}} \quad (5)$$

Cutler & Flanagan (1994) mostraron que uno puede medir y predecir la masa del chirp con un error de $\frac{\Delta \mathcal{M}_c}{\mathcal{M}_c} \approx 10^{-3}$ demostrando, que la mayoría de la información contenida en la onda gravitacional es la masa del chirp (Cutler & Flanagan , 1994).

Simulación de ondas gravitacionales (relatividad numérica)

Simular una GW como si fuera observada por un interferómetro es un proceso compuesto por tres etapas:



1. Generación de la onda: Generar una GW sin ruido, $h(t)$, utilizando modelos de relatividad numérica como el IMRPhenomP.
2. Simulación de ruido: Simular el ruido del interferómetro, $n(t)$, como un proceso estacionario de ruido gaussiano caracterizado por su densidad espectral (PSD).
3. Inyección de la onda: Combinar la onda y el ruido para crear la señal observada, $d(t) = h(t) + n(t)$.

La onda generada modela la evolución del sistema, regresando una serie de tiempo representativa del evento. Por ejemplo, el modelo IMRPhenomP, definido en parte por las ecuaciones (6) y (7), describe el evento completo de inspiral, merger y ringdown de un par de agujeros negros.

Generación de la onda: relatividad numérica

La relatividad numérica es una rama interdisciplinaria que combina la relatividad general con la física computacional, empleando métodos numéricos para resolver las ecuaciones de campo de Einstein. Uno de los focos de investigación actual en este campo es la simulación de sistemas binarios compactos (CBCs) y las ondas gravitacionales que emiten.

Para modelar estos fenómenos, se utilizan diversas aproximaciones, como, la familia Effective One Body (EOB), que incluye modelos como SEOBNv2 y SEOBNv2T (Damour , 2008) o las aproximaciones de la serie IMRPhenomP, como IMRPhenomPv2 e IMRPhenomXPHM (Khan et al. , 2016). Cada uno de estos modelos representa un enfoque distinto para abordar numéricamente las complejas ecuaciones de Einstein, equilibrando precisión y eficiencia computacional.

Para ejemplificar de manera general, el funcionamiento de estas aproximaciones, tomamos como ejemplo la serie IMRPhenomP. Esta serie descompone la señal gravitacional en fases (inspiral, intermedia y merger-ringdown), modeladas mediante funciones de transición suaves. Para una frecuencia f , la fase Φ y una amplitud A se expresan como:

$$\Phi_{\text{IMR}}(f) = \phi_{\text{Ins}}(f)\theta_{f_1}^- + \theta_{f_1}^+\phi_{\text{Int}}(f)\theta_{f_2}^- + \theta_{f_2}^+\phi_{\text{MR}}(f), \quad (6)$$



$$A_{\text{IMR}}(f) = A_{\text{Ins}}(f)\theta_{f_1}^- + \theta_{f_1}^+ A_{\text{Int}}(f)\theta_{f_2}^- + \theta_{f_2}^+ A_{\text{MR}}(f), \quad (7)$$

donde:

- $\phi_{\text{Ins}}, \phi_{\text{Int}}$ y ϕ_{MR} son las fases de inspiral, intermedia y merger-ringdown.
- $A_{\text{Ins}}, A_{\text{Int}}$ y A_{MR} son las amplitudes correspondientes.
- $\theta_{f_i}^\pm$ son funciones de transición que garantizan continuidad entre fases.

Donde $\phi_{\text{Ins}}, \phi_{\text{Int}}$ y ϕ_{MR} se describen como:

$$\phi_{\text{Ins}} = \phi_{\text{TF2}}(Mf; \Xi) + \frac{1}{\eta}(\sigma_0 + \sigma_1 f + \frac{3}{4}\sigma_2 f^{4/3} + \frac{3}{5}\sigma_3 f^{5/3} + \frac{1}{2}\sigma_4 f^2) \quad (8)$$

$$\phi_{\text{Int}} = \frac{1}{\eta}(\beta_0 + \beta_1 f + \beta_2 \text{Log}(f) - \frac{\beta_3}{3} f^{-3}) \quad (9)$$

$$\phi_{\text{MR}} = \frac{1}{\eta} \left\{ \alpha_0 + \alpha_1 f - \alpha_2 f^{-1} + \frac{4}{3} \alpha_3 f^{3/4} + \alpha_4 \tan^{-1} \left(\frac{f - \alpha_5 f_{\text{RD}}}{f_{\text{damp}}} \right) \right\} \quad (10)$$

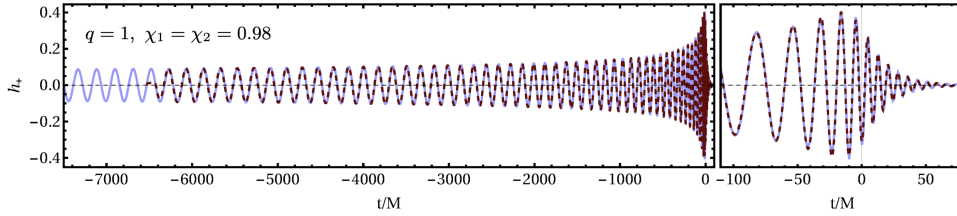


Figura 4: Ejemplo de una onda gravitacional de la serie IMRPhenomP (Khan et al. , 2016).

Para un análisis detallado, se puede consultar el artículo '*Frequency-domain gravitational waves from non-precessing black-hole binaries. II. A phenomenological model for the advanced detector era*' (Khan et al. , 2016).



Simulación del ruido

PSD

La simulación del ruido en un interferómetro requiere una caracterización espectral precisa, la cual se obtiene a partir de su Densidad Espectral de Potencia (PSD). La PSD es una función que describe la distribución de la potencia promedio de una señal en el dominio de la frecuencia (Madisetti, 2010). Para un proceso estocástico estacionario, la PSD teórica, $S_{gg}(f)$, se define como la transformada de Fourier de su función de autocorrelación $R_{gg}(\tau)$.

En la práctica, la PSD debe estimarse a partir de un conjunto finito de observaciones discretas del proceso, $\{g[n]\}$ para $n = 0, 1, \dots, N-1$, donde $g[n] = g[n+N]$. Un estimador fundamental de la PSD es el periodograma, el cual se calcula de la siguiente manera (Madisetti, 2010):

Si asumimos una señal $\{g[n]\}$ con periodo N , es decir

$$g[n] = g[n+N] \quad (11)$$

Utilizamos la transformada discreta de Fourier (DFT) para representar la señal en el dominio de la frecuencia

$$G(f_k) = \sum_{n=0}^{N-1} g[n]e^{-j2\pi f_k n}, \quad f_k = \frac{k}{N}, \quad k \in \{0, 1, \dots, N-1\} \quad (12)$$

El periodograma, denotado como $\hat{S}_{gg}(f_k)$ se define entonces como la magnitud al cuadrado de la DFT, normalizada por la longitud de la secuencia

$$\hat{S}_{gg}(f_k) = \frac{1}{N} |G(f_k)|^2 \quad (13)$$

Esto para un sistema discreto, para un sistema continuo se utiliza la transformada de Fourier, es decir

$$\tilde{n}_T = \int_{-\infty}^{\infty} g(t)e^{-j2\pi f_k t} dt \quad (14)$$



$$S_{xx}(f_k) = \lim_{T \rightarrow \infty} \frac{1}{T} |\tilde{n}_T(f)|^2 \quad (15)$$

Relación señal-ruido óptimo

Consideremos el flujo de datos $d(t)$ que contiene una señal GW, de modo que $d = s + n$, donde s es la señal limpia y n es el ruido. Ahora, consideremos una plantilla de onda, $h(\theta)$ parametrizada por un conjunto de parámetros intrínsecos y extrínsecos θ (por ejemplo, masas m_1, m_2 , espines χ_1, χ_2 , etc.).

La detectabilidad y la distinguibilidad de una señal se cuantifican mediante su relación señal-ruido (SNR). La SNR esperada, u *óptima*, para una plantilla $h(\theta)$ se define como la raíz cuadrada del producto interno de la plantilla consigo misma:

$$\rho_{\text{opt}}^2 = \langle h|h \rangle, \quad (16)$$

donde el producto interno $\langle a|b \rangle$ tiene en cuenta la sensibilidad al ruido del detector a través de su PSD:

$$\langle a|b \rangle = 4, \text{Re} \int_{f_{\min}}^{f_{\max}} \frac{\tilde{a}(f) \tilde{b}^*(f)}{\tilde{S}_{xx}(f)} df. \quad (17)$$

La SNR óptima representa el límite superior teórico para detectar una señal h en presencia de ruido caracterizado por $S_{xx}(f)$.

Inyección de la onda

Una vez determinada la relación señal-ruido (SNR) óptima, se procede a inyectar la forma de onda teórica, $h(t)$, en el ruido simulado, $n(t)$, para generar la señal observada $d(t) = h(t) + n(t)$. Este proceso se ilustra en la Figura 5, donde pueden observarse tanto el ruido calculado según la SNR óptima como la señal limpia $h(t)$ inmersa en dicho ruido.

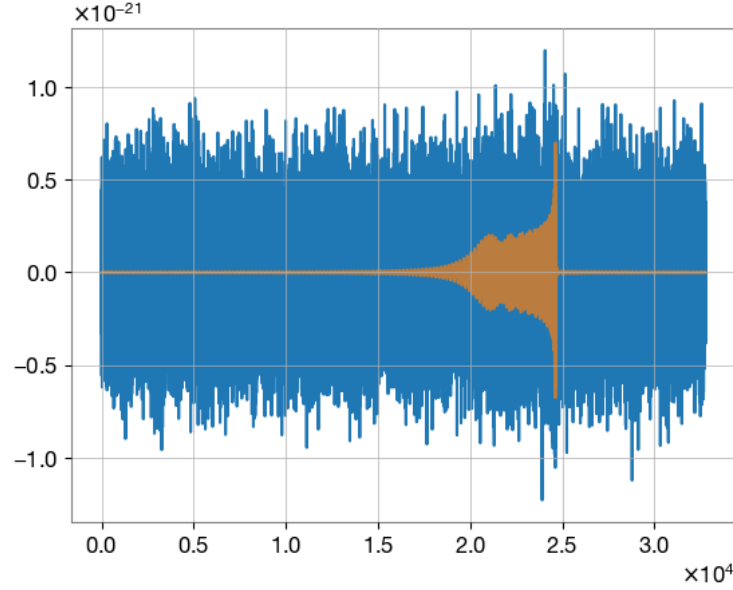


Figura 5: Serie de tiempo para el componente de ruido y el componente limpio de una señal GW procedente de una fusión binaria de agujeros negros, con una relación señal-ruido óptima de 15,6 db.

Parámetros

La noción habitual de *fuerza gravitatoria* desaparece en la relatividad general, sustituida por la idea de que los cuerpos en caída libre siguen geodésicas en el espacio-tiempo. Dada una métrica del espacio-tiempo g_{ab} y un conjunto de coordenadas del espacio-tiempo x^a , las trayectorias geodésicas vienen dadas por la ecuación (Schutz , 2018):

$$\frac{d^2 x^a}{d\tau^2} + \Gamma_{bc}^a \frac{dx^b}{d\tau} \frac{dx^c}{d\tau} = 0 \quad (18)$$

donde τ es el tiempo propio medido por un observador que viaja a lo largo de la geodésica.

Para que un interferómetro GW detecte ondas gravitacionales, es necesario evaluar (18) en el marco de referencia propio local del interferómetro. Esto se puede lograr utilizando un sistema de coordenadas ortogonal a la onda, lo que se denomina *calibre transversal*.

$$\partial_i h_{ij} = 0 \quad (19)$$

Esta es la representación de la propagación de la onda y su equivalente a (4), lo que significa que



podemos representar la polarización de la onda como

$$h_{xx}^{TT} = -h_{yy}^{TT} = h_{+}(t - z) \quad (20)$$

$$h_{xy}^{TT} = h_{yx}^{TT} = h_{\times}(t - z) \quad (21)$$

Por lo tanto, podemos reescribir la ecuación (18) como

$$\delta L^i(t) = \frac{1}{2} h_{ij}^{TT}(t) L_0^j \quad (22)$$

Donde L es la distancia entre el interferómetro y la fuente. La señal que recibe el interferómetro también viene dada por la posición de su fuente en el cielo, θ y ϕ , por lo tanto, el conocer estos parámetros es necesario para simular la señal observada.

Transformada Wavelet y Transformada de Wavelet Discreta (DWT)

Cuando se trabaja con señales, lo más común es hacerlo con señales en el dominio del tiempo (*time-domain*), las cuales representan la información en términos de amplitud en función del tiempo. Esta representación proporciona los datos 'en bruto' de la señal, pero a menudo es necesario procesarla para extraer información más detallada.

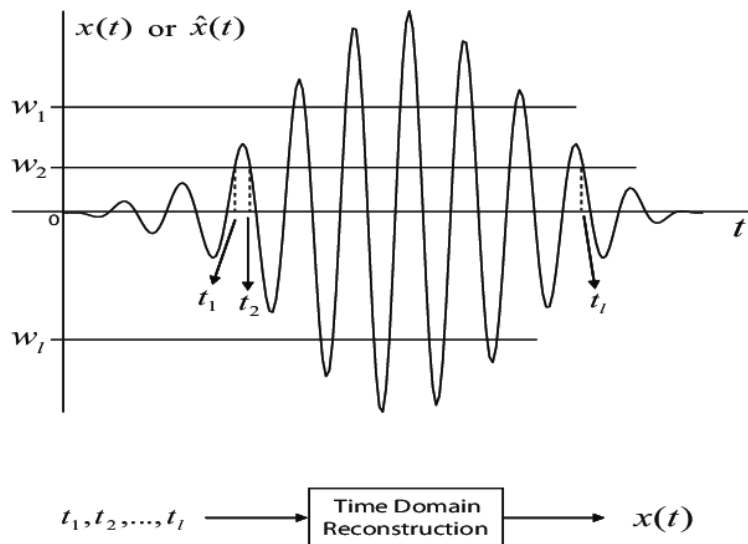


Figura 6: Esquema básico de una señal en el dominio del tiempo. (Kumaresan & Panchal , 2010)



Para analizar estas señales, se aplican transformaciones matemáticas que permiten estudiar sus características en otros dominios (como la frecuencia). Una señal que ha sido sometida a este tipo de transformación se denomina señal procesada (Polikar , 1996), y su análisis facilita la identificación de patrones, frecuencias dominantes y otras propiedades que no son evidentes cuando se analiza la señal original en el dominio del tiempo.

Entre las herramientas más comunes destacan las transformadas de Fourier (FT), que permiten descomponer una señal en sus componentes frecuenciales. Sin embargo, una limitación clave de la FT es que proporciona información sobre las frecuencias presentes en la señal, pero no indica cuándo ocurren temporalmente. Esta característica hace que la FT sea inadecuada para analizar señales no estacionarias (Polikar , 1996).

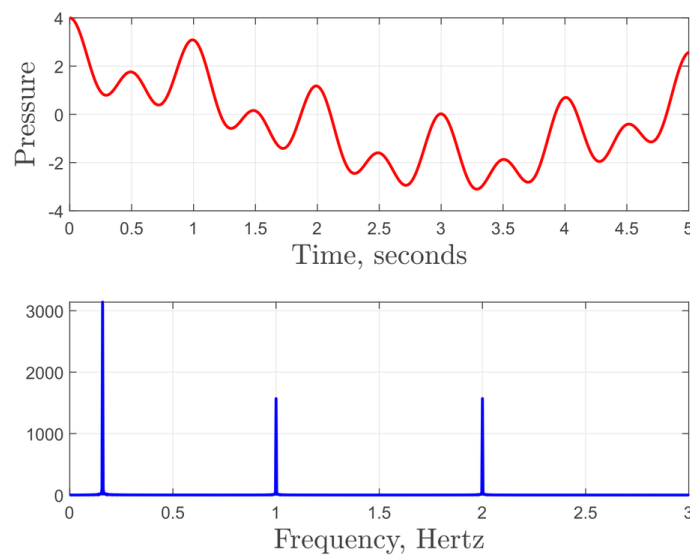


Figura 7: Transformada de Fourier de una función f . (White , 2020)

Una solución a este problema es la transformada de Fourier de tiempo corto (STFT). Su principio básico consiste en calcular la FT en segmentos o ventanas lo suficientemente pequeñas como para asumir que la señal es estacionaria dentro de ellas. Al desplazar esta ventana a lo largo de la señal, se obtiene un análisis tiempo-frecuencia (TFR) que revela cómo varían las componentes espectrales en el tiempo (Polikar , 1996).

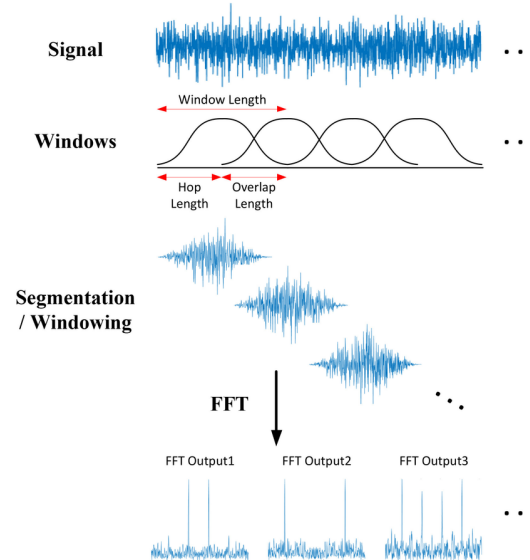


Figura 8: Ejemplo de una STFT. (Jeon et al. , 2020)

La STFT también presenta limitaciones inherentes, esta vez de origen físico. El principio de incertidumbre, formulado por Heisenberg en 1927, establece que no es posible determinar simultáneamente la posición y el momento de una partícula. Trasladado al análisis de señales, esto implica que no podemos conocer con precisión la frecuencia de una señal en un instante exacto de tiempo; en su lugar, solo podemos aproximar qué frecuencias están presentes dentro de un intervalo temporal determinado (Polikar , 1996).

Como consecuencia, la STFT debe equilibrar dos aspectos: si se emplea una ventana más grande, se obtiene una mejor resolución en frecuencia, pero se pierde resolución temporal. Por el contrario, con una ventana más pequeña, se gana resolución en el tiempo, pero se reduce la precisión en frecuencia. Es en este contexto donde surge la transformada wavelet como alternativa.

A diferencia de la STFT, que emplea ventanas de tamaño fijo, lo que la hace vulnerable al principio de incertidumbre, la transformada wavelet utiliza ventanas de resolución variable. Esta flexibilidad le permite adaptarse a las variaciones de la señal, ofreciendo una representación más precisa para los distintos componentes frecuenciales presentes (Polikar , 1996).

La transformada de Wavelet continua (CWT) se define como la suma del producto de una señal



$x(t)$ continua y una 'wavelet madre' $\Psi(t)$:

$$CWT_x^\Psi(a, b) = \Psi_x^\Psi(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \Psi\left(\frac{t-b}{a}\right) dt \quad (23)$$

Para que la wavelet madre sea válida debe de cumplir con la condición de admisibilidad (Valens , 1999):

$$C_\Psi = \int_{\mathbb{R}} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < \infty \quad (24)$$

Donde $\Psi(\omega)$ es la transformada de Fourier de $\psi(x)$. Esta condición implica que $\Psi(\omega)$ se desvanece en la frecuencia cero (Valens , 1999):

$$|\Psi(\omega)|^2 \Big|_{\omega=0} = 0 \quad (25)$$

$$0 = \Psi(0) = \int \psi(t) dt \quad (26)$$

En otras palabras, la wavelet madre $\Psi(t)$ debe de ser una onda (Valens , 1999) (Polikar , 1996).

La wavelet madre recibe este nombre porque sirve como función prototipo a partir de la cual se generan todas las transformaciones mediante los parámetros de escala (a) y desplazamiento (b). El parámetro a controla el escalamiento (dilatación/compresión) de la función, mientras que b determina su traslación en el tiempo (Polikar , 1996).

Para su implementación computacional, la transformada wavelet debe discretizarse. Esto se logra muestreando los parámetros a y b en valores discretos (UNAM , n.d.):

$$a = 2^{-j}$$

$$b = k2^{-j}$$

Donde k y j deben ser valores enteros.



De esta manera la transformada de wavelet discreta (DWT) se puede definir como:

$$DWT_{j,k} = 2^{j/2} \int_{-\infty}^{\infty} x(t) \psi(2^j t - k) dt \quad (27)$$

Finalmente, para la reconstrucción de la señal se utiliza la siguiente ecuación (UNAM , n.d.):

$$x(t) = \sum_k \sum_j c_{j,k} \phi(t) + \sum_k \sum_j d_{j,k} \psi(t) \quad (28)$$

Donde $\phi(t)$ es la función escala, $\psi(t)$ es la función wavelet, $c_{j,k}$ son los coeficientes de escala y $d_{j,k}$ son los coeficientes de wavelet.

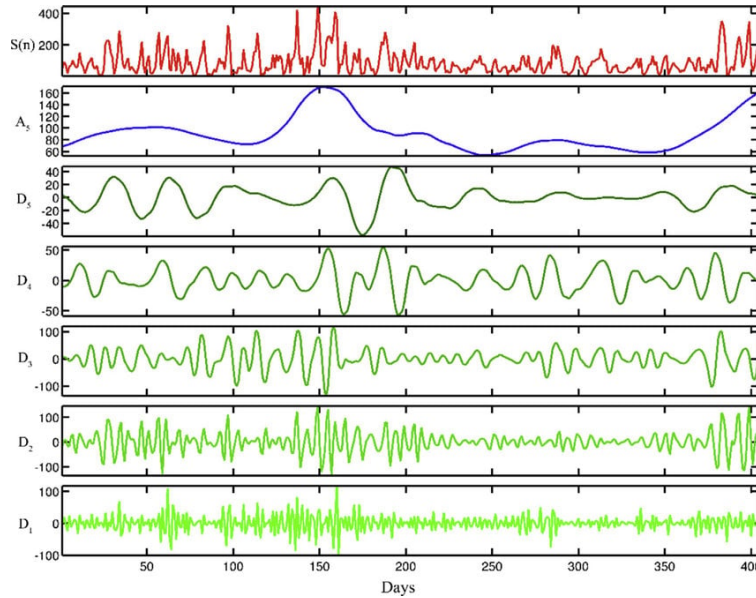


Figura 9: Descomposición de wavelet de una serie de tiempo en distintos coeficientes. (Feng et al. , 2015)

Filtros Digitales

Los filtros digitales tienen dos usos principales: La separación de señales y la reconstrucción de señales que han sido modificadas de alguna manera (Smith , 1997). La separación de señales es necesaria cuando una señal se ha contaminado por interferencia, ruido, alguna otra señal, etc.

La manera directa de implementar un filtro digital es mediante la convolución de la señal de entrada con la respuesta impulso del filtro. El filtro recibe una señal $x[n]$ de entrada y produce una señal



$y[n]$ de salida, donde n es el tiempo y $y[n]$ es la convolución de $x[n]$ con la respuesta impulso del filtro $h[n]$ (Smith , 1997).

Filtros FIR (Finite Impulse Response)

Los filtros FIR (Respuesta de Impulso Finita) son un tipo de filtros digitales cuya respuesta impulso $h[n]$ tiene duración finita, es decir, es distinta de cero únicamente para un número limitado de muestras (Smith , 1997). Para un filtro FIR de orden M , la respuesta impulso satisface:

$$h[n] = 0 \quad \text{para } n < 0 \text{ y } n > M \quad (29)$$

La salida $y[n]$ se calcula mediante la convolución discreta de la señal de entrada $x[n]$ con la respuesta impulso $h[n]$:

$$y[n] = \sum_{k=0}^M h[k]x[n-k] \quad (30)$$

donde $M + 1$ es la longitud del filtro (número de coeficientes). Esta ecuación muestra que la salida en el instante n depende únicamente de los valores actuales y pasados de la entrada, sin retroalimentación.

Una propiedad fundamental de los filtros FIR es que son incondicionalmente estables, ya que, al no existir retroalimentación, una entrada acotada produce siempre una salida acotada (criterio BIBO) (Ambardar , 1999). Además, es posible diseñar filtros FIR con fase lineal, lo que garantiza grupo constante en todas las frecuencias y preserva la forma de onda de la señal filtrada. Estas características los hacen especialmente adecuados para aplicaciones de procesamiento de señales donde la distorsión de fase es un factor crítico (Smith , 1997).

En la práctica, los filtros FIR se clasifican según su banda de paso: pasa-bajos, pasa-altos, pasa-banda y rechaza-banda. El diseño de estos filtros consiste en determinar los coeficientes $h[k]$ que aproximen una respuesta en frecuencia deseada, tarea que puede abordarse mediante técnicas como el método de ventanas o la aproximación de mínimos cuadrados (Smith , 1997).



$$h[k] = B_0\delta[k] + B_1\delta[k-1] + B_2\delta[k-2] + \dots + B_M\delta[k-M] \quad (31)$$

o reescrito como:

$$h[k] = \{B_0, B_1, \dots, B_M\} \quad (32)$$

Donde $\delta[k]$ es el impulso de entrada y $B_0, B_1, \dots, B_M$ son los coeficientes del filtro (Ambardar , 1999).

Energía

En el procesamiento de señales, una señal se modela matemáticamente como una función del tiempo (Viswanathan , 2019). El concepto de *tamaño* de una señal cuantifica su intensidad o magnitud, siendo fundamental disponer de métricas para caracterizarlo.

Una de las medidas más utilizadas es la energía de la señal, definida en el dominio continuo como el área bajo el cuadrado de su magnitud:

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (33)$$

Esta integral, análoga a la energía en sistemas físicos pero aplicable a señales, proporciona una medida del tamaño independientemente de la naturaleza física de la fuente (Viswanathan , 2019).

Para señales discretas, la energía se define mediante la norma euclídea al cuadrado:

$$E_x \triangleq |x|^2 = \sum_{n=-\infty}^{\infty} |x[n]|^2 \quad (34)$$

Redes Neuronales

Las redes neuronales artificiales (RNA) son algoritmos de machine learning bioinspirados que emulan el funcionamiento de las redes neuronales biológicas. Estos algoritmos procesan información de manera distribuida a través de nodos interconectados (neuronas artificiales), permitiendo



la resolución de problemas complejos.(IBM , 2021).

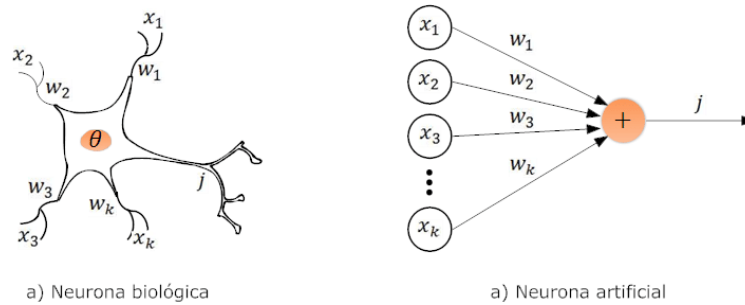


Figura 10: Neurona biologica versus neurona artificial (Fernández et al. , 2013)

Estructuralmente, una RNA se compone de capas secuenciales: una capa de entrada que recibe los datos, múltiples capas ocultas donde se realizan transformaciones no lineales, y una capa de salida que produce los resultados. Cada neurona mantiene conexiones con otras neuronas y posee un umbral de activación. La interacción entre estos elementos permite a la red aprender patrones y generalizar conocimiento a partir de ejemplos (IBM , 2021).

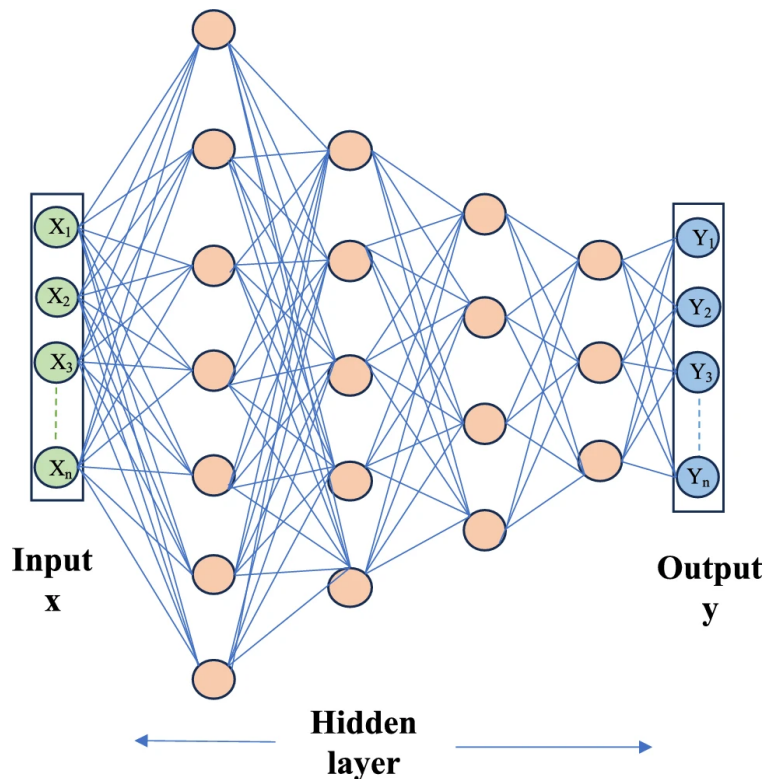


Figura 11: Gráfico de una red neuronal (Reza et al. , 2026)



Redes Neuronales Convolucionales (CNN)

Las redes neuronales convolucionales (CNN) son un tipo de modelo de deep learning. A diferencia de los enfoques clásicos de machine learning que requieren extracción manual de características, las CNN automatizan este proceso mediante el aprendizaje jerárquico de patrones directamente desde los datos en su forma cruda (Vadapalli , 2025).

Las CNNs consisten de múltiples capas. La capa de convolución, donde se le aplican filtros (kernels) a los datos para poder extraer las características mas relevantes. La capa de pooling se encarga de reducir la dimensionalidad de los tensores generados, manteniendo la información mas relevante. La capa fully connected recibe todas las características extraídas de los datos, y se encarga de la clasificación o regresión del problema (The Lifesciences Magazine , n.d.) (Vadapalli , 2025).

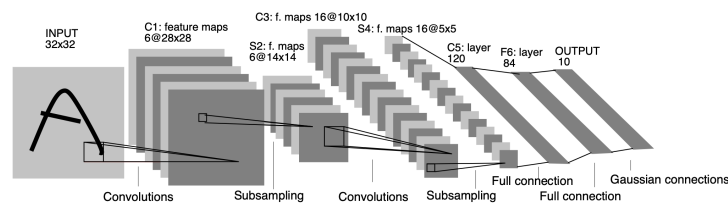


Figura 12: Arquitectura de LeNet-5, CNN desarrollada por LeCun et al. (1998)

Long Short Term Memory (LSTM)

Las redes neuronales recurrentes (RNN) son arquitecturas de redes neuronales que permiten el aprendizaje de secuencias de datos, estos son utilizados generalmente para los problemas de series temporales (Géron , 2023).

Estas presentan un problema principal, el desvanecimiento de gradiente. Esto quiere decir, que conforme se avanza en la secuencia, el modelo RNN pierde la información de los datos anteriores (Géron , 2023).

Para solucionar este problema, Hochreiter & Schmidhuber (1997) propusieron la arquitectura *Long Short-Term Memory* (LSTM), una variante de las RNN diseñada específicamente para capturar dependencias de largo plazo en secuencias de datos. La clave de la arquitectura LSTM es la

introducción de un *estado de celda* $\mathbf{c}_t$, que actúa como una memoria de largo plazo y fluye a lo largo de la secuencia con modificaciones controladas por mecanismos de compuerta (*gates*) (Géron , 2023).

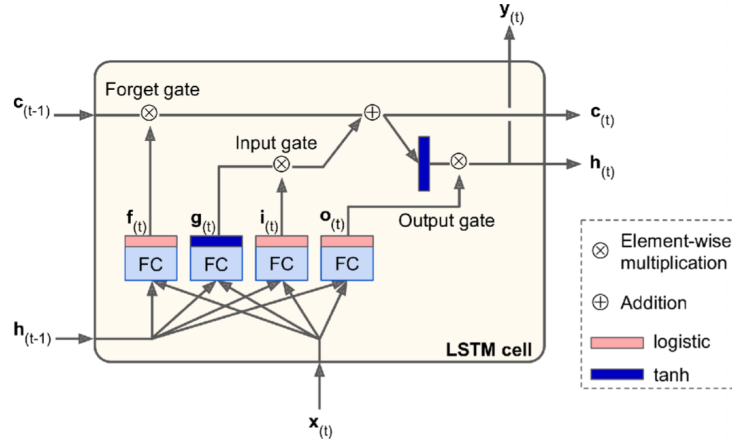


Figura 13: Arquitectura de una LSTM (Géron , 2023)

Una celda LSTM cuenta con tres compuertas que regulan el flujo de información:

- Compuerta de olvido (forget gate): determina qué fracción del estado de celda anterior $\mathbf{c}_{t-1}$ se conserva. Su activación está dada por:

$$\mathbf{f}_t = \sigma(\mathbf{W}_f [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) \quad (35)$$

- Compuerta de entrada (input gate): controla cuánta información nueva $\tilde{\mathbf{c}}_t$ se escribe en el estado de celda:

$$\mathbf{i}_t = \sigma(\mathbf{W}_i [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i), \quad \tilde{\mathbf{c}}_t = \tanh(\mathbf{W}_c [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_c) \quad (36)$$

- Compuerta de salida (output gate): regula qué parte del estado de celda actualizado $\mathbf{c}_t$ se expone como estado oculto $\mathbf{h}_t$:

$$\mathbf{o}_t = \sigma(\mathbf{W}_o [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o) \quad (37)$$

La actualización del estado de celda y del estado oculto en cada paso de tiempo t se define como:

$$\mathbf{c}_t = \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tilde{\mathbf{c}}_t \quad (38)$$

$$\mathbf{h}_t = \mathbf{o}_t \odot \tanh(\mathbf{c}_t) \quad (39)$$

donde $\sigma(\cdot)$ es la función sigmoide, $\tanh(\cdot)$ es la tangente hiperbólica, $\odot$ denota el producto elemento a elemento (Hadamard), $\mathbf{x}_t$ es la entrada en el instante t , y $\mathbf{W}_{\{\cdot\}}$, $\mathbf{b}_{\{\cdot\}}$ son los pesos y sesgos entrenables de cada compuerta (Hochreiter & Schmidhuber , 1997).

Autoencoders

Los autoencoders son arquitecturas neuronales diseñadas para aprender representaciones latentes (codificaciones densas) de los datos de entrada mediante aprendizaje no supervisado (Géron , 2023). Estructuralmente, están optimizados para reconstruir sus entradas a través de un proceso de codificación-decodificación, donde la capa intermedia (latente) captura características esenciales de los datos. Cuando esta arquitectura contiene múltiples capas ocultas en sus componentes codificador y decodificador, se denominan deep autoencoders (Géron , 2023).

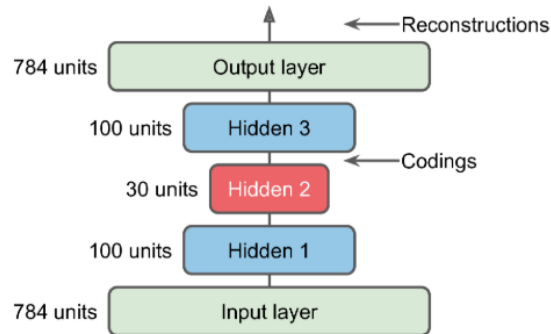


Figura 14: Arquitectura de un deep autoencoder (Géron , 2023)

Otra aplicación fundamental de los autoencoders es la tarea de denoising (eliminación de ruido), donde el modelo aprende a reconstruir señales limpias a partir de entradas con ruido. En este enfoque, conocido como denoising autoencoder (DAE), la red neuronal recibe como entrada datos contaminados con ruido (típicamente gaussiano) y es entrenada para recuperar la señal original no



contaminada. Este proceso fuerza al autoencoder a aprender características robustas y representaciones latentes invariantes al ruido (Géron , 2023).

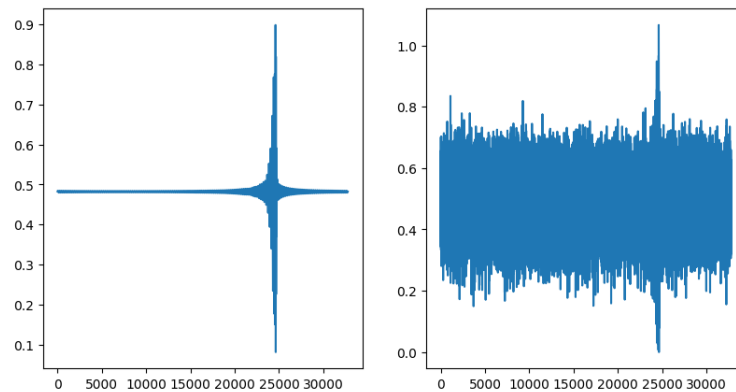


Figura 15: Datos de entrada con (derecha) y sin ruido (izquierda) para un autoencoder

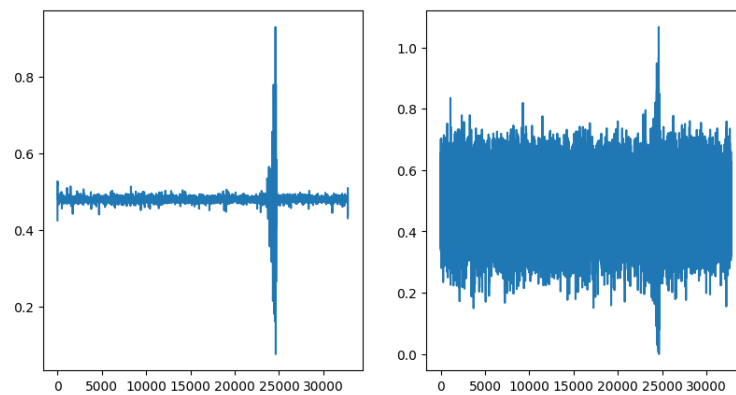


Figura 16: Datos de salida (izquierda) dada una entrada con ruido (derecha) de un Autoencoder entrenado

Optimización por nube de partículas (PSO)

El método de optimización por nube de partículas (PSO) es un algoritmo bioinspirado propuesto por Kennedy & Eberhart (1995) En el algoritmo PSO, es necesario modelar un conjunto de soluciones potenciales las cuales serán recorridas por las partículas para encontrar la solución óptima. Durante el recorrido, las partículas ajustan su posición de acuerdo a su propia experiencia (búsqueda local) y la experiencia de las partículas vecinas (búsqueda global).

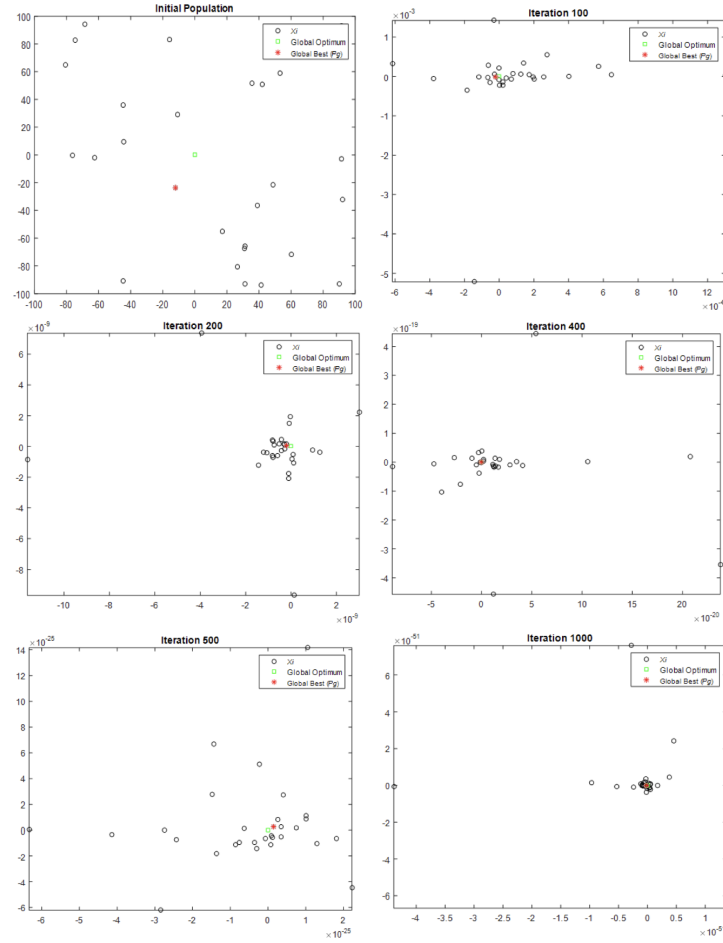


Figura 17: Ejemplo de PSO en 2D (Karim et al. , 2020)

Un algoritmo PSO consiste en un proceso iterativo. La posición de cada partícula representa una solución potencial al problema. Generalmente, una partícula p_i está compuesta de tres vectores y dos valores de fitness

- El vector $x_i = \{x_1, x_2, x_3, \dots, x_i\}$ almacena la posición actual de la partícula en el espacio.
- El vector $pBest_i = \{p_1, p_2, p_3, \dots, p_i\}$ almacena la mejor posición encontrada en la iteración.
- El vector $v_i = \{v_1, v_2, v_3, \dots, v_i\}$ almacena el gradiente en la cual se moverá la partícula.
- El valor $fitness_x_i$ almacena el desempeño de la solución actual (x_i).
- El valor $fitness_pBest$ almacena el desempeño de la mejor solución encontrada en la iteración ($pBest$).

Para inicializar el algoritmo, se generan valores iniciales a las posiciones y velocidades de cada partícula. Las posiciones se pueden generar aleatoriamente, de forma regular, o una combinación de ambas. Una vez generadas las posiciones, se calcula su desempeño y se actualizan los vectores $fitness_x$ y $fitness_pBest$.

Las velocidades se generan aleatoriamente, con cada componente en el intervalo $[-V_{max}, V_{max}]$, donde V_{max} es la velocidad máxima que puede tomar una partícula. Este valor no es conveniente dejarlo en 0, dado que no entrega buenos resultados.

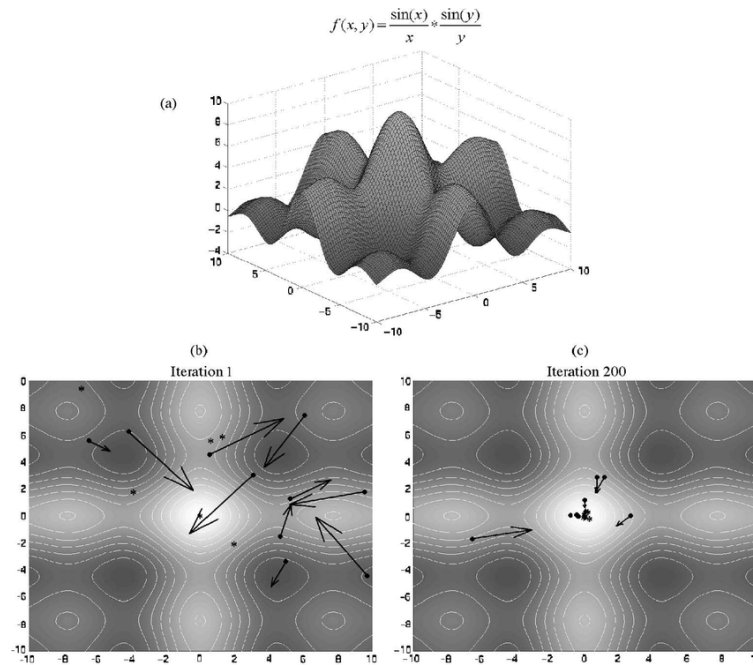


Figura 18: Ejemplo de PSO en 3D (Robinson & Rahmat-Samii , 2004)

Una vez que se inicializa la nube de partículas, se procede al proceso iterativo. Se inicia en el movimiento de las partículas en el espacio de búsqueda. Esto se realiza calculando un nuevo vector de posición x_{i+1} realizando una suma de la posición actual x_i y su velocidad v_i .

$$x_{i+1} \leftarrow x_i + v_i \quad (40)$$

Con la nueva posición calculada, se actualiza el $fitness_x$ y se evalúa si el resultado ha sido el mejor obtenido, para actualizar $pBest$. El vector de velocidad va a ser actualizado en cada iteración, utilizando la velocidad anterior y lo que se conoce como ζ componente cognitivo y ζ componente



social”. Esto representado con las siguientes ecuaciones.

$$v_i(k+1) = w \cdot v_i(k) + \varphi_1 \cdot rand_1 \cdot (pBest_i - x_i(k)) + \varphi_2 \cdot rand_2 \cdot (g_i - x_i(k)) \quad (41)$$

$$x_i(k+1) = x_i(k) + v_i(k+1) \quad (42)$$

donde

- $v_i(k)$ es la velocidad de la partícula i en la k -ésima iteración
- w es un factor de inercia
- φ_1, φ_2 son las tasas de aprendizaje para los componentes cognitivo y social respectivamente
- $rand$ es un valor aleatorio entre 0 y 1
- $x_i(k)$ es la posición actual para la partícula i en la k -ésima iteración.
- $pBest_i$ es la mejor solución para la partícula i
- g_i representa la posición de la partícula con el mejor $fitness_pBest$ del entorno p_i (local) o de toda la nube (globalbest)

El componente cognitivo $\varphi_1 \cdot rand_1 \cdot (pBest_i - x_i(k))$, representa la distancia entre la posición actual y la mejor conocida por esta partícula, es decir, la decisión que tomará la misma con base en su propia experiencia. El modelo social está modelado por el factor $\varphi_2 \cdot rand_2 \cdot (g_i - x_i(k))$ y representa la distancia entre la posición actual y la mejor posición de los vecinos, es decir, la decisión será influenciada por el resto de la nube.

Un problema habitual del PSO es que la magnitud de la velocidad suele llegar a ser muy grande, y por ende, las partículas se mueven muy rápido. Para evitar este excesivo crecimiento, se utiliza el factor de inercia w . Este factor se reduce gradualmente en cada iteración.

$$w = w_{max} - \frac{w_{max} - w_{min}}{iter_{max}} \cdot iter \quad (43)$$



Metodología

El proceso metodológico se organiza en tres etapas fundamentales:

- Creación de la base de datos
- Procesamiento de las señales
- Entrenamiento de las redes neuronales

Esta estructuración permite desarrollar el flujo metodológico de manera coherente y sistemática, abarcando desde la generación de los datos iniciales hasta la obtención de resultados finales.

El flujo completo del sistema puede describirse como sigue. Primero se simulan ondas gravitacionales utilizando *Bilby* como generador, almacenando los pares de señal limpia $h(t)$ y señal contaminada $d(t) = h(t) + n(t)$ junto con sus parámetros físicos en una base de datos. A continuación, una señal de entrada $d(t)$ es sometida a la cadena de procesamiento FIR-Wavelet-FIR para extraer el vector compacto de desviaciones estándar energéticas $\sigma = [\sigma_0, \dots, \sigma_{W-1}]$, que alimenta al clasificador LSTM para determinar si contiene información relevante (onda gravitacional) o es simplemente ruido.

Si la señal es clasificada como positiva, se procesa mediante el autoencoder convolucional para obtener la señal reconstruida $\hat{h}(t)$. Finalmente, sobre $\hat{h}(t)$ se aplica la transformada wavelet y el análisis de diferencias de energía para localizar la región de interés, que es entregada al regresor CNN para inferir la masa del chirp (Cutler & Flanagan, 1994):

Simulación de GWs y creación de la base de datos

Simulación

Como demuestran los resultados de LIGO en 2016 (Abbott et al., 2016), la clave para detectar ondas gravitacionales con éxito es disponer de datos de alta calidad, tal y como exige el método de filtrado adaptado.



Parámetro	Rango	Unidad
M_1, M_2	$[5, 100]$	$M_\odot$
χ_1, χ_2	$[0, 0.8]$	—
θ_1, θ_2	$[0, \pi]$	rad
$\delta\phi, \phi_{JL}$	$[0, 2\pi]$	rad
d_L	$[100, 5000]$	Mpc
α	$[0, 2\pi]$	rad
$\delta\alpha$	$[-\frac{\pi}{2}, \frac{\pi}{2}]$	rad
ι	$[0, \pi]$	rad
ψ	$[0, \pi]$	rad
ϕ_c	$[0, 2\pi]$	rad

Cuadro 1: Rango de valores de los parametros manejados por *Bilby* (Ashton et al., 2019)

Del mismo modo, si se pretende utilizar técnicas de machine learning como otro enfoque para la detección de ondas gravitacionales, es de gran importancia trabajar con un conjunto de datos de alta calidad para entrenar dichos modelos.

Una de las herramientas más utilizadas para la simulación de ondas gravitacionales es *Bilby* (Ashton et al., 2019). Esta librería fue concebida específicamente para realizar inferencia bayesiana a gran escala de los parámetros de las ondas gravitacionales. Su arquitectura le permite procesar tanto datos observacionales reales de GW como simular una amplia gama de eventos astrofísicos capaces de emitir este tipo de radiación.

En *Bilby*, la onda es una función de sus parámetros θ , los cuales consisten en al menos 8 parámetros intrínsecos, y 7 parámetros extrínsecos. Los parámetros intrínsecos son aquellos que dependen directamente del evento astrofísico, mientras que los extrínsecos son aquellos que dependen de los instrumentos utilizados para detectarlo.

Los parámetros intrínsecos manejados por *Bilby* son, las masas de los agujeros negros ($M_{1,2}$), la magnitud del spin de los agujeros negros ($\chi_{1,2}$), el ángulo de inclinación entre el spin y el momento



Parámetro	Valores	Unidad
M_1, M_2	(25), (35)	$M_\odot$
χ_1, χ_2	(0), (0)	—
θ_1, θ_2	(0,5), (1)	rad
$\delta\phi, \phi_{JL}$	(1,7), (0,3)	rad
d_L	500	Mpc
ra	1,375	rad
dec	−1,2108	rad
ι	0,4	rad
ψ	0,6	rad
ϕ_c	1,3	rad

Cuadro 2: Valores ejemplo para la simulación de una GW.

angular orbital ($\theta_{1,2}$), y dos vectores de espín que describen el ángulo acimutal que separa los vectores de espín $\delta\phi$ y el cono de precesión alrededor del momento angular del sistema ϕ_{JL} (Ashton et al., 2019).

De la misma manera, los parámetros extrínsecos son aquellos que dependen del instrumento utilizado para detectar el evento. Estos parámetros son, la distancia de luminosidad d_L , la ascensión ra y declinación dec, el ángulo de inclinación entre la línea de visión del observador y el momento angular orbital ι , el ángulo de polarización ψ y la fase de coalescencia ϕ_c (Ashton et al., 2019). Estos parámetros trabajan dentro de un rango específico de valores, como se muestra en la Tabla 1.

Una vez definidos estos parámetros, Bilby puede generar simulaciones de ondas gravitacionales. Tomemos como ejemplo, los valores paramétricos presentados en la Tabla 2. Estos valores sirven como entrada para Bilby, que produce como resultado la forma de onda mostrada en la Figura 19.

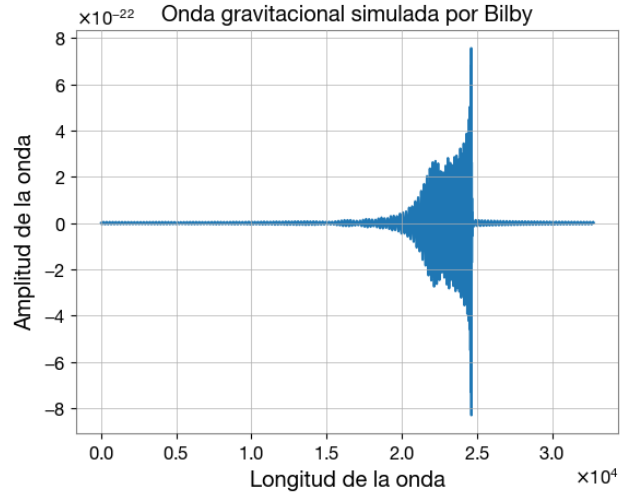


Figura 19: Simulación de una GW con los parámetros de la Tabla 2

Bilby proporciona un conjunto de interferómetros preconfigurados con las características necesarias para simular su respuesta. En este estudio, se utilizó la configuración correspondiente al observatorio LIGO Hanford.

Una vez generada la forma de onda (strain), se procede a calcular la relación señal-ruido óptima SNR_{opt} del evento. La señal se evalúa en el detector seleccionado para caracterizar su comportamiento en condiciones realistas. Con esta información, es posible calcular la densidad espectral de potencia (PSD) del detector y, posteriormente, determinar el valor de SNR_{opt} (Ashton et al., 2019).

Retomando el ejemplo de la Figura 19, al evaluar la señal en el detector LIGO Hanford se obtiene una relación señal-ruido óptima SNR_{opt} de 16.2 dB. Como resultado de este proceso, se genera la señal mostrada en la Figura 20, representativa de la señal final $h(t)$.

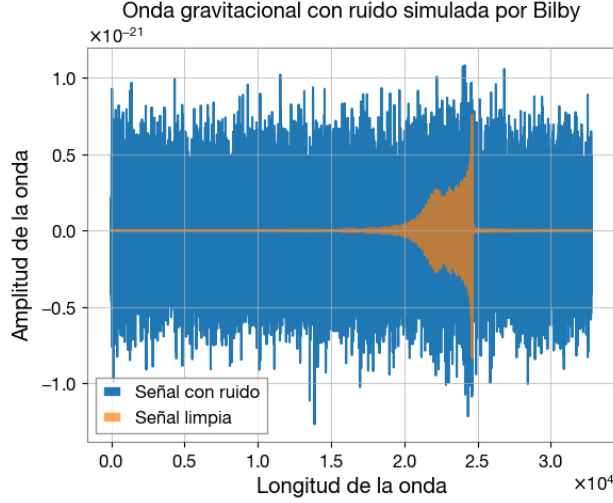


Figura 20: Simulación de una GW con los parámetros de la Tabla 2 y un SNR_{opt} de 16.2 dB

Base de datos

Para construir un conjunto de datos significativo, los eventos deben representarse en diversas condiciones de SNR. El SNR óptimo sirve como referencia para este fin. Podemos generar una señal con una SNR específica y deseada escalando la plantilla.

Dado que la SNR óptima es una función de la amplitud de la plantilla, una plantilla $h'(t) = \alpha h(t)$ tendrá una SNR óptima de $\rho'_{\text{opt}} = \alpha \rho_{\text{opt}}$. (Cutler & Flanagan , 1994)

Para alcanzar la relación señal-ruido deseada (convertida de decibeles a una escala lineal), el factor de escala requerido α es:

$$\alpha = \sqrt{\frac{\rho_{\text{opt}}}{10^{\text{SNR}_{\text{db}}/10}}} \quad (44)$$

donde SNR_{db} es la SNR deseada en decibeles. Al aplicar este factor a la plantilla original, $h'(t) = \alpha h(t)$, se obtiene una forma de onda cuya SNR óptima coincide con el valor lineal deseado de SNR.

A partir de la configuración mostrada en la Figura 19, pueden generarse distintas versiones del mismo evento variando el SNR, obteniendo así una representación completa del comportamiento de la señal. Como se aprecia en la Figura 21, un SNR bajo (0 dB) resulta en el predominio del



ruido, mientras que un SNR alto (30 dB) permite observar la señal limpia con mínima interferencia.

La construcción de la base de datos requiere la implementación de un proceso iterativo que permita generar y almacenar sistemáticamente un volumen significativo de señales bajo distintas configuraciones paramétricas. Esto se realiza mediante el cálculo de todas las combinaciones posibles de los eventos que se pueden generar.

Dada la explosión combinatoria que resultaría de calcular todas las posibles combinaciones de parámetros, se optó por una estrategia de muestreo selectivo. La mayoría de los parámetros se mantuvieron en sus valores predeterminados, variando únicamente las masas de los agujeros negros y el SNR deseado.

Una vez determinados los parámetros a variar, es necesario definir sus respectivos dominios de valores. Para las masas, los valores de m_1 y m_2 oscilan entre $10M_\odot$ y $100M_\odot$, conforme a los límites establecidos en la Tabla 1. Adicionalmente, se impone una restricción en la relación de masas, definida en la ecuación (45), que garantiza que ninguna de las componentes supere el doble de la masa de su acompañante.

$$\frac{\max(m_1, m_2)}{\min(m_1, m_2)} < 2 \quad (45)$$

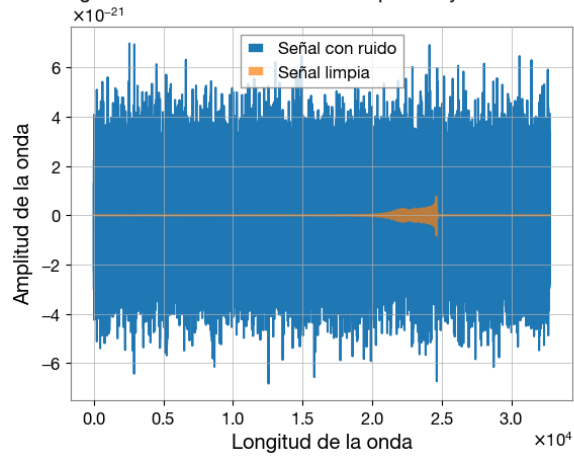
Para el SNR deseado, se establece un dominio de 10 dB a 50 dB.

Si tomamos un incremento de masa de $2M_\odot$, el número total de muestras sería 6125. Del mismo modo, para la SNR, establecemos un incremento de paso de 2 dB. De esta manera, el número total de muestras generadas sería de 122,500 señales.

Las señales generadas se extraen como vectores de puntos que representan series temporales. Se almacenan dos versiones de cada señal: la señal contaminada con ruido $n(t)$ y la señal limpia original $h(t)$. Ambos vectores se incorporan a la base de datos.

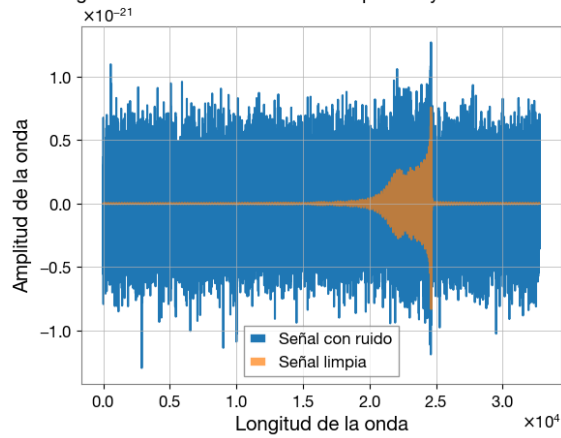


Onda gravitacional con ruido simulada por Bilby con un SNR de 0



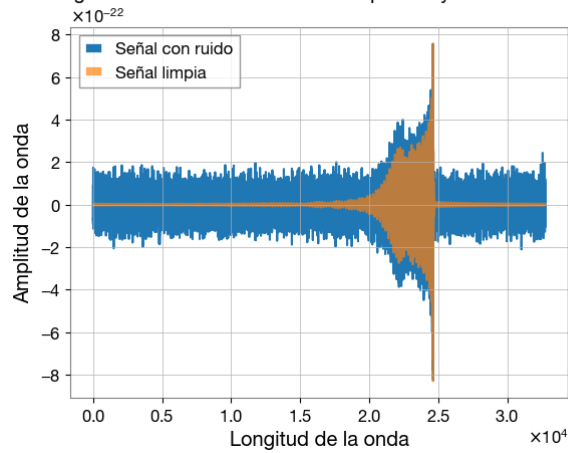
(a) Señal generada con una SNR de 0 dB.

Onda gravitacional con ruido simulada por Bilby con un SNR de 16.2



(b) Señal generada con una SNR de 16.2 dB.

Onda gravitacional con ruido simulada por Bilby con un SNR de 30



(c) Señal generada con una SNR de 30 dB.

Figura 21: Diferentes comportamientos de la misma señal generada bajo diferentes valores de SNR.



Parámetro	Rango/Valor	Unidad
Señal con ruido	$[n(t)0, \dots, n(t)N]$	-
Señal limpia	$[h(t)0, \dots, h(t)N]$	-
Masa primaria (M_1)	$[10, 100]$	$M_\odot$
Masa secundaria (M_2)	$[10, 100]$	$M_\odot$
Espín de componente 1 (χ_1)	0 (fijo)	-
Espín de componente 2 (χ_2)	0 (fijo)	-
Relación señal-ruido (SNR)	$[10, 50]$	dB
Ángulo de inclinación 1 (θ_1)	0,5 (fijo)	rad
Ángulo de inclinación 2 (θ_2)	1,0 (fijo)	rad
Ángulo acimutal ($\delta\phi$)	1,7 (fijo)	rad
Ángulo de precesión (ϕ_{JL})	0,3 (fijo)	rad
Distancia de luminosidad (d_L)	500 (fijo)	Mpc
Ascensión recta (ra)	1,375 (fijo)	rad
Declinación (dec)	-1,2108 (fijo)	rad
Ángulo de inclinación orbital (ι)	0,4 (fijo)	rad
Ángulo de polarización (ψ)	0,6 (fijo)	rad
Fase de coalescencia (ϕ_c)	1,3 (fijo)	rad

Cuadro 3: Configuración de parámetros utilizada en la generación de la base de datos mediante *Bilby* (Ashton et al., 2019). Los parámetros marcados como fijos mantienen un valor constante, mientras que los especificados con un intervalo varían dentro del rango indicado.

De esta forma, la configuración actual de la base de datos puede resumirse mediante los siguientes parámetros y sus respectivos rangos de variación de la tabla 3:



Configuración final de la base de datos

Transformada de Wavelet

Como puede observarse en las Figuras 21a y 21b, las señales se encuentran enmascaradas predominantemente por ruido. Entrenar un modelo de machine learning directamente con estas muestras resultaría ineficaz, dado que la información relevante permanece prácticamente oculta. Por esta razón, se hace indispensable un procesamiento de señales que permita resaltar los componentes más relevantes, atenuar el ruido y generar representaciones adecuadas para el entrenamiento de modelos de machine learning.

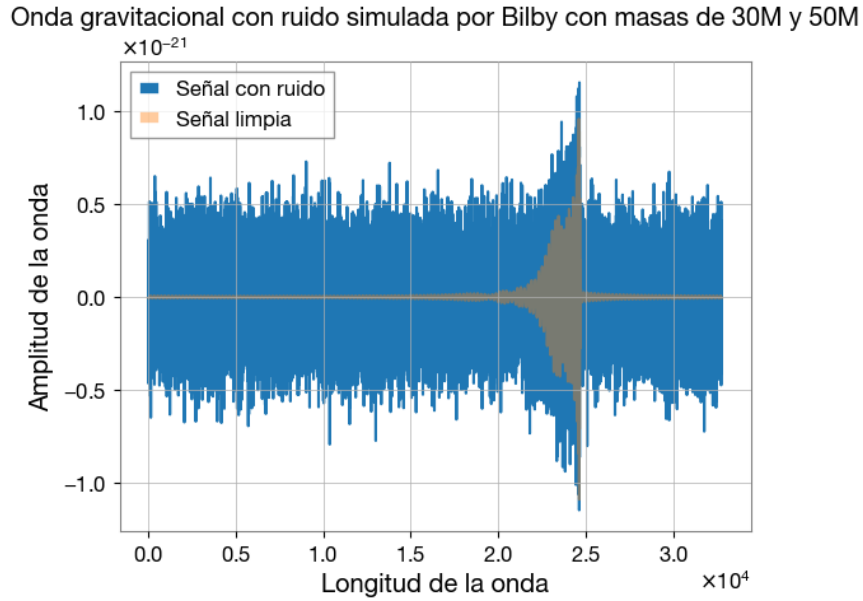


Figura 22: Simulación de una GW con masas de 30 $M_{\odot}$ y 50 $M_{\odot}$, con un SNR de 20 dB.

Una estrategia efectiva para extraer información relevante de señales contaminadas por ruido consiste en emplear la transformada wavelet discreta (DWT) (Polikar , 1996). Retomando la ecuación (27), la DWT se define mediante el análisis de una señal $x(t)$ y una wavelet madre $\psi(t)$ evaluada a diferentes escalas j y posiciones k .

$$DWT_{j,k} = 2^{j/2} \int_{-\infty}^{\infty} x(t) \psi(2^j t - k) dt$$



Este proceso genera una descomposición multirresolución de $x(t)$, permitiendo aislar componentes de interés en distintos niveles de detalle, para finalmente, reconstruir la señal original $x(t)$, utilizando la ecuación (28).

$$x(t) = \sum_k \sum_j c_{j,k} \phi(t) + \sum_k \sum_j d_{j,k} \psi(t)$$

La prueba de diferentes familias de wavelets $\psi(t)$ arrojó, que la familia de Haar es la que mejor representaba la estructura de la señal. La ecuación (28) nos dice que, $x(t)$ requiere la función de escala $\phi(t)$ y la función wavelet $\psi(t)$, que para la wavelet de Haar vienen dadas por

$$\psi(x) = \begin{cases} 1 & \text{si } 0 \leq x < 1/2 \\ -1 & \text{si } 1/2 \leq x < 1 \\ 0 & \text{en los demás casos} \end{cases}, \quad (46)$$

$$\phi(x) = \begin{cases} 1 & \text{si } 0 \leq x < 1 \\ 0 & \text{en caso contrario} \end{cases} \quad (47)$$

Esta transformación permite descomponer y comprimir la señal, separando sus componentes característicos en diferentes bandas de frecuencia. Consideremos como referencia la señal original mostrada en la Figura 22. En la Figura 23 se ilustra el resultado de la descomposición multinivel mediante transformada wavelet, donde puede observarse cómo cada nivel de resolución captura distintos aspectos de la señal: desde las componentes de alta frecuencia (detalles finos) hasta las de baja frecuencia (tendencia principal). Esta descomposición facilita la identificación y extracción de características relevantes para el análisis posterior.

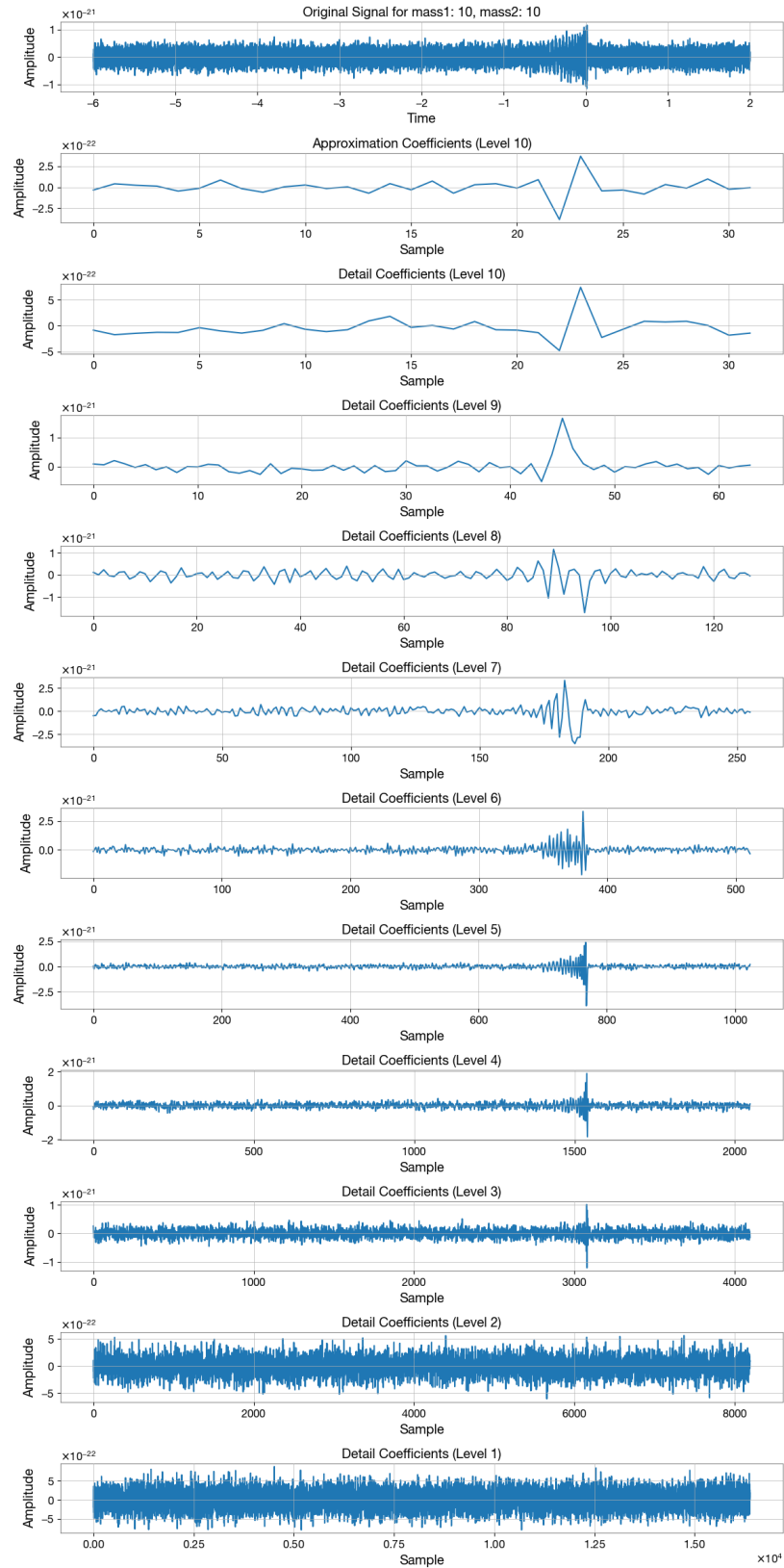


Figura 23: Descomposición multinivel de la señal mediante transformada wavelet.



De esta manera, la descomposición multinivel permite aislar escalas de resolución específicas donde la firma del merger del sistema binario se manifiesta con mayor claridad. Al analizar selectivamente estos niveles, es posible identificar con mayor precisión la zona de interés correspondiente al instante de coalescencia, incluso en señales contaminadas con ruido.

Por lo tanto, se puede seleccionar el nivel de descomposición que mejor se adapte a la señal. En este caso, un nivel 4 será suficiente para localizar la zona de interés

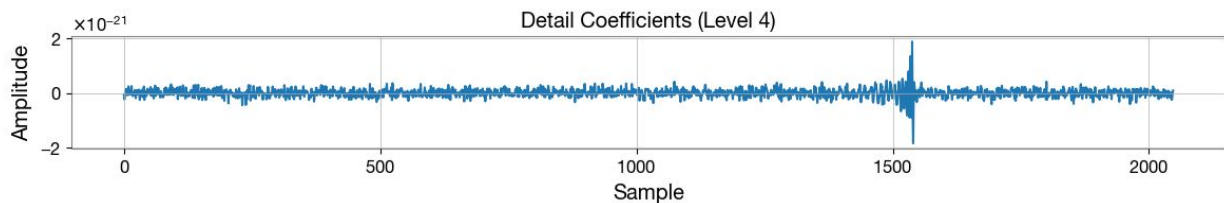


Figura 24: Descomposición nivel 4 de la señal mediante transformada wavelet.

Un beneficio adicional de la transformada wavelet es la significativa reducción en la dimensionalidad de las señales. Mientras que la señal original de la Figura 22 contiene un vector de aproximadamente 3×10^4 puntos muestrales, su representación wavelet en la Figura 24 se comprime a únicamente alrededor de 2000 coeficientes relevantes. Esta compresión efectiva reduce la complejidad computacional de los datos y acelera considerablemente el entrenamiento de modelos de machine learning, sin comprometer la información fundamental de la señal.

Diferencia de energía

Si bien la transformada wavelet constituye una herramienta poderosa para la descomposición de señales, la representación obtenida en la Figura 24 resulta insuficiente para identificar de manera robusta la zona de interés. Esta limitación evidencia la necesidad de incorporar una etapa adicional de procesamiento: el análisis de diferencias de energía.

Una señal en el procesamiento de señales se modela como una función del tiempo, y su intensidad se cuantifica mediante medidas de tamaño, como la energía de la señal. Para una señal en tiempo continuo, esto se define como el área bajo el cuadrado de su magnitud:



$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (48)$$

La definición análoga para una señal discreta utiliza la norma euclídea al cuadrado:

$$E_x = \sum_{n=-\infty}^{\infty} |x[n]|^2 \quad (49)$$

Para localizar regiones de alta actividad energética dentro de la señal, se implementa un análisis de diferencias de energía. Este método calcula la energía en segmentos consecutivos, identificando la ventana con el valor máximo como la zona de interés más prominente (Farret et al. , 2017). El tamaño de la ventana w se adapta a la resolución de la señal mediante la relación:

$$w = \frac{L_s}{50} \quad (50)$$

donde L_s es la longitud de la señal.

El procedimiento consiste en dividir la señal en segmentos sucesivos de longitud w muestras. Para cada segmento, se calcula la energía de la señal. La ventana de análisis se desplaza en incrementos de w muestras, repitiendo el cálculo de energía hasta cubrir toda la extensión de la señal. Finalmente, se selecciona la ventana con el valor máximo de energía, identificando así la región de mayor interés dentro de la señal.

Retomando el análisis de la Figura 24, la aplicación del método de diferencias de energía produce un efecto equivalente a un *zoom* analítico sobre la zona de interés, como se evidencia en la Figura 25.

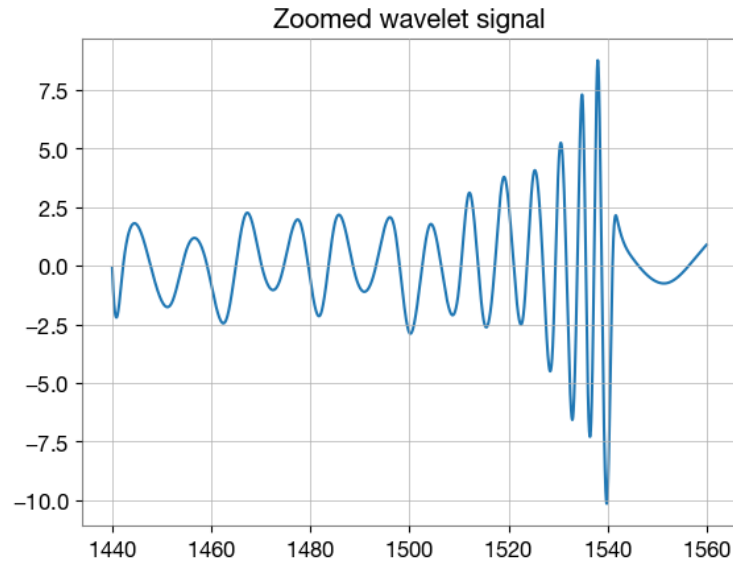


Figura 25: Realce de la zona de interés mediante análisis de diferencias de energía aplicado a la descomposición wavelet nivel 4.

Este procedimiento reduce adicionalmente la complejidad computacional del análisis, ya que el procesamiento se concentra exclusivamente en una ventana de w muestras.

En el caso de la Figura 26, se observa el efecto de la aplicación del análisis de diferencias de energía sobre una señal sin ruido. Se puede apreciar que las zonas de interés presentan un comportamiento similar.

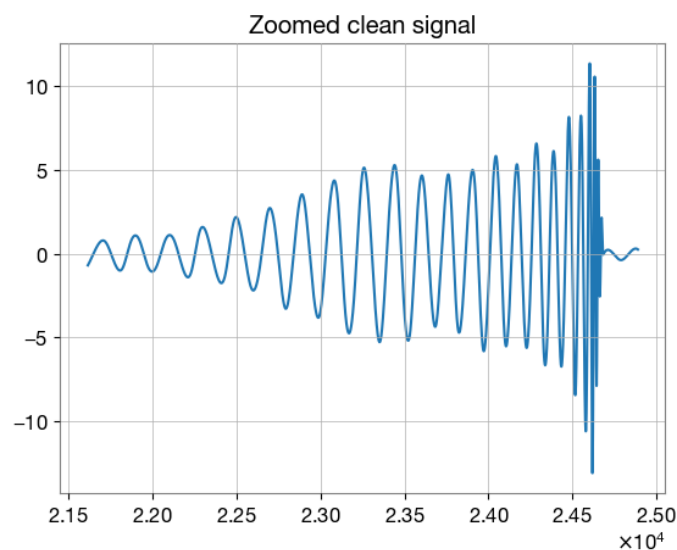


Figura 26: Zona de interés identificada para una señal sin ruido.



Parámetro	Rango/Valor	Unidad
Señal con ruido	$[n(t)0, \dots, n(t)N]$	-
Señal limpia	$[h(t)0, \dots, h(t)N]$	-
Masa primaria (M_1)	$[10, 100]$	$M_\odot$
Masa secundaria (M_2)	$[10, 100]$	$M_\odot$
Espín de componente 1 (χ_1)	0 (fijo)	-
Espín de componente 2 (χ_2)	0 (fijo)	-
Relación señal-ruido (SNR)	$[10, 50]$	dB
Ángulo de inclinación 1 (θ_1)	0,5 (fijo)	rad
Ángulo de inclinación 2 (θ_2)	1,0 (fijo)	rad
Ángulo acimutal ($\delta\phi$)	1,7 (fijo)	rad
Ángulo de precesión (ϕ_{JL})	0,3 (fijo)	rad
Distancia de luminosidad (d_L)	500 (fijo)	Mpc
Ascensión recta (ra)	1,375 (fijo)	rad
Declinación (dec)	-1,2108 (fijo)	rad
Ángulo de inclinación orbital (i)	0,4 (fijo)	rad
Ángulo de polarización (ψ)	0,6 (fijo)	rad
Fase de coalescencia (ϕ_c)	1,3 (fijo)	rad
Zoom (Zona de interés)	$[Z(t)0, \dots, Z(t)N]$	-

Cuadro 4: Configuración final de la base de datos.

Si bien el análisis de diferencias de energía aplicado a la descomposición wavelet arroja resultados favorables e influirá en el desarrollo metodológico, es crucial considerar su integración con el entrenamiento de modelos de machine learning. Por este motivo, se decidió extraer las regiones de interés identificadas mediante el zoom analítico de las señales limpias, las cuales se utilizarán para el posterior entrenamiento de los modelos.

De este modo se obtiene la configuración final de la base de datos, cuyas características se resumen en la Tabla 4. Este conjunto de datos servirá como base para el desarrollo y entrenamiento de los modelos de machine learning.



Análisis de la señal

Cadena de filtrado FIR y extracción de características para detección

Para la tarea de detección binaria de ondas gravitacionales, se implementa una cadena de procesamiento que combina filtros digitales de respuesta finita al impulso (FIR) con la descomposición wavelet discreta. El objetivo es atenuar el ruido de banda ancha y resaltar las características energéticas de la señal que resulten discriminativas para un clasificador.

Esta cadena se describe en el método signal filtering de la clase Detection y consta de tres etapas secuenciales aplicadas sobre la señal de entrada $x[n]$ muestreada a $f_s = 4,096$ Hz.

Tengamos como referencia la figura 27. La cual se compone por dos cuerpos de $10M_{\odot}$ y con un SNR de 20 dB.

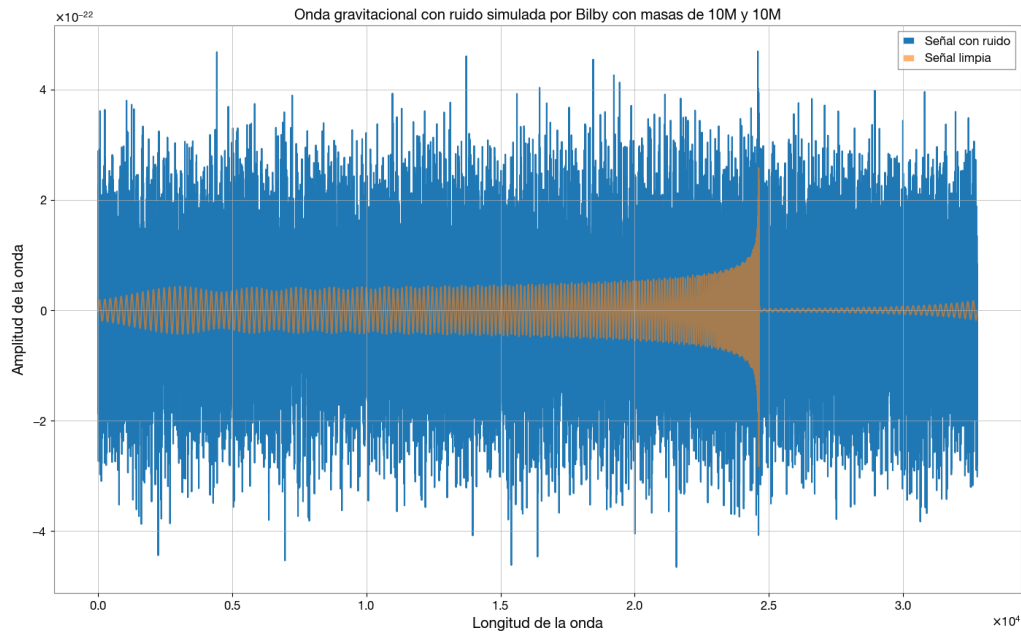


Figura 27: Señal original

Primer filtro FIR paso-bajas. Se diseña un filtro FIR de $N_1 = 1,500$ coeficientes mediante el método de ventana de Hann, con frecuencia de corte normalizada $f_{c1} = 1000/4096 \approx 0,244$, equivalente a



1,000 Hz. Este primer filtro elimina las componentes de alta frecuencia no asociadas con el chirp gravitacional.

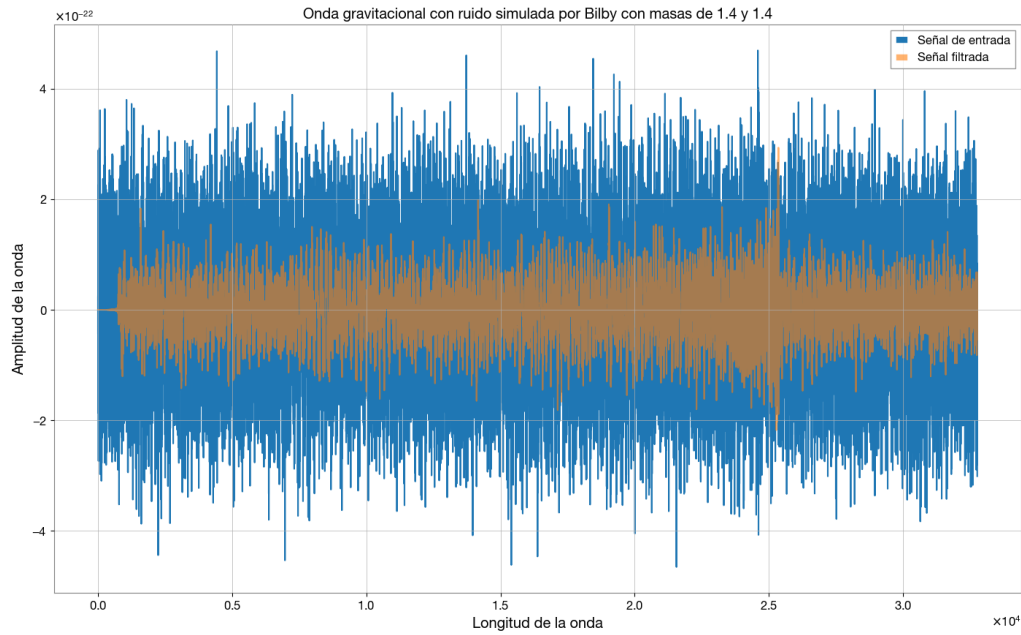


Figura 28: Primer filtro FIR paso-bajas

Una vez aplicado el primer filtro, la señal es procesada mediante una descomposición de wavelet. La señal filtrada, se procesa mediante la DWT con una wavelet madre de Haar hasta el nivel de descomposición $j = 9$. De esta manera, se extrae una señal correspondiente a dicho nivel, y se reduce la dimensionalidad de la señal.

Sobre la señal reconstruida desde el nivel 9 de descomposición wavelet, se aplica un segundo filtro FIR de $N_2 = 1,000$ coeficientes con ventana de Hann y frecuencia de corte normalizada $f_{c2} = 550/4096 \approx 0,134$, equivalente a 550 Hz. Este segundo filtro suprime los artefactos de alta frecuencia introducidos durante la reconstrucción wavelet, produciendo la señal procesada final $\tilde{x}[n]$.

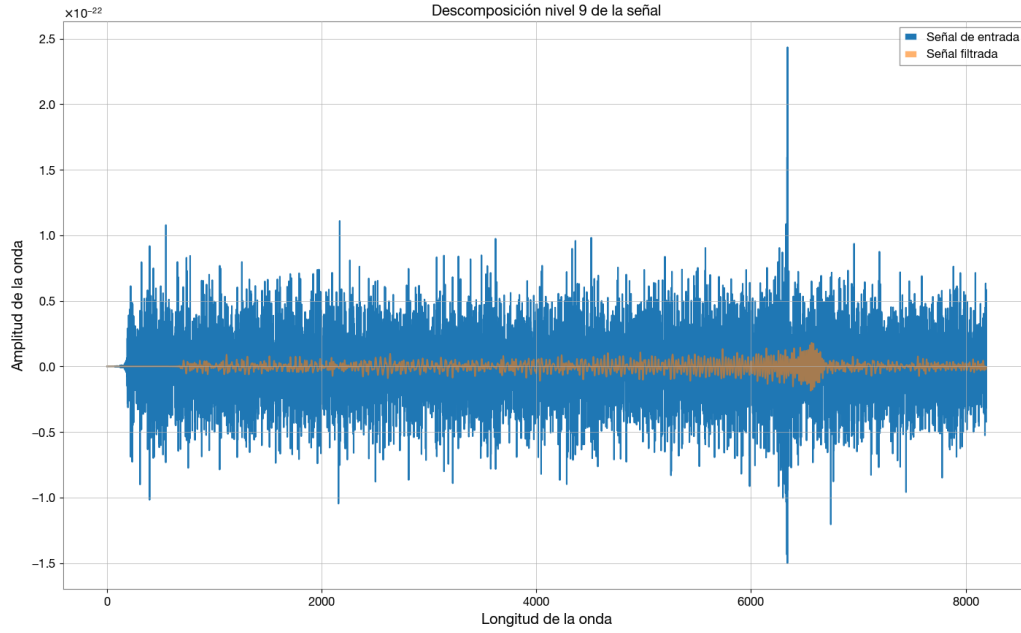


Figura 29: Segundo filtro FIR paso-bajas para la señal procesada por wavelets

Optimización conjunta del tamaño de ventana mediante PSO

La señal procesada $\tilde{x}[n]$ se divide en W segmentos de igual longitud. Para cada segmento i , se calcula la energía instantánea mediante una ventana deslizante de tamaño w_s :

$$E_i[n] = \sum_{k=0}^{w_s-1} \tilde{x}_i[n-k]^2 \quad (51)$$

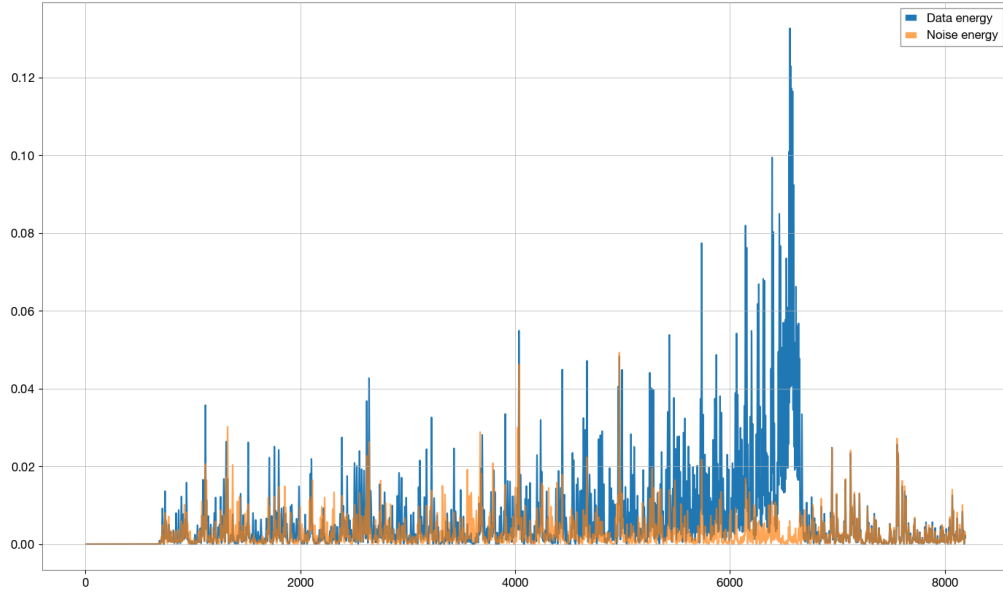


Figura 30: Energía instantánea de cada segmento

donde $\tilde{x}_i[n]$ es la subsecuencia del i -ésimo segmento. A partir de la energía calculada, se extrae la desviación estándar de cada segmento:

$$\sigma_i = \sqrt{\frac{1}{L_i} \sum_{n=0}^{L_i-1} (E_i[n] - \bar{E}_i)^2} \quad (52)$$

donde L_i es la longitud del segmento y $\bar{E}_i$ su energía media. El vector resultante $\sigma = [\sigma_0, \sigma_1, \dots, \sigma_{W-1}]$ constituye la representación compacta de la señal que se entrega como secuencia al clasificador LSTM.

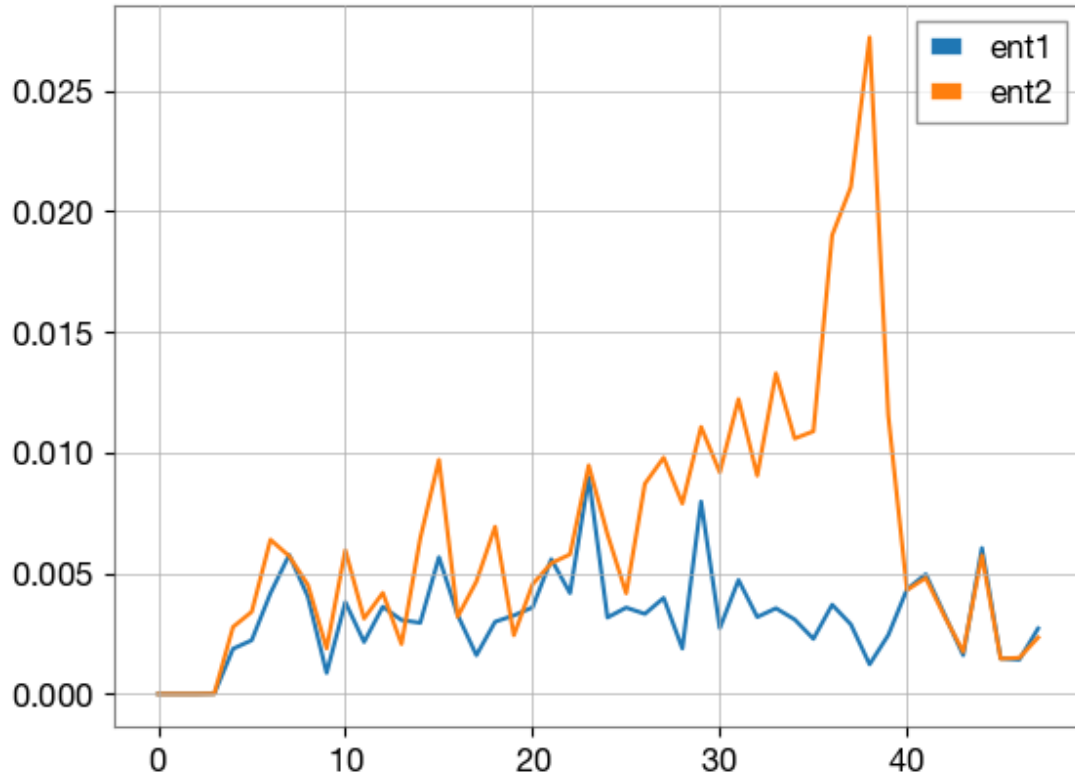


Figura 31: Desviación estándar de cada segmento

Dos parámetros son críticos en la etapa de extracción de características para el clasificador LSTM: el tamaño de la ventana deslizante w_s y el número de segmentos W en que se divide la señal procesada $\tilde{x}[n]$. Un valor de w_s demasiado pequeño captura fluctuaciones de alta frecuencia irrelevantes, mientras que uno excesivamente grande suaviza la región de interés y reduce el contraste entre señal gravitacional y ruido puro. De forma análoga, un número reducido de segmentos W diluye la información temporal, en tanto que un valor excesivo fragmenta el evento gravitacional entre múltiples ventanas, degradando la representación σ .

Para encontrar simultáneamente los valores óptimos de ambos parámetros, el problema se formula como una optimización bidimensional:



$$(w_s^*, W^*) = \arg \max_{w_s \in [w_{\min}, w_{\max}], W \in [W_{\min}, W_{\max}]} \max_i \sigma(E_i(w_s, W)) \quad (53)$$

$$E_i(w_s, W)[n] = \sum_{k=0}^{w_s-1} \tilde{x}_i^{(W)}[n-k]^2 \quad (54)$$

donde $\tilde{x}_i^{(W)}[n]$ denota la subsecuencia del i -ésimo segmento al dividir $\tilde{x}[n]$ en W partes de igual longitud. La función objetivo selecciona la desviación estándar máxima sobre los W segmentos, de modo que el par (w_s^*, W^*) maximiza el contraste energético entre segmentos, resaltando la región donde se concentra el evento gravitacional.

Ambas variables se tratan como continuas durante la búsqueda y se redondean al entero más cercano en cada evaluación de la función objetivo. El dominio de búsqueda es $w_s \in [2, 100]$ muestras y $W \in [8, 128]$ segmentos.

Los hiperparámetros del PSO se resumen en la Tabla 5:

Cuadro 5: Hiperparámetros del PSO para optimización conjunta de w_s y W .

Hiperparámetro	Valor
Número de partículas (N_p)	20
Dimensiones (d)	2
Peso de inercia (w)	0.5
Coefficiente cognitivo (c_1)	0.2
Coefficiente social (c_2)	0.8
Iteraciones máximas	500
Dominio de w_s	[2, 100]
Dominio de W	[8, 128]

El valor elevado de c_2 respecto a c_1 refleja que la convergencia hacia la mejor solución global es preferible sobre la exploración individual de cada partícula. El peso de inercia $w = 0,5$ equilibra la



velocidad de exploración con la capacidad de explotación.

Modelos de machine learning

Red LSTM para clasificación binaria de ondas gravitacionales

La detección de ondas gravitacionales se formula como un problema de clasificación binaria: determinar si una señal de entrada contiene una onda gravitacional (clase 1) o corresponde únicamente a ruido instrumental (clase 0). Para abordar esta tarea se propone un clasificador basado en redes de memoria a largo y corto plazo (LSTM) que procesa la representación energética compacta $\sigma = [\sigma_0, \sigma_1, \dots, \sigma_{T-1}]$ obtenida mediante la cadena de filtrado FIR y segmentación descrita anteriormente.

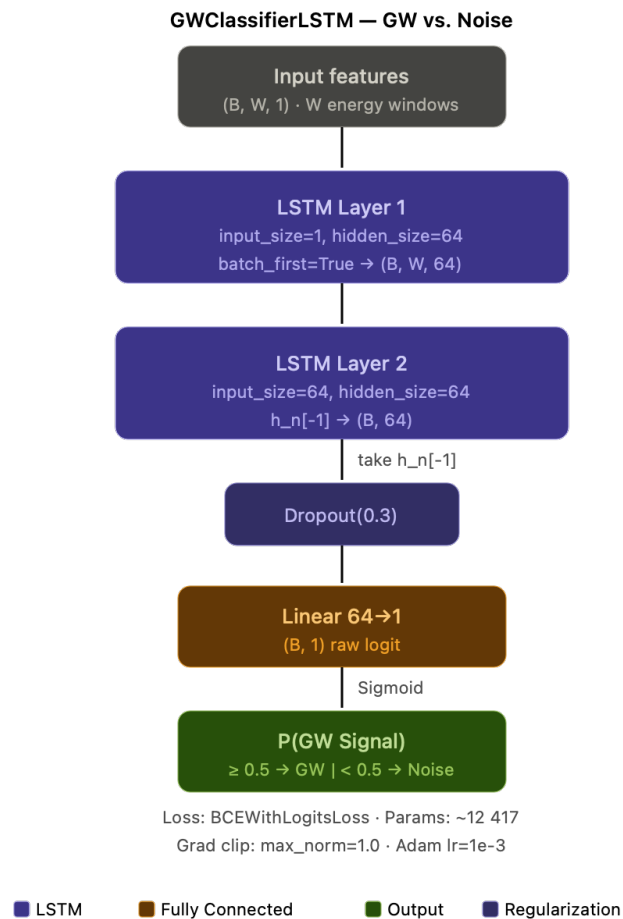




Figura 32: Arquitectura del clasificador GWClassifierLSTM para la detección binaria de ondas gravitacionales.

El modelo, denominado GWClassifierLSTM e ilustrado en la Figura 32, recibe como entrada una secuencia de $T = W$ desviaciones estándar de ventanas energéticas, representadas como T pasos temporales con una característica escalar cada uno, es decir, un tensor de dimensión $(T, 1)$. La arquitectura comprende los siguientes componentes:

Dos capas LSTM recurrentes apiladas con $H = 64$ unidades ocultas cada una. El estado oculto de la última capa en el último paso temporal, $\mathbf{h}_T \in \mathbb{R}^{64}$, condensa la información temporal de toda la secuencia. Se aplica *Dropout* con probabilidad $p = 0,3$ al estado oculto $\mathbf{h}_T$ para reducir el sobreajuste durante el entrenamiento. Una capa lineal $\mathbb{R}^{64} \rightarrow \mathbb{R}^1$ transforma el estado oculto en un logit escalar z .

La decisión de clasificación se obtiene aplicando la función sigmoide al logit:

$$\hat{y} = \sigma(z) = \frac{1}{1 + e^{-z}}, \quad \text{clase} = \begin{cases} 1 \text{ (GW)} & \text{si } \hat{y} \geq 0,5 \\ 0 \text{ (Ruido)} & \text{si } \hat{y} < 0,5 \end{cases} \quad (55)$$

Entrenamiento del clasificador LSTM

Los datos de entrenamiento se generan aplicando la cadena de procesamiento descrita anteriormente a cada señal de la base de datos: por cada muestra se extrae el vector σ , etiquetado como clase 1 (GW) o clase 0 (ruido) según corresponda. El conjunto de datos se dividió de forma estratificada en tres subconjuntos: entrenamiento (70 %), validación (15 %) y prueba (15 %). Las características se normalizaron mediante un *StandardScaler* ajustado exclusivamente sobre el conjunto de entrenamiento, para evitar fuga de información hacia los conjuntos de validación y prueba.

La función de pérdida empleada es la entropía cruzada binaria con logits (*BCEWithLogitsLoss*), que combina numéricamente de manera estable la función sigmoide con la pérdida de entropía cruzada:



$$\mathcal{L} = -\frac{1}{n} \sum_{i=1}^n [y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)] \quad (56)$$

El optimizador utilizado es Adam con tasa de aprendizaje $\alpha = 10^{-3}$ y regularización L2 con peso $\lambda = 10^{-4}$. Se implementó una política de reducción de tasa de aprendizaje (*ReduceLROnPlateau*) con factor 0,5 y paciencia de 5 épocas sin mejora en la pérdida de validación (Subramanian et al. , 2025).

Adicionalmente, se aplica *early stopping* con una paciencia de 15 épocas para detener el entrenamiento cuando la pérdida de validación deja de decrecer. Durante el entrenamiento se aplica recorte de gradiente (*gradient clipping*) con norma máxima de 1,0 para estabilizar la actualización de pesos en la red recurrente (Zhang et al. , 2019). El mejor modelo, determinado por la menor pérdida de validación, se almacena y se utiliza para la evaluación final sobre el conjunto de prueba. Los hiperparámetros se resumen en la Tabla 6.

Cuadro 6: Hiperparámetros del clasificador LSTM para detección de GWs.

Hiperparámetro	Valor
Longitud de secuencia (T)	W
Unidades ocultas LSTM (H)	64
Capas LSTM	2
Dropout	0.3
Tamaño de batch	32
Tasa de aprendizaje (η)	10^{-3}
Regularización L2 (λ)	10^{-4}
Épocas máximas	150
Paciencia <i>early stopping</i>	15
División entrenamiento / validación / prueba	70 % / 15 % / 15 %



Autoencoder

Uno de los desafíos más significativos en el análisis de señales de ondas GWs, es la presencia dominante del ruido en ellas. Si bien se ha demostrado la efectividad de un esquema de procesamiento que combina la transformada wavelet con análisis de diferencias de energía para identificar la region de interés, la presencia de ruido continúa representando una limitación crítica para el entrenamiento de modelos de machine learning.

Tomemos como ejemplo la Figura 33, donde se observa una señal GW contaminada por un ruido con SNR de 10 dB.

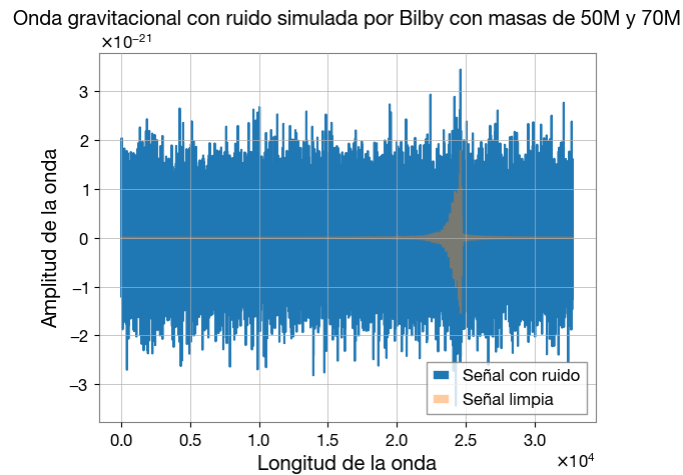


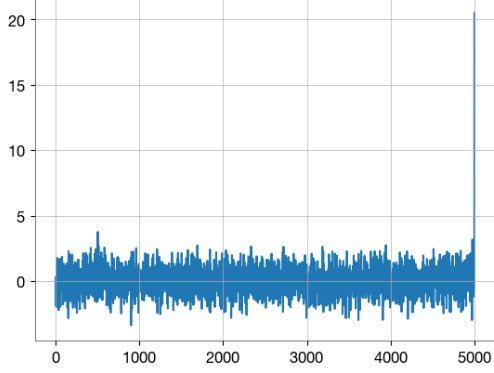
Figura 33: Señal GW contaminada por un ruido con SNR de 10 dB.

En escenarios donde el ruido domina significativamente la señal, el análisis de diferencias de energía aplicado a la descomposición wavelet pierde efectividad para identificar la región de interés. Esto se evidencia en la Figura 34, donde la descomposición wavelet produce una señal completamente distorsionada y caótica, divergiendo notablemente de la señal original debido a la alta contaminación de ruido.

Las redes neuronales, particularmente los autoencoders (AE) y sus variantes para eliminación de ruido (DAE), han demostrado ser herramientas efectivas para la supresión de ruido en señales, como avalan diversas investigaciones (Vincent et al. , 2008; ?). En el contexto específico de ondas gravitacionales, estudios previos han validado la eficacia de los DAEs, logrando una supresión sig-

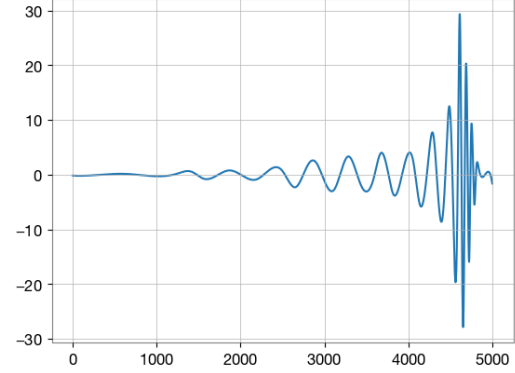


Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(a) Análisis de diferencias de energía aplicado a la descomposición wavelet para una señal con SNR de 10 dB.

Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(b) Análisis de diferencias de energía aplicado a la misma señal sin ruido.

Figura 34: Comparación entre el análisis de diferencias de energía aplicado a la descomposición wavelet para una señal con SNR de 10 dB y la misma señal sin ruido.

nificativa de ruido dentro de un rango de SNR de [6,5,26] dB (Shen et al. , 2019).

La arquitectura del Denoising Autoencoder (DAE) implementada sigue una estructura de tipo U-Net (Ulyanov et al. , 2020), ilustrada en la Figura 35. Esta arquitectura se distingue de los autoencoders convencionales por la presencia de conexiones de salto (*skip connections*) entre el codificador y el decodificador, que permiten recuperar información de alta resolución perdida durante el submuestreo.

El codificador consta de cuatro etapas convolucionales sucesivas. Cada etapa aplica una capa *Conv1D* con kernel de tamaño 3, seguida de *BatchNorm1D*, activación *ReLU* y una operación *MaxPool1D* con factor 2. La progresión de canales es $1 \rightarrow 16 \rightarrow 32 \rightarrow 64 \rightarrow 128$, de modo que la resolución temporal se reduce a $L/16$ en el espacio latente. mientras la profundidad de características aumenta a 128 canales.

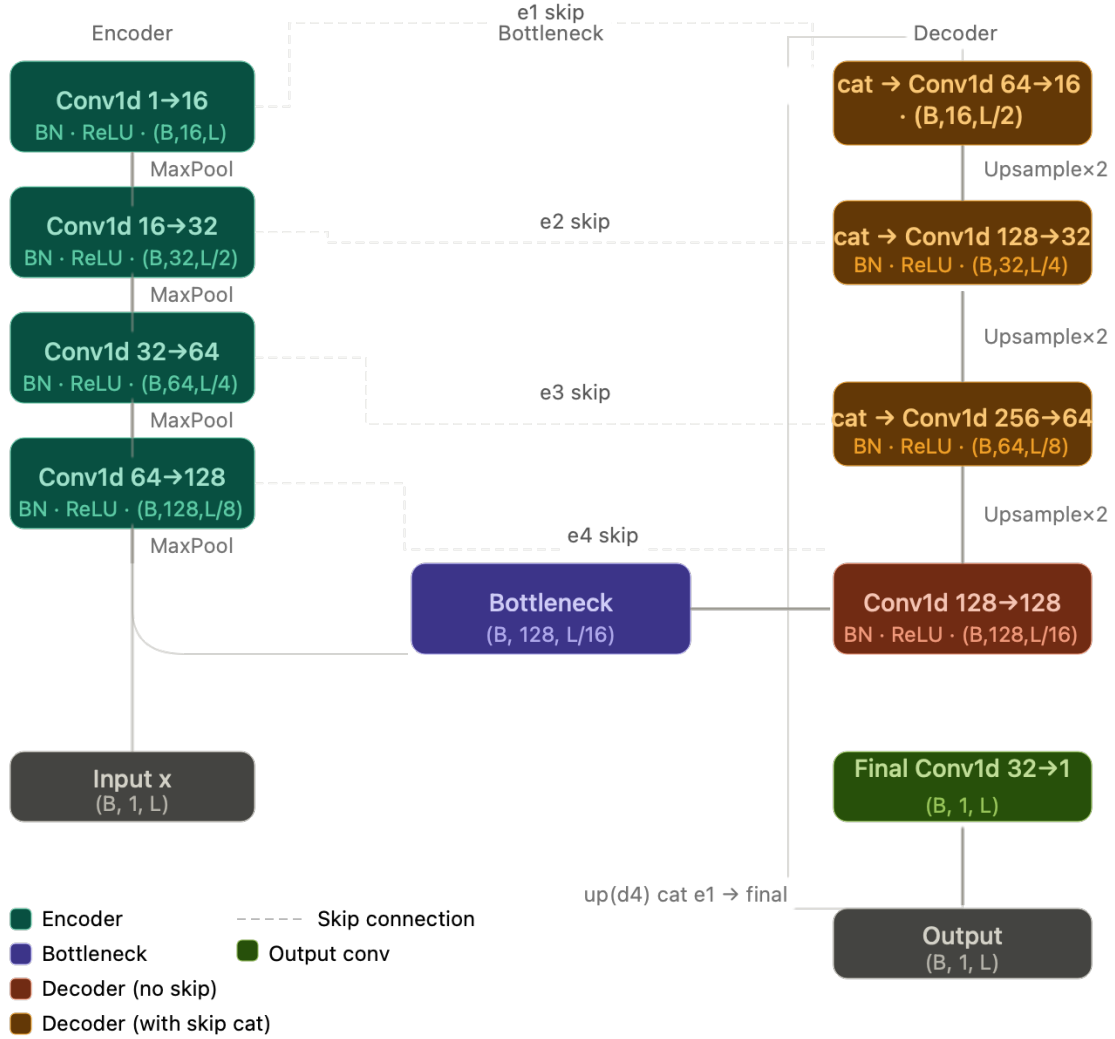


Figura 35: Arquitectura del Denoising Autoencoder implementado. El codificador (izquierda) reduce la resolución temporal mediante convoluciones y maxpooling; el decodificador (derecha) recupera la resolución original mediante upsampling con conexiones de salto desde el codificador.

El decodificador es simétrico al codificador y consta igualmente de cuatro etapas. Cada etapa aplica una operación de *Upsample* con factor 2 para recuperar la resolución temporal, seguida de una concatenación con el mapa de características correspondiente del codificador (conexión de salto) y una capa *Conv1D* para reducir los canales. La progresión de canales en el decodificador es $(128+128) \rightarrow 64 \rightarrow (64+64) \rightarrow 32 \rightarrow (32+32) \rightarrow 16 \rightarrow (16+16)$. Una capa convolucional final *Conv1D*(1) produce la señal reconstruida $\hat{h}(t)$ con la misma longitud temporal L que la entrada.



En el estudio de Shen et al. (2019), los autores destacan la efectividad de entrenar un DAE mediante Curriculum Learning (Bengio et al., 2009). Esta estrategia consiste en exponer progresivamente el modelo a señales con diferentes relaciones de SNR, comenzando con señales de alto SNR y gradualmente introduciendo señales más ruidosas. Este enfoque permite que el modelo aprenda de manera incremental a reconstruir eficazmente señales con bajo SNR, mejorando su robustez y capacidad de generalización.

Entrenamiento del autoencoder

El modelo fue implementado en *PyTorch*. Cada instancia de entrenamiento consiste en un par (X, y) , donde X es la señal contaminada con ruido y y la señal limpia de referencia. La normalización se aplicó mediante escalado min-máx por muestra, conforme a la ecuación (57).

$$x_{\text{norm}} = \frac{x - \text{mín}(x)}{\text{máx}(x) - \text{mín}(x)} \quad (57)$$

La función de pérdida combina dos términos complementarios:

$$\mathcal{L} = \mathcal{L}_{\text{Huber}} + 0,1 \cdot \mathcal{L}_{\text{espectral}} \quad (58)$$

donde $\mathcal{L}_{\text{espectral}} = \frac{1}{N} \sum_k ||\hat{H}[k]| - |H[k]||$ penaliza las diferencias en la amplitud del espectro de Fourier entre la señal reconstruida y la limpia, incentivando la fidelidad en el dominio frecuencial.

Se empleó la estrategia de *Curriculum Learning* (Bengio et al., 2009), comenzando el entrenamiento exclusivamente con muestras de alto SNR (fáciles) y expandiendo progresivamente el conjunto activo de entrenamiento mediante una función de ritmo de raíz cuadrada:

$$f_{\text{activa}}(e) = f_0 + (1 - f_0) \sqrt{\frac{e}{E}} \quad (59)$$

donde $f_0 = 0,1$ es la fracción inicial, e la época actual y E el número total de épocas. El optimizador utilizado es *AdamW* con tasa de aprendizaje $\alpha = 10^{-3}$, y se aplica *ReduceLROnPlateau* para ajuste adaptativo de la tasa. La Tabla 7 resume los hiperparámetros del DAE.

Cuadro 7: Hiperparámetros del DAE para eliminación de ruido.

Hiperparámetro	Valor
Etapas del codificador	4
Canales (codificador)	$1 \rightarrow 16 \rightarrow 32 \rightarrow 64 \rightarrow 128$
Conexiones de salto	Sí
Pérdida	Huber + $0,1 \times \text{espectral}$
Optimizador	AdamW
Tasa de aprendizaje	10^{-3}
Tamaño de batch	64
Curriculum Learning	Sí ($f_0 = 0,1$)

Retomando la señal de la Figura 33, cuya descomposición wavelet y análisis de diferencias de energía no permitió identificar la región de interés, aplicamos ahora el proceso de reconstrucción mediante el DAE. El resultado, mostrado en la Figura 36, evidencia una notable reducción del ruido presente en la señal original.

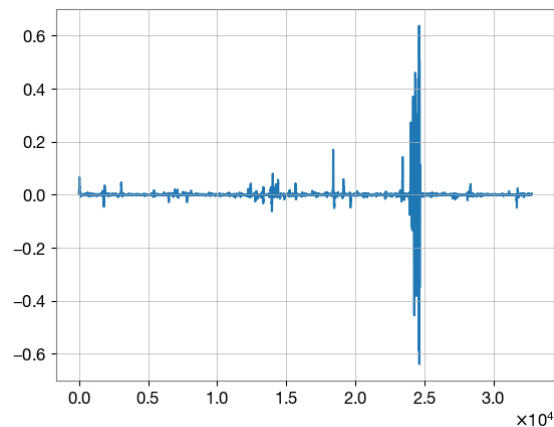
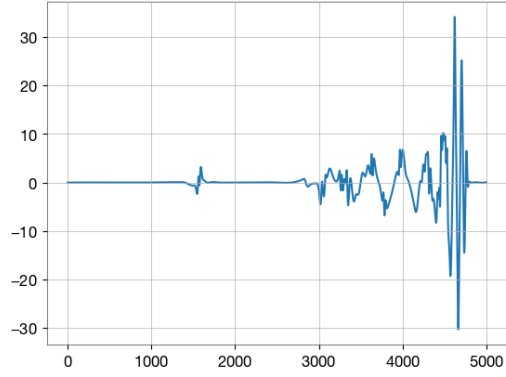


Figura 36: Reconstrucción de la señal de la Figura 33 mediante el DAE propuesto.

Al aplicar el procedimiento de descomposición wavelet y análisis de diferencias de energía a la señal reconstruida por el DAE, se obtienen los resultados mostrados en la Figura 37. La comparación entre la señal procesada por el DAE (Figura 37a) y la señal limpia original (Figura 37b) revela una similitud notable en la identificación de la zona de interés.

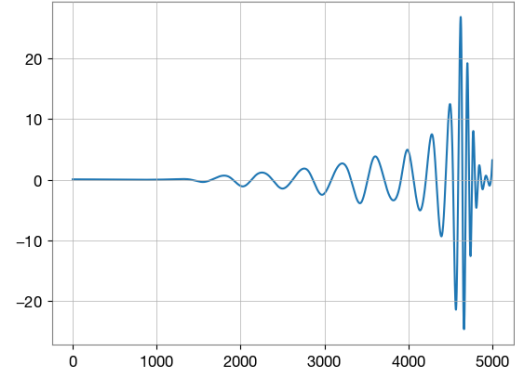


Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(a) Análisis aplicado a la señal reconstruida.

Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(b) Análisis aplicado a la señal limpia original.

Figura 37: Comparación del análisis de diferencias de energía entre la señal reconstruida por el DAE y la señal original sin ruido.

Estos resultados demuestran cuantitativamente la efectividad del DAE como herramienta para la supresión de ruido en señales de ondas gravitacionales. La reconstrucción generada se aproxima significativamente a la señal original sin ruido, permitiendo una identificación precisa de la zona de interés que no era posible con la señal contaminada original.

Modelo CNN para la regresión de parámetros

La estimación de parámetros físicos del sistema binario constituye la etapa final del pipeline. Dado que la masa del chirp $\mathcal{M}_c$ definida en la ecuación (5) condensa la mayor parte de la información de fase de la señal en la banda de frecuencias observada por LIGO (Flanagan & Hughes , 2005) (Cutler & Flanagan , 1994), se tomó como variable objetivo de la regresión en lugar de las masas individuales.

El modelo implementado, ilustrado en la Figura 38, es una red neuronal convolucional pura que opera directamente sobre la región de interés de 5000 puntos extraída por el análisis wavelet-energía de la señal reconstruida.

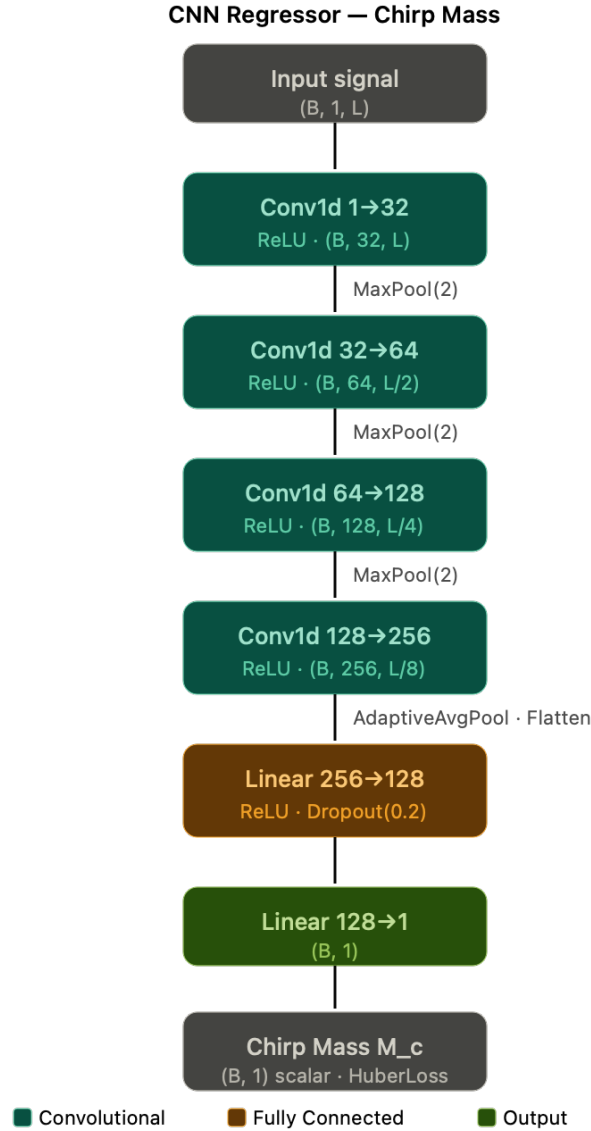


Figura 38: Arquitectura del regresor CNN para estimación de la masa del chirp $\mathcal{M}_c$.

La señal de entrada (un solo canal) pasa por cuatro bloques convolucionales sucesivos. Cada bloque aplica una capa *Conv1D* con kernel de tamaño 3, seguida de activación *ReLU* y *MaxPool1D(2)* (excepto el último bloque). La progresión de canales es $1 \rightarrow 32 \rightarrow 64 \rightarrow 128 \rightarrow 256$, de modo que la señal se comprime espacialmente mientras se incrementa la profundidad de características.

Tras el último bloque convolucional, se aplica *Global Average Pooling* (*AdaptiveAvgPool1D(1)*) para colapsar la dimensión temporal y obtener un vector de 256 características independiente de la longitud de la entrada. Este vector alimenta una cabeza de regresión compuesta por una capa



completamente conectada $256 \rightarrow 128$ con activación *ReLU* y *Dropout*(0,2), seguida de una capa final $128 \rightarrow 1$ que produce la estimación escalar de $\mathcal{M}_c$ en unidades de $M_\odot$.

Entrenamiento del modelo CNN

El conjunto de datos se dividió en subconjuntos de entrenamiento (80%) y prueba (20%). Como variable característica (X) se emplearon las regiones de interés $Z(t)$ (5000 puntos) extraídas mediante la transformada wavelet y el análisis de diferencias de energía aplicados a las señales limpias. La variable objetivo (y) es la masa del chirp $\mathcal{M}_c$ calculada según la ecuación (5) a partir de las masas de cada evento.

Las señales de entrada se normalizaron mediante la ecuación (57) y los valores de $\mathcal{M}_c$ se estandarizaron con *StandardScaler* ajustado sobre el conjunto de entrenamiento, para evitar fuga de información.

La implementación en PyTorch utilizó la función de pérdida de Huber con $\delta = 1,5$ y el optimizador *NAdam* con tasa de aprendizaje $\alpha = 10^{-3}$. El entrenamiento se condujo durante 200 épocas con tamaño de batch de 128 muestras.



Resultados

Para evaluar el desempeño de la metodología completa se seleccionaron tres eventos con condiciones contrastantes: un caso extremo con ruido dominante sobre la señal gravitacional, un segundo escenario en el que el ruido sigue siendo dominante pero con masas considerablemente mayores, y un tercer caso con bajo nivel de ruido en el que la señal es claramente distinguible.

Estos tres escenarios permiten describir de forma sistemática el comportamiento de la metodología a lo largo de todo su flujo, desde la generación de la base de datos hasta el entrenamiento de las redes neuronales.

Simulación de las ondas

Las señales gravitacionales son generadas mediante *Bilby* (Ashton et al., 2019), configurando los parámetros de cada evento de forma independiente. El evento 1 corresponde al caso extremo, con dos masas iguales de $10M_{\odot}$ y un SNR de 10 dB. El evento 2 aumenta la masa de los objetos a $30M_{\odot}$ y $60M_{\odot}$, manteniendo el mismo SNR de 10 dB. El evento 3 considera las masas $30M_{\odot}$ y $60M_{\odot}$, con un SNR de 20 dB.

Parámetro	Valores	Unidad
M_1, M_2	$(10), (10)$	$M_\odot$
χ_1, χ_2	$(0), (0)$	—
θ_1, θ_2	$(0,5), (1)$	rad
$\delta\phi, \phi_{JL}$	$(1,7), (0,3)$	rad
d_L	500	Mpc
ra	1,375	rad
dec	−1,2108	rad
ι	0,4	rad
ψ	0,6	rad
ϕ_c	1,3	rad
SNR	10	dB

Cuadro 8: Valores para la simulación del evento 1.

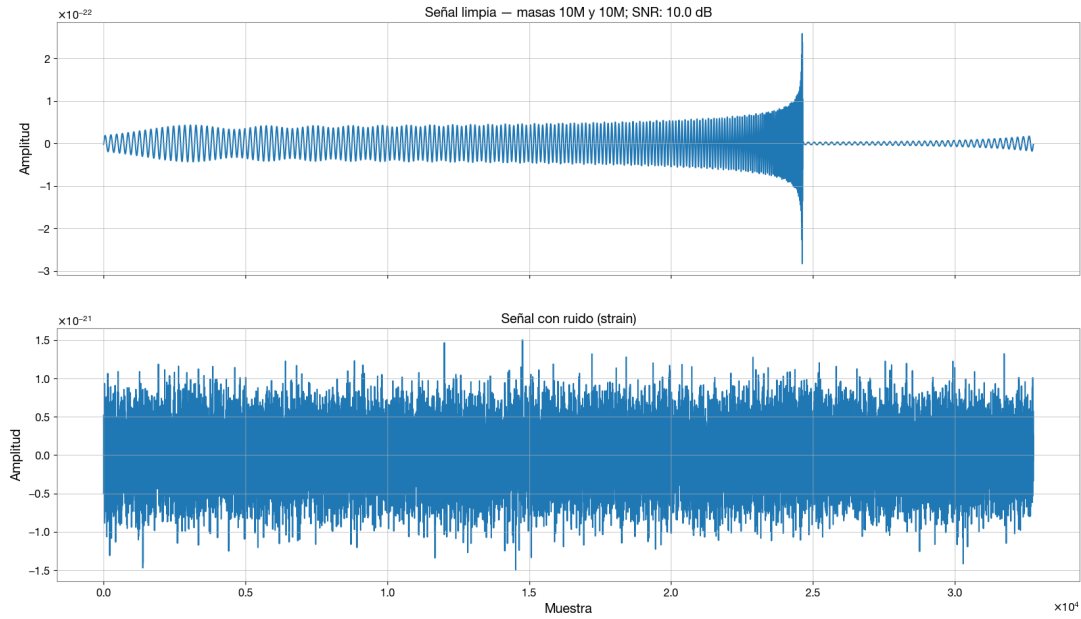


Figura 39: Simulación de una GW para el evento 1.



Parámetro	Valores	Unidad
M_1, M_2	(30), (60)	$M_\odot$
χ_1, χ_2	(0), (0)	—
θ_1, θ_2	(0,5), (1)	rad
$\delta\phi, \phi_{JL}$	(1,7), (0,3)	rad
d_L	500	Mpc
ra	1,375	rad
dec	-1,2108	rad
ι	0,4	rad
ψ	0,6	rad
ϕ_c	1,3	rad
SNR	10	dB

Cuadro 9: Valores para la simulación del evento 2.

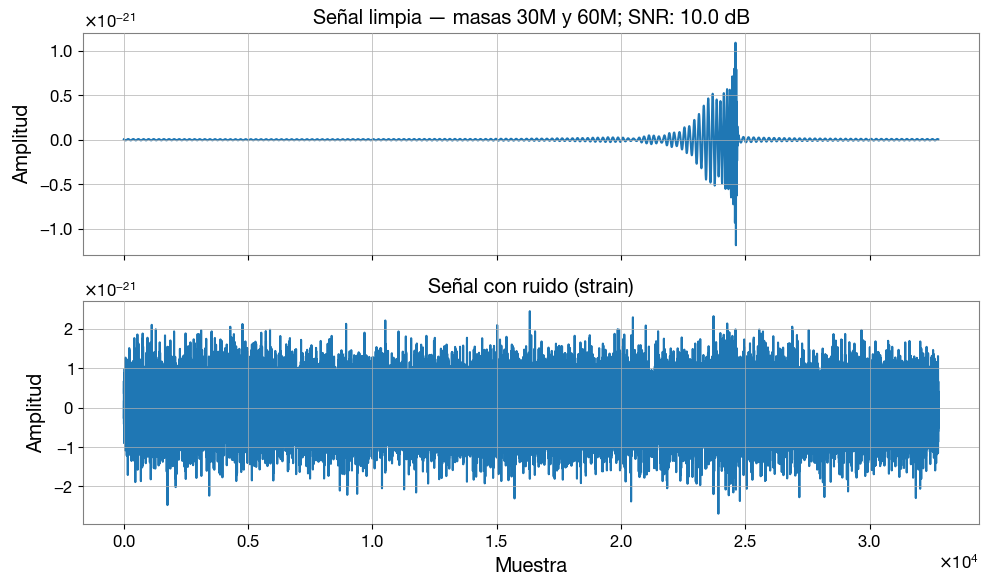


Figura 40: Simulación de una GW para el evento 2.



Parámetro	Valores	Unidad
M_1, M_2	(30), (60)	$M_\odot$
χ_1, χ_2	(0), (0)	—
θ_1, θ_2	(0,5), (1)	rad
$\delta\phi, \phi_{JL}$	(1,7), (0,3)	rad
d_L	500	Mpc
ra	1,375	rad
dec	-1,2108	rad
ι	0,4	rad
ψ	0,6	rad
ϕ_c	1,3	rad
SNR	20	dB

Cuadro 10: Valores para la simulación del evento 3.

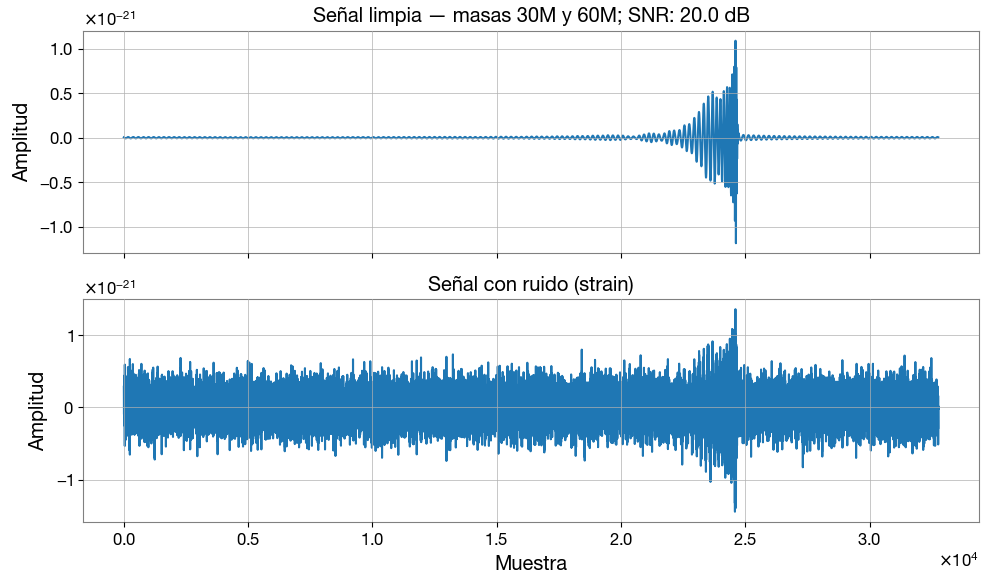


Figura 41: Simulación de una GW para el evento 3.

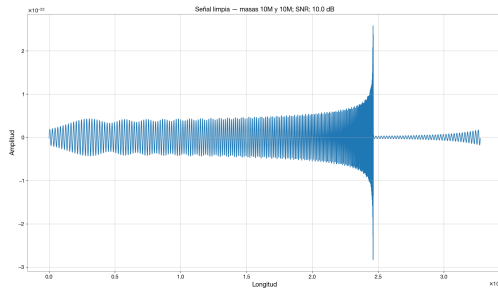
De acuerdo con la Tabla 4, la configuración final de la base de datos incluye los parámetros descritos en las Tablas 8, 9 y 10, así como tres listas de señales: la señal limpia $[h(t_0), \dots, h(t_N)]$, la señal contaminada $[n(t_0), \dots, n(t_N)]$, y una versión procesada de la señal limpia obtenida mediante la transformada de wavelets y la diferencia de energías, empleada para delimitar la región de interés del evento ($[z(t_0), \dots, z(t_N)]$).

Análisis preliminar para la creación de la base de datos

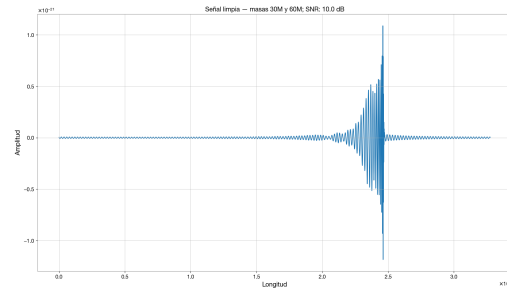
Para identificar la región de interés de cada señal simulada, se descompone la señal en sus componentes de frecuencia mediante la transformada de wavelets y se calcula la diferencia de energías entre niveles, con el fin de delimitar la región de interés en la que se concentra el evento gravitacional.

Este análisis es fundamental, dado que las señales resultantes constituyen la entrada de los modelos de *machine learning* entrenados en etapas posteriores.

Transformada de Wavelet



(a) Evento 1.



(b) Eventos 2 y 3.

Figura 42: Señales gravitacionales simuladas sin ruido para los tres eventos.

Las figuras 42a y 42b muestran la señal limpia de los tres eventos. Cabe señalar que la única diferencia entre los eventos 2 y 3 es la intensidad del ruido añadido.



Al aplicar la transformada de wavelets a la señal, se obtiene una descomposición por niveles como la ilustrada en la Figura 43, donde cada nivel corresponde a una banda de frecuencia distinta.

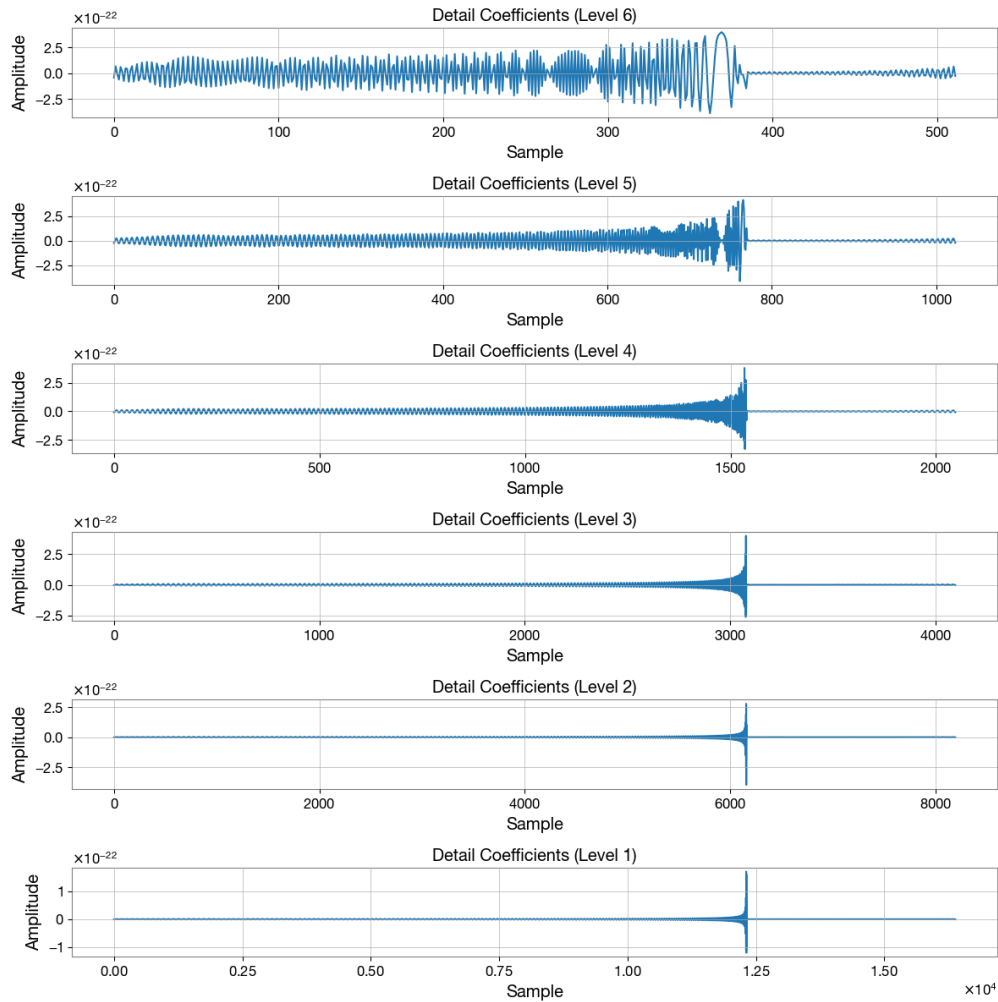
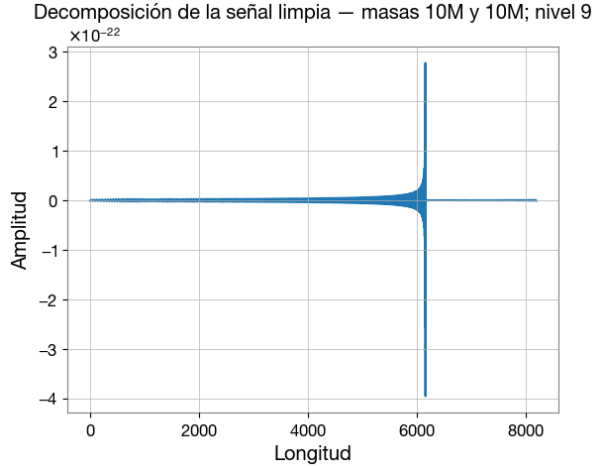
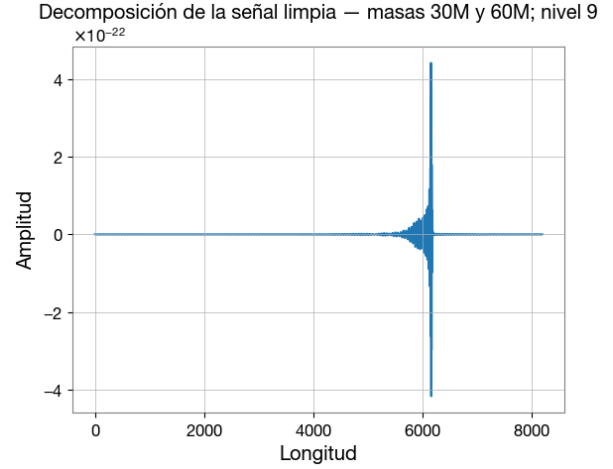


Figura 43: Descomposición de wavelet de la señal limpia del evento 1.

Dado que el evento gravitacional se concentra en una banda específica, se selecciona la descomposición de nivel 9 para el análisis, como se muestra en la Figura 44.



(a) Evento 1.

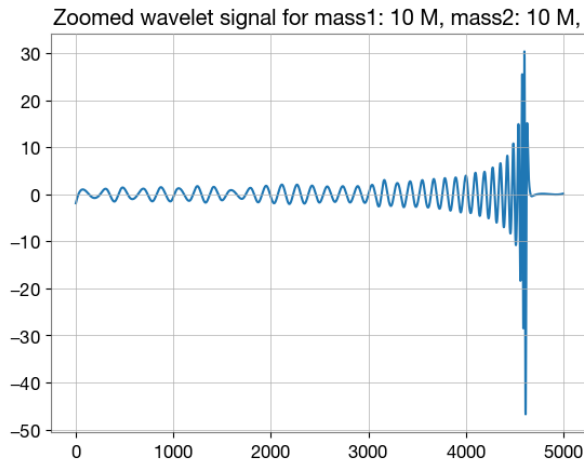


(b) Eventos 2 y 3.

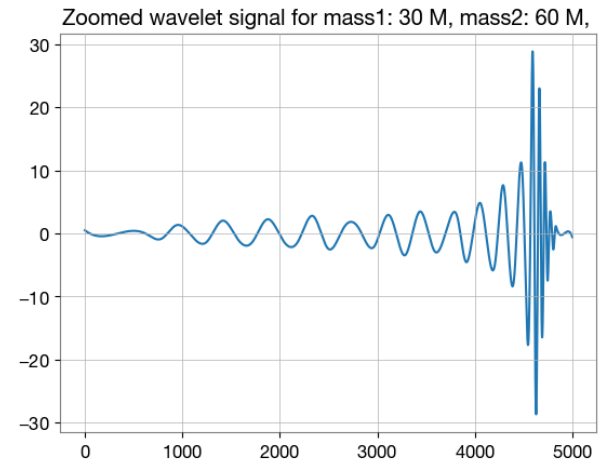
Figura 44: Descomposición nivel 9 de la transformada de wavelets de la señal limpia.

Diferencia de energía

A partir de la descomposición de nivel 9, se aplica la diferencia de energías definida en la ecuación (49) para localizar la región temporal en la que se concentra el evento gravitacional.



(a) Evento 1.



(b) Eventos 2 y 3.

Figura 45: Diferencia de energías de la descomposición nivel 9 de la transformada de wavelets.

La señal resultante, ilustrada en la Figura 45, corresponde a la diferencia de energías $[z(t_0), \dots, z(t_N)]$.

Esta señal se incorpora a la base de datos final y constituye una de las entradas para el entrena-



miento de los modelos de *machine learning*.

Este proceso se repite para todos los eventos simulados, obteniendo de esta manera, la base de datos final que se observa en la tabla 4.

Análisis de la señal

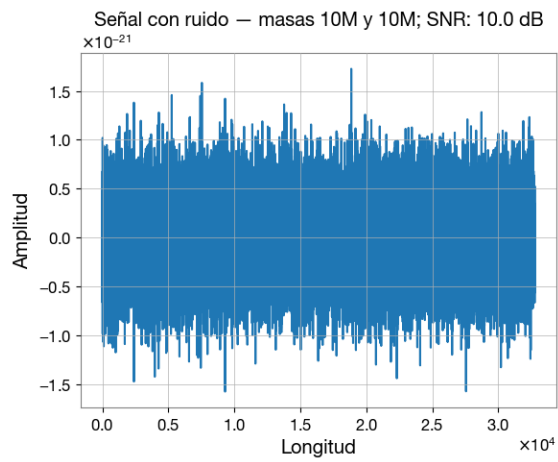
Con la base de datos final disponible, se procede al análisis de las señales. Esta etapa consiste en procesar cada señal para extraer un vector de características que permita discriminar los eventos gravitacionales del ruido de fondo, el cual se utiliza posteriormente para entrenar un modelo de *machine learning* orientado a la clasificación.

Cadena de filtrado FIR

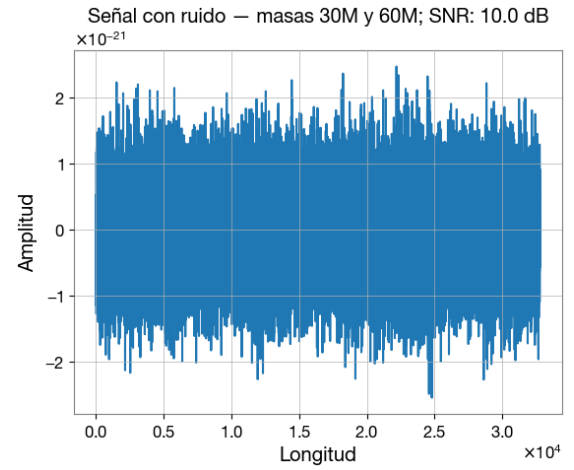
El análisis parte de la señal contaminada con ruido, ilustrada en las Figuras 46a, 46b y 46c.

Para discriminar los eventos gravitacionales del ruido de fondo, se extraen como características la energía de la señal y su desviación estándar. Tal como se describió en la sección de metodología, la cadena de filtrado comprende la aplicación sucesiva de un filtro FIR, una descomposición de wavelet y un segundo filtro FIR.

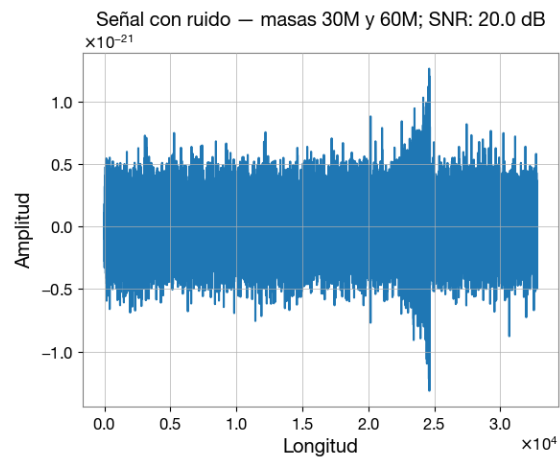
En las Figuras 47, 48 y 49, se puede observar el comportamiento de la cadena de filtrado para los tres eventos.



(a) Evento 1.



(b) Evento 2.



(c) Evento 3.

Figura 46: Señales gravitacionales simuladas con ruido para los tres eventos.

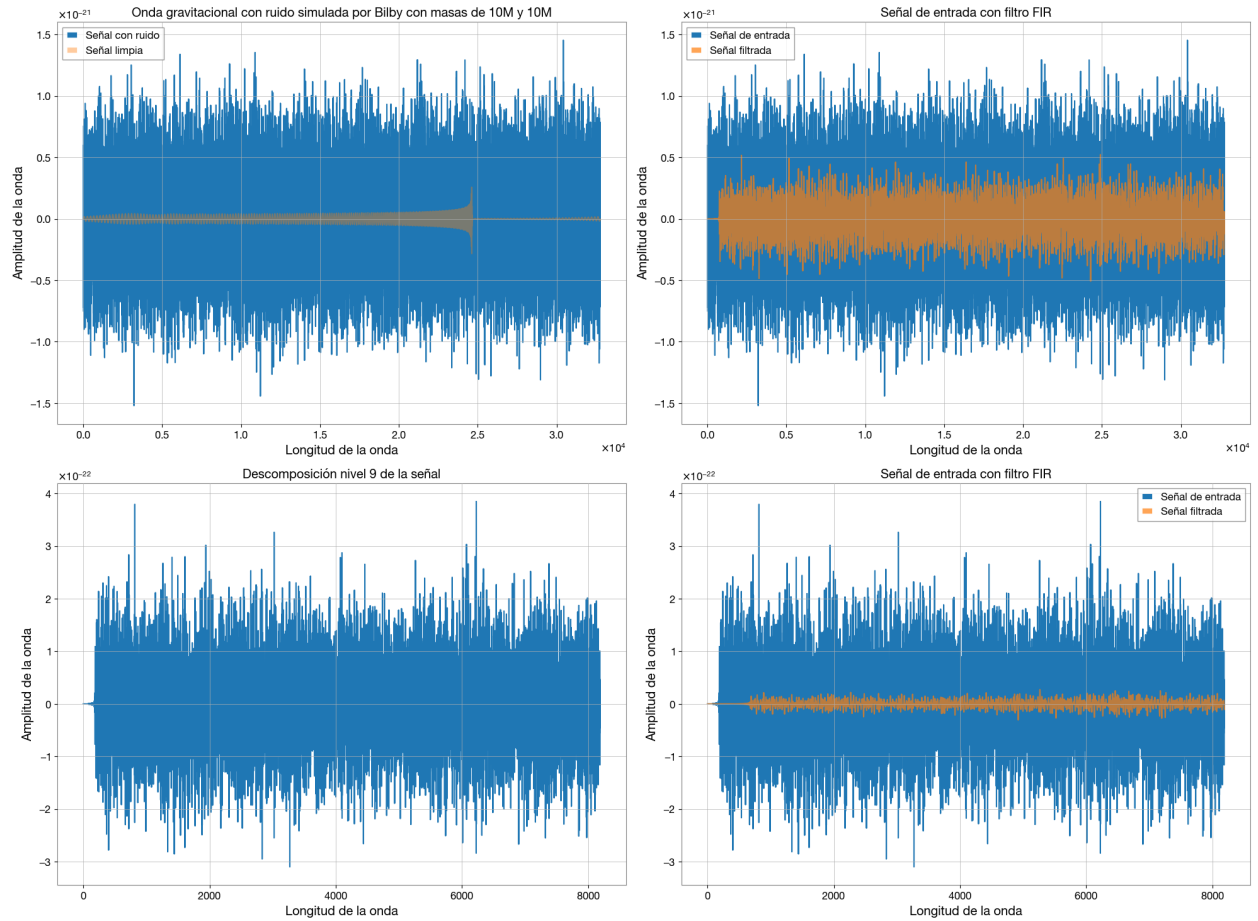


Figura 47: Cadena de filtrado del evento 1. En la primera gráfica se observa la señal limpia (naranja) y contaminada (azul). En la segunda gráfica podemos observar el efecto del primer filtro FIR (naranja), sobre la señal con ruido. En la tercera gráfica se muestra la transformada de wavelet nivel 9 para la señal filtrada un paso anterior. En la cuarta gráfica podemos observar el efecto del segundo filtro FIR (naranja) sobre la transformada de wavelet.

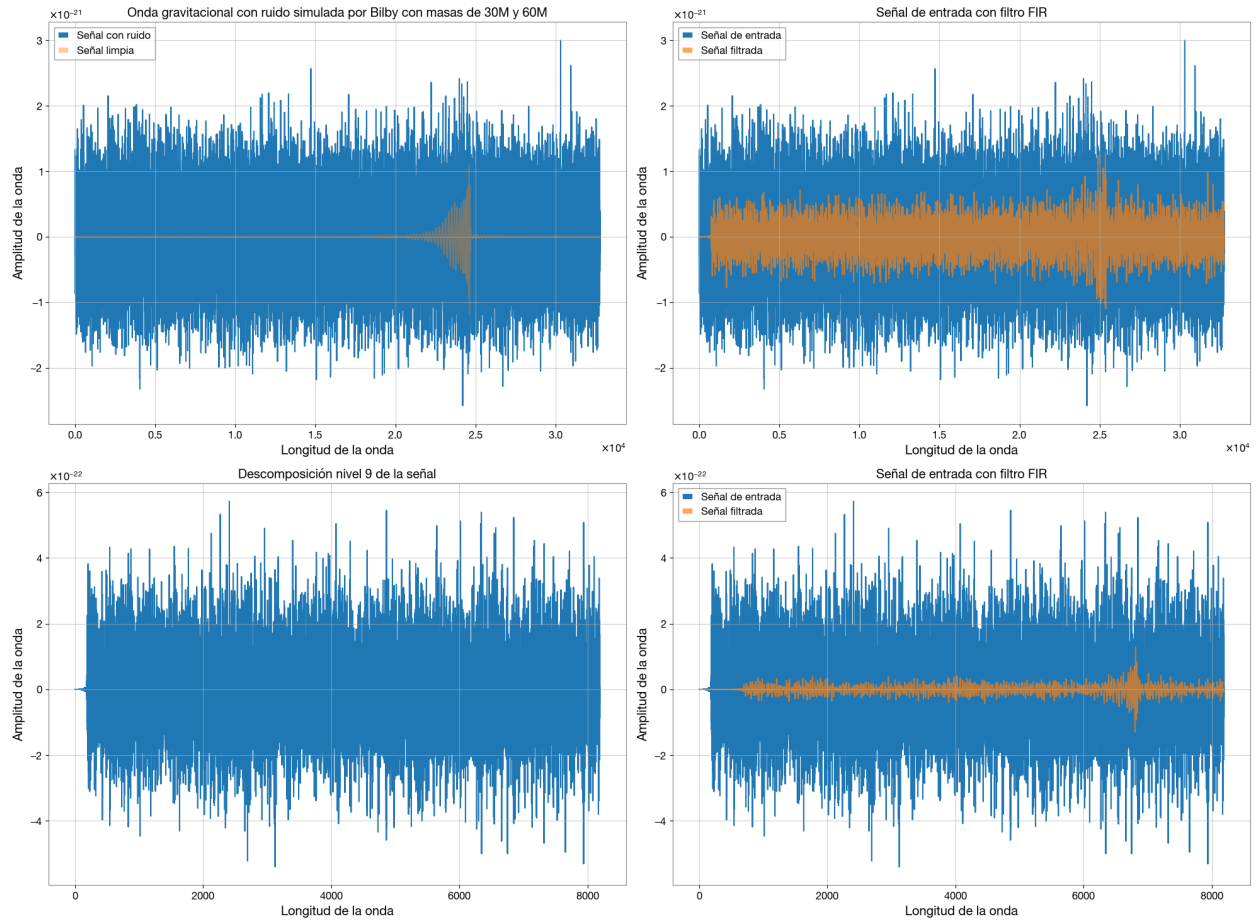


Figura 48: Cadena de filtrado del evento 2.

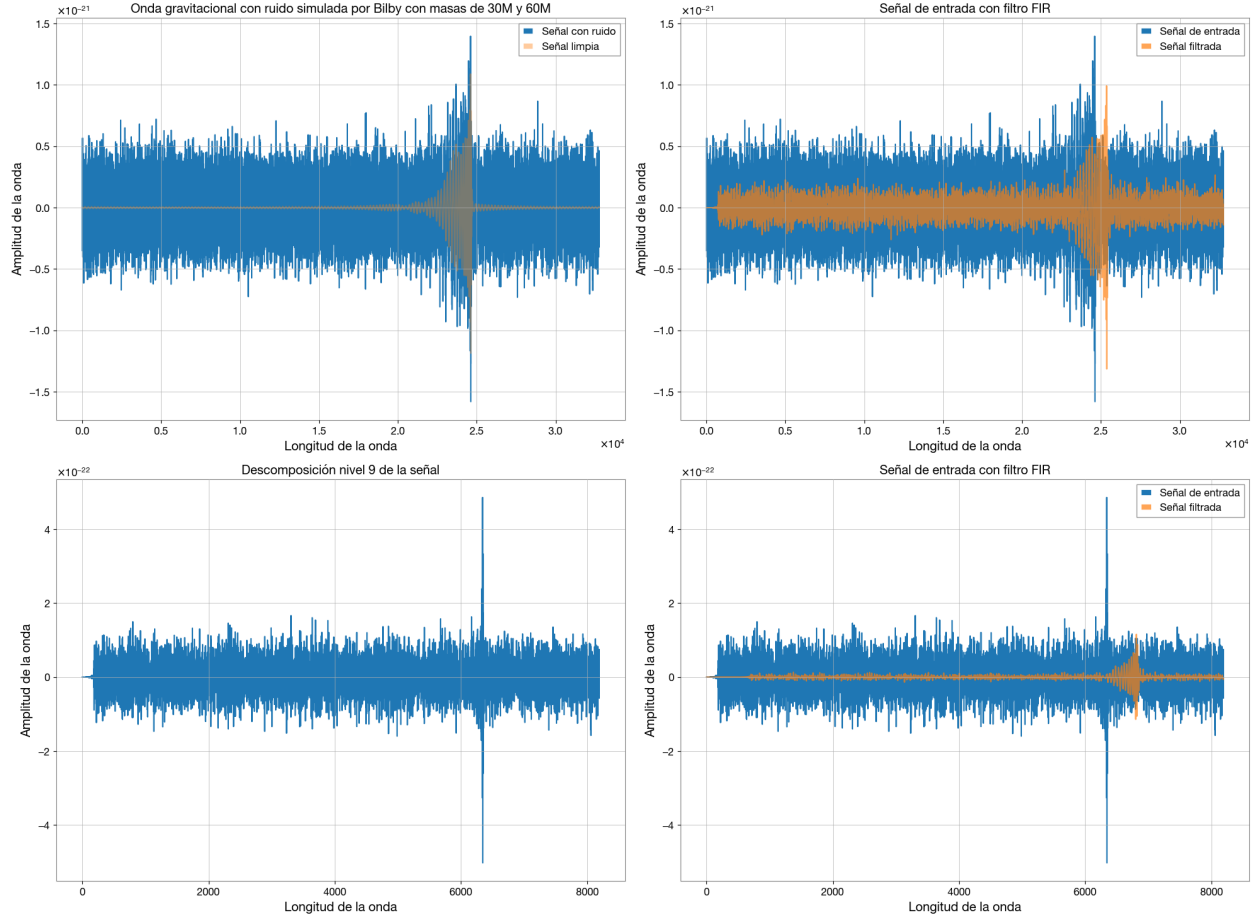


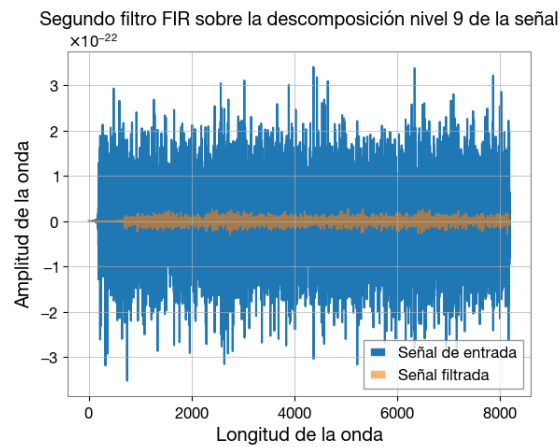
Figura 49: Cadena de filtrado del evento 3.

Extracción de características para detección

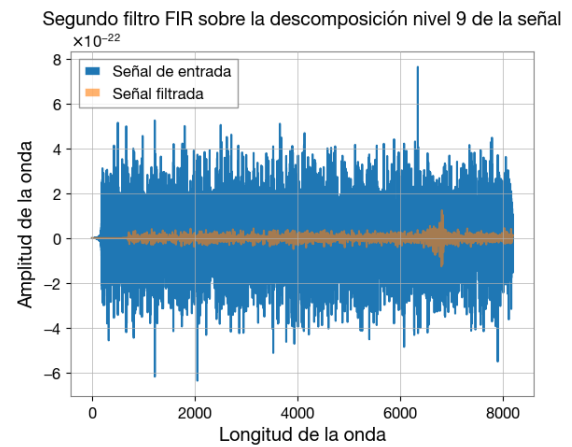
Para extraer las características relevantes de la señal, tomaremos la señal resultante del segundo filtro FIR, como se observa en la Figura 50. y calcularemos su energía, así como su desviación estándar. La finalidad de este proceso es la de discriminar los eventos gravitacionales del ruido de fondo.

Energía.

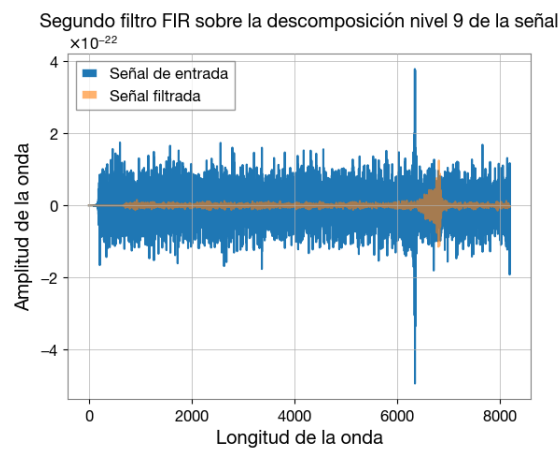
Para calcular la energía de la señal, vamos a segmentar la señal en ventanas de tamaño w_s y calcular la energía de cada segmento con la ecuación 51. Obteniendo de esta manera la energía de la señal resultante, ilustrada en la Figura 51.



(a) Evento 1.

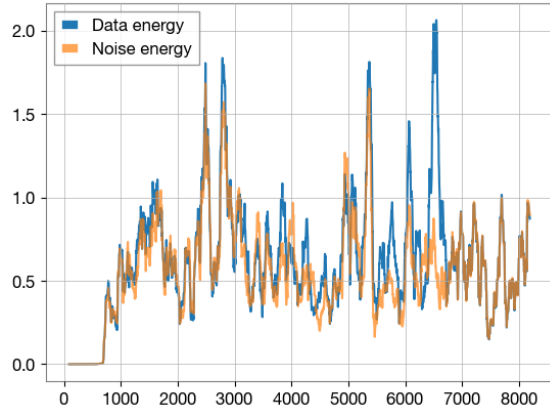


(b) Evento 2.

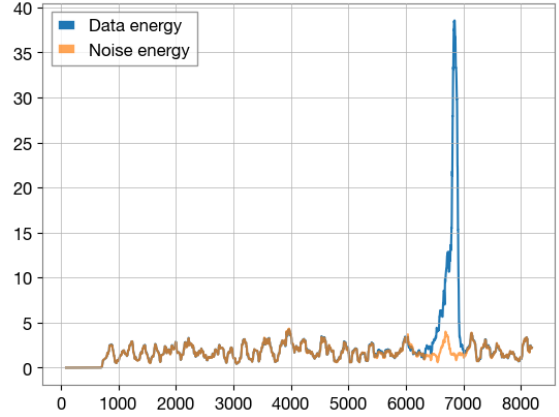


(c) Evento 3.

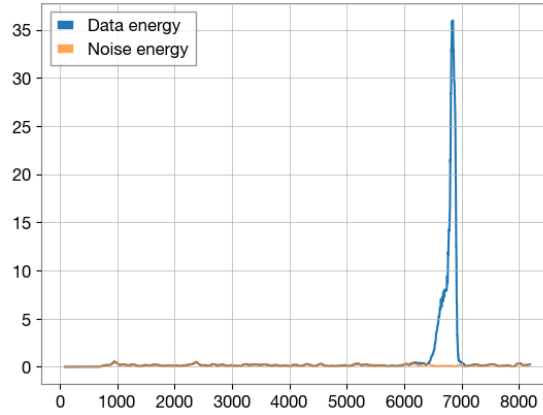
Figura 50: Señales resultantes de la cadena de filtrado para los tres eventos.



(a) Evento 1.



(b) Evento 2.



(c) Evento 3.

Figura 51: Energías de las señales resultantes de la cadena de filtrado para los tres eventos.

Podemos observar como la energía de la señal demuestra ser un buen indicador de la presencia de un evento gravitacional. De la misma manera, este resultado nos da información muy importante. Recordemos que las señales que originan al evento 1 y el evento 2, tienen un mismo valor de ruido, o mismo SNR, pero sus masas son distintas. Como podemos ver en la figura 46, las señales de los eventos 1 y 2 son muy similares, pero al momento de procesar las señales, observamos que el evento 1 es más difícil de distinguir (figura 51a) que el evento 2 (figura 51b).

Esto podría ser debido a que la mayoría de la información de la señal se encuentra en la masa del chirp, es decir, la onda gravitacional solo va a depender de la combinación de las masas de los

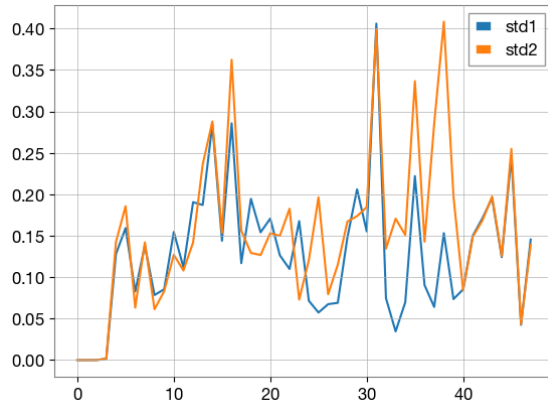


objetos que la generan, como demostraron Cutler & Flanagan (1994). De esta manera, podemos esperar que mientras mayor sea la masa del chirp, mayor será la energía de la señal y su detección será más fácil.

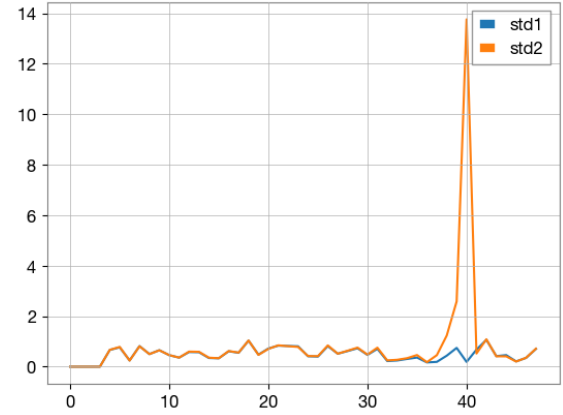
Desviación estándar

Una vez obtenida la energía de la señal, podemos calcular su desviación estándar mediante la ecuación 52. Al igual que el cálculo de la energía, vamos a segmentar la señal en ventanas de tamaño W y calcular la desviación estándar de cada segmento.

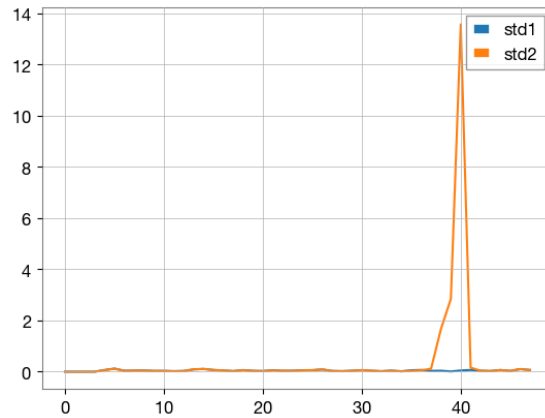
Este paso tiene dos objetivos principales, el primero es determinar donde se encuentran los cambios principales de la energía, para casos como el de la figura 51a, donde la energía de la señal con información es similar a la energía de la señal con ruido. El segundo objetivo es la reducción del tamaño de la señal resultante, para de esta manera reducir el tiempo de entrenamiento del clasificador.



(a) Evento 1.



(b) Evento 2.



(c) Evento 3.

Figura 52: Desviación estándar de las señales resultantes de la cadena de filtrado para los tres eventos. Para este ejemplo se asume que el tamaño de ventana es $W = 48$, pero como vamos a ver más adelante, este tamaño se va a modificar para obtener el valor óptimo.



Optimización conjunta del tamaño de ventana mediante PSO

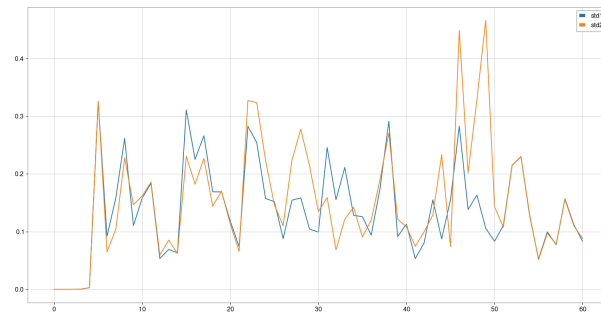
Como se mencionó previamente, el cálculo de la energía de la señal y la desviación estándar involucra el uso de ventanas de diferentes tamaños, cuyos valores influyen directamente en el resultado final de la detección, tal como se discutió en la metodología.

Con el fin de determinar de manera óptima estos parámetros, se implementa un algoritmo de PSO que busca los tamaños de ventana w_s y W que maximicen la desviación estándar de la señal de detección.

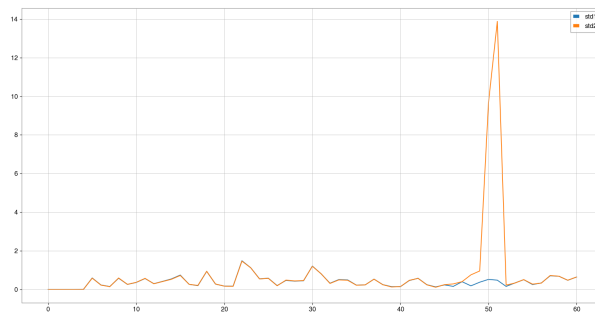
El algoritmo PSO se entrenó con los parámetros especificados en la Tabla 5, utilizando tres señales seleccionadas aleatoriamente como conjunto de entrenamiento. Para cada señal, el algoritmo explora el espacio de búsqueda y determina los tamaños de ventana óptimos w_s y W que maximizan la desviación estándar de la señal de detección. Dado que los valores óptimos pueden variar entre señales, el tamaño final de cada ventana se estima como la media aritmética de los resultados obtenidos en las tres ejecuciones.

Como resultado del proceso de optimización, se obtienen los valores $w_s = 97$ y $W = 61$. Estos tamaños de ventana maximizan la separación entre regiones de actividad y ruido en la señal. El resultado obtenido es mostrado en la figura 53.

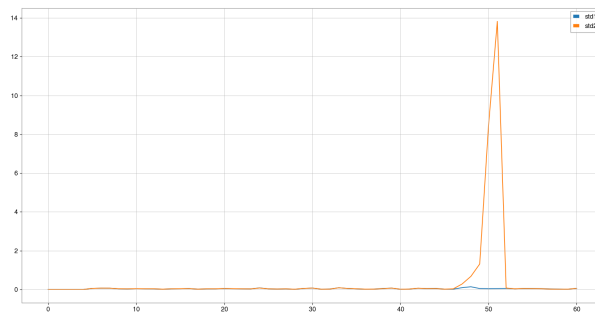
Con los tamaños de ventana óptimos determinados, se construye una nueva base de datos que contiene, para cada registro, la desviación estándar de la señal calculada con dichos parámetros, la masa del objeto que la genera, el SNR y una etiqueta binaria que indica si el registro corresponde a una señal gravitacional o a ruido. Esta base de datos es utilizada para el entrenamiento del clasificador.



(a) Evento 1.



(b) Evento 2.



(c) Evento 3.

Figura 53: Desviación estándar de las señales resultantes de la cadena de filtrado para los tres eventos. En esta figura, podemos ver el efecto de la optimización conjunta del tamaño de ventana.



Clasificador LSTM

La base de datos empleada para el entrenamiento del clasificador se describe en la Tabla 11. Con estos datos y los hiperparámetros definidos en la Tabla 6, se entrena el clasificador LSTM.

Parámetro	Rango/Valor	Unidad
Desviación estándar	$[\sigma_0, \dots, \sigma_N]$	-
Masa primaria (M_1)	$[10, 100]$	$M_\odot$
Masa secundaria (M_2)	$[10, 100]$	$M_\odot$
Relación señal-ruido (SNR)	$[10, 26]$	dB
Etiqueta binaria	$[0, 1]$	-

Cuadro 11: Base de datos para el entrenamiento del clasificador LSTM.

La base de datos consiste en 19600 muestras, con un balance de 9800 muestras de ruido y 9800 muestras de señales gravitacionales. Los datos se dividen en 70 % de entrenamiento, 15 % de validación y 15 % de prueba. El entrenamiento se realiza en batches de 32 muestras, con una tasa de aprendizaje de 10^{-3} y 150 épocas.

Cuadro 12: Reporte de clasificación del modelo LSTM en el conjunto de prueba.

Clase	Precisión	Recall	F1-score	Soporte
Ruido (0)	0.99	1.00	1.00	1470
Señal GW (1)	1.00	0.99	1.00	1470
Accuracy		1.00		2940
Promedio macro	1.00	1.00	1.00	2940
Promedio ponderado	1.00	1.00	1.00	2940
ROC-AUC: 0,9999				

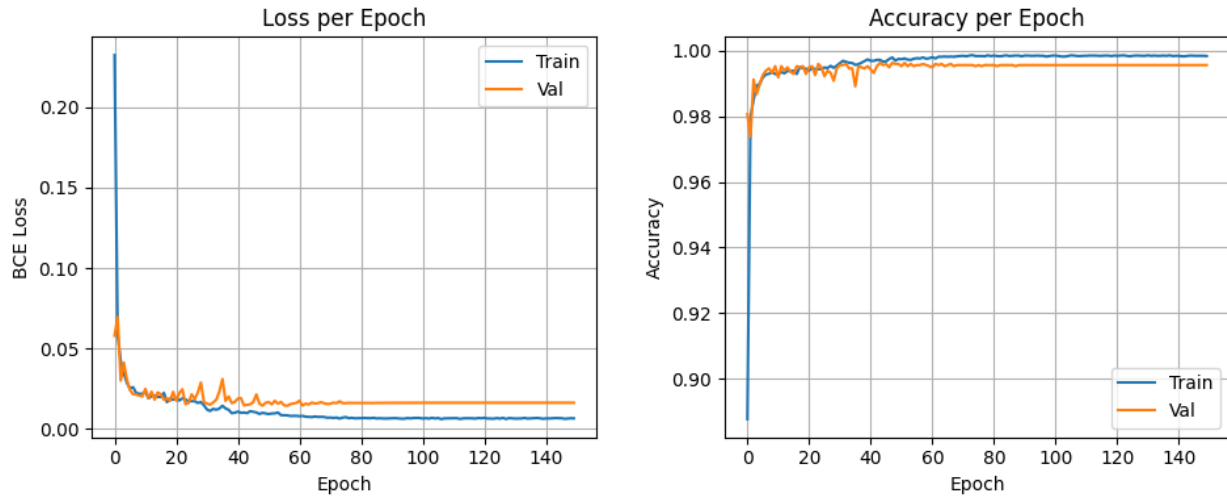


Figura 54: Gráfica de pérdida y precisión por época durante el entrenamiento.

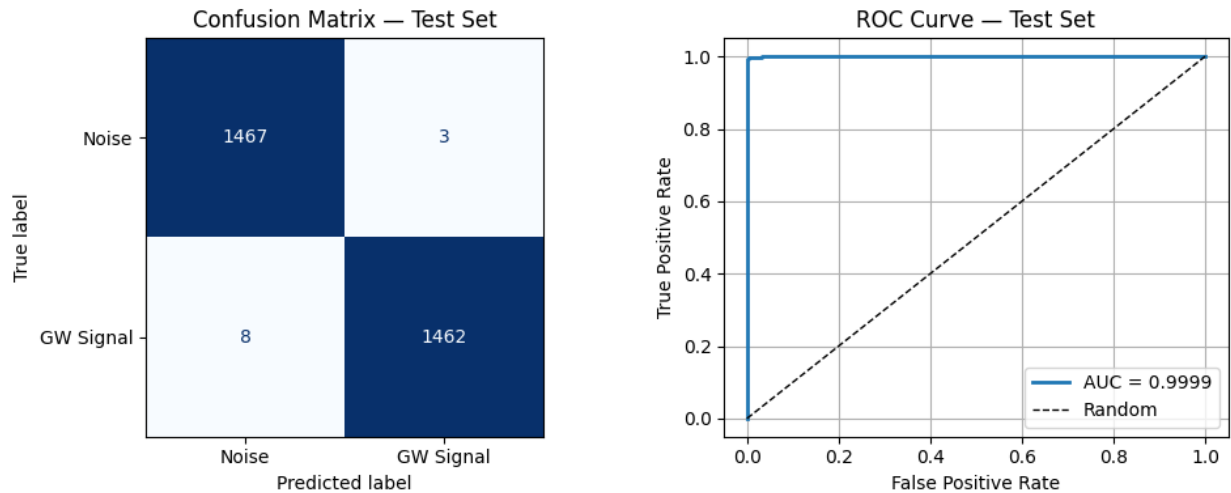


Figura 55: Matriz de confusión del clasificador LSTM en el conjunto de prueba y curva ROC.

La Tabla 12 presenta el reporte de clasificación del modelo LSTM evaluado sobre el conjunto de prueba, compuesto por 2,940 muestras balanceadas. El clasificador alcanza una exactitud global del 100%, con valores de F1-score de 1,00 para ambas clases. Para la clase de ruido, se obtiene una precisión de 0,99 y un recall de 1,00, lo que indica que prácticamente todas las muestras de ruido son correctamente identificadas, con una mínima tasa de falsos negativos. Para la clase de señal gravitacional, se obtiene una precisión de 1,00 y un recall de 0,99, lo que significa que casi la totalidad de las señales reales son detectadas sin generar falsos positivos. Complementariamente, el valor de ROC-AUC de 0,9999 confirma la capacidad discriminativa del modelo, situándose



prácticamente en el máximo posible.

Estos resultados se ilustran visualmente en la Figura 54, que muestra la evolución de la pérdida y la exactitud durante el entrenamiento, y en la Figura 55, que presenta la matriz de confusión y la curva ROC del modelo sobre el conjunto de prueba.

Cuadro 13: Resultado de la detección para la señal del evento 1.

Parámetro / Resultado	Valor
Masa 1 ($M_{\odot}$)	10
Masa 2 ($M_{\odot}$)	10
SNR	10
Detección (umbral)	No
Probabilidad de señal GW	0.1402
Clase predicha	Ruido

Cuadro 14: Resultado de la detección para la señal del evento 2.

Parámetro / Resultado	Valor
Masa 1 ($M_{\odot}$)	30
Masa 2 ($M_{\odot}$)	60
SNR	10
Detección (umbral)	Sí
Probabilidad de señal GW	1.0
Clase predicha	Señal GW

Las Tablas 13, 14 y 15 presentan los resultados de la detección para los eventos 1, 2 y 3, respectivamente. El clasificador identifica correctamente las señales gravitacionales en los eventos 2 y 3, mientras que el evento 1 es clasificado como ruido.



Cuadro 15: Resultado de la detección para la señal del evento 3.

Parámetro / Resultado	Valor
Masa 1 ($M_{\odot}$)	30
Masa 2 ($M_{\odot}$)	60
SNR	20
Detección (umbral)	Sí
Probabilidad de señal GW	1.0
Clase predicha	Señal GW

Adicionalmente, se analiza el efecto del tamaño de la ventana de entrada sobre el rendimiento del clasificador LSTM, con el fin de evaluar la influencia de este parámetro en el proceso de optimización mediante PSO.

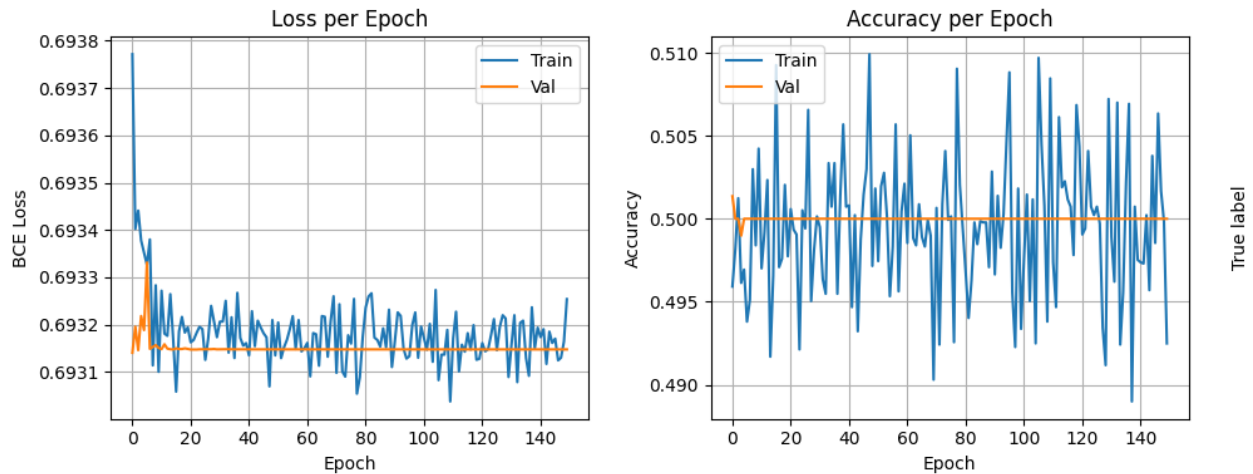


Figura 56: Gráfica de pérdida y precisión por época durante el entrenamiento para datos de entrada no optimizados.

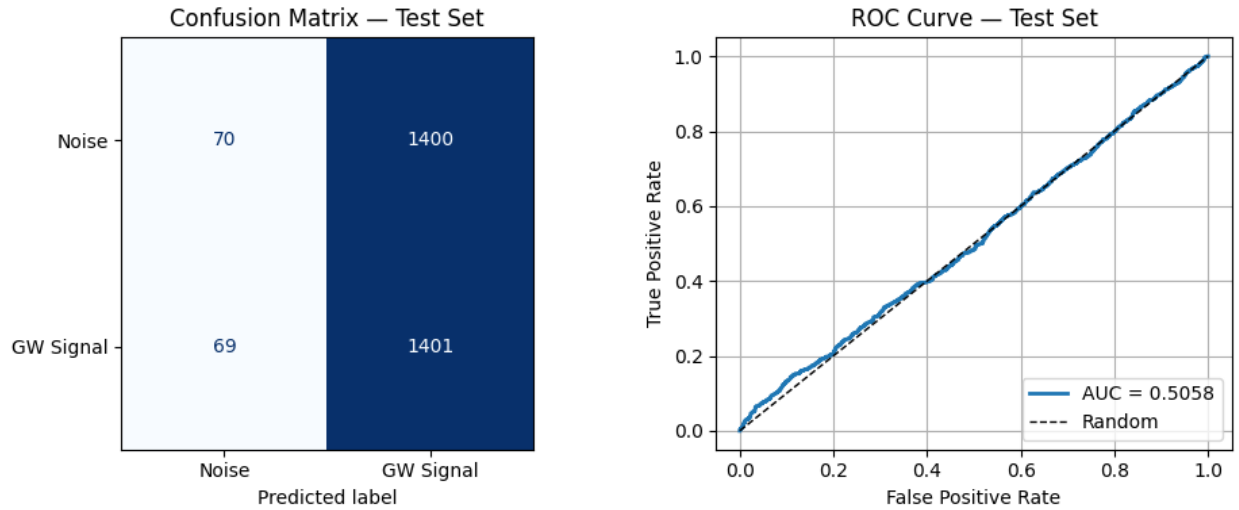


Figura 57: Matriz de confusión del clasificador LSTM en el conjunto de prueba y curva ROC para datos de entrada no optimizados.

Cuadro 16: Reporte de clasificación del modelo LSTM en el conjunto de prueba no optimizado.

Clase	Precisión	Recall	F1-score	Soporte
Ruido (0)	0.50	0.05	0.09	1470
Señal GW (1)	0.50	0.95	0.66	1470
Accuracy		0.50		2940
Promedio macro	0.50	0.50	0.37	2940
Promedio ponderado	0.50	0.50	0.37	2940
ROC-AUC: 0,5058				

Las Figuras 56 y 57 muestran las curvas de entrenamiento y la matriz de confusión del clasificador LSTM al emplear una ventana de entrada de 105 muestras ($w_s = 48, W = 105$). Los resultados del reporte de clasificación (Tabla 16) revelan una degradación severa del rendimiento: el modelo alcanza únicamente un 50 % de exactitud y un ROC-AUC de 0.5058, valores equivalentes a una clasificación aleatoria, lo que indica que esta configuración carece de capacidad discriminativa.



Con esto concluye la primera etapa del proceso de detección. Si el clasificador determina que la señal corresponde a ruido, el procesamiento se detiene en este punto. En caso contrario, cuando se detecta una señal gravitacional, el proceso continúa hacia la segunda etapa: la reconstrucción mediante el autoencoder.



Autoencoder

El autoencoder tiene como objetivo reconstruir la señal gravitacional original a partir de la versión contaminada con ruido. El modelo fue entrenado con los parámetros de la Tabla 7, utilizando el 80 % de los datos para entrenamiento y el 20 % restante para prueba.

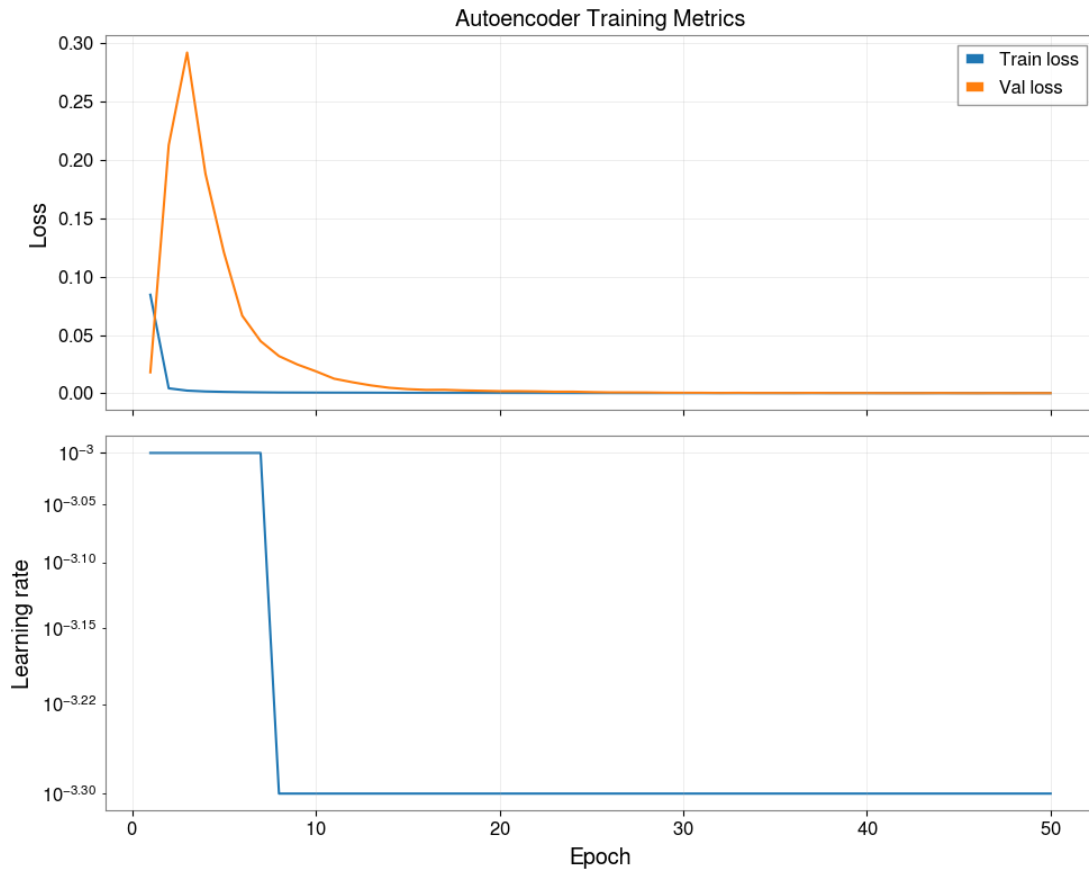
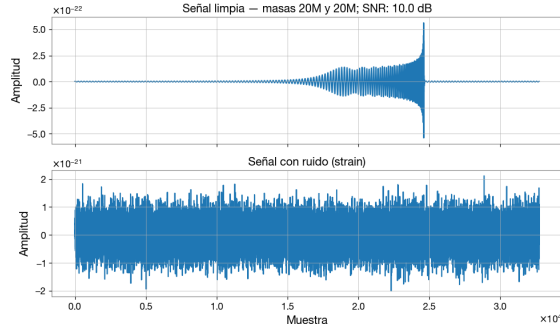
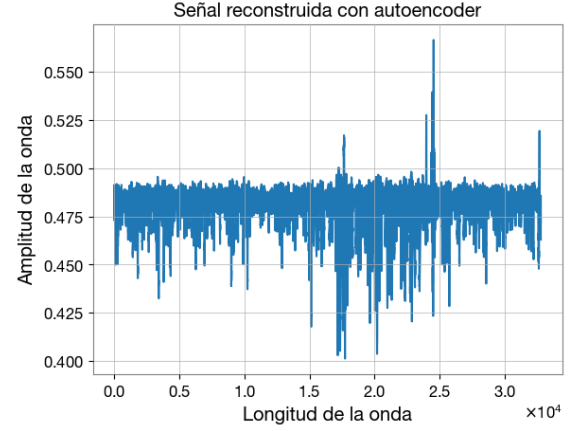


Figura 58: Gráfica de pérdida y precisión por época durante el entrenamiento del autoencoder. La segunda gráfica muestra el cambio del learning rate a lo largo de las épocas.

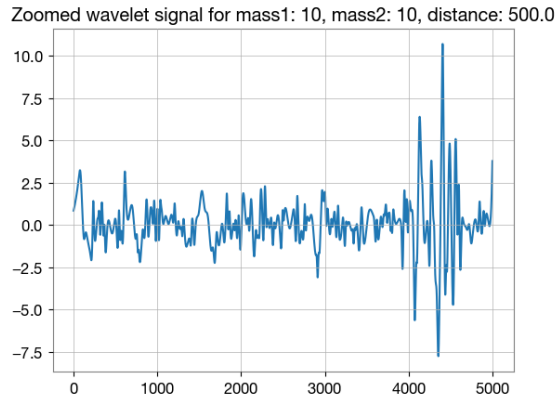
En la Figura 58 se ilustra la evolución de la pérdida durante el entrenamiento. Podemos apreciar cómo la pérdida para el conjunto de entrenamiento y validación disminuyen con el paso de las épocas hasta llegar a un valor cercano a 0 ($1,3 \times 10^{-4}$ para el conjunto de entrenamiento y $1,9 \times 10^{-4}$ para el conjunto de validación).



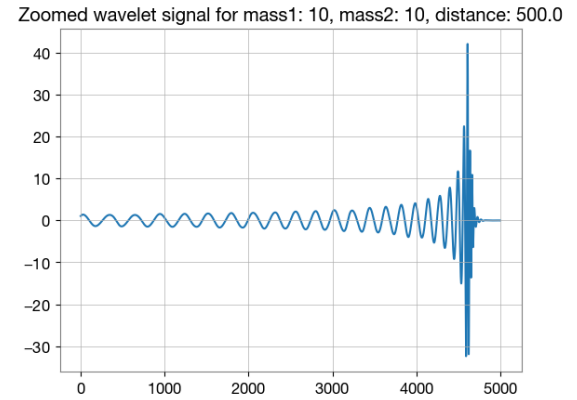
(a) Onda gravitacional simulada limpia y con ruido.



(b) Señal reconstruida con el autoencoder.



(c) Diferencia de energía para la señal reconstruida.

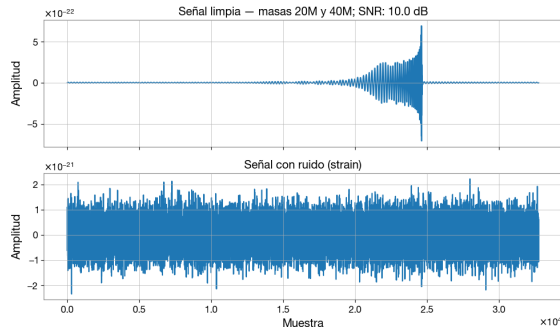


(d) Diferencia de energía para la señal limpia.

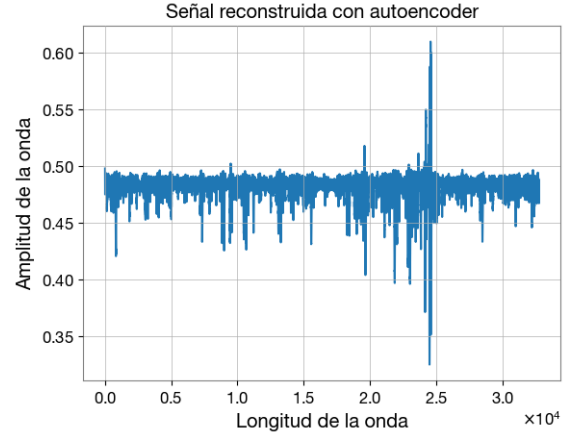
Figura 59: Proceso de reconstrucción mediante el autoencoder para un evento de masas de $20M_{\odot}$ y SNR de 10 dB.

Retomando el ejemplo del evento 1, de dos objetos de $10, M_{\odot}$ con SNR de 10 dB, que el clasificador LSTM no logra detectar. Al incrementar las masas a $20, M_{\odot}$ manteniendo el mismo nivel de ruido, la señal es correctamente detectada y puede ser reconstruida por el DAE.

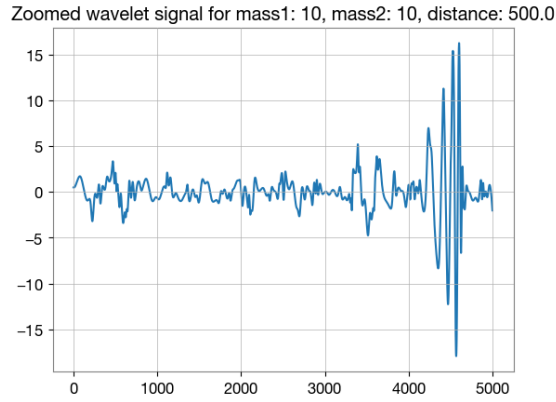
La Figura 59 muestra el proceso completo de reconstrucción. A partir de la señal contaminada (Figura 59a), se extrae la desviación estándar y se clasifica mediante el LSTM; al confirmarse la presencia de una señal gravitacional, el DAE produce la señal reconstruida (Figura 59b). Finalmente, se aplica el análisis de diferencia de energías tanto a la señal reconstruida (Figura 59c) como a la señal limpia de referencia (Figura 59d), permitiendo evaluar la reconstrucción.



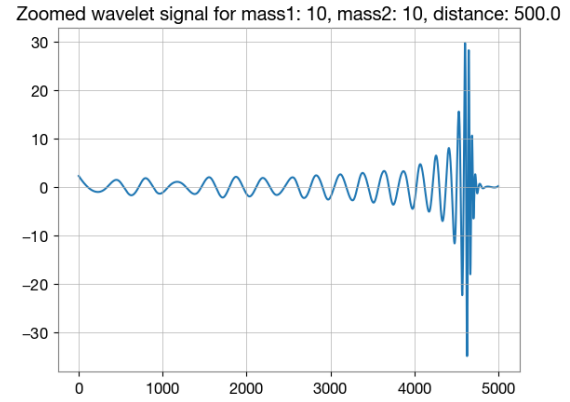
(a) Onda gravitacional simulada limpia y con ruido.



(b) Señal reconstruida con el autoencoder.



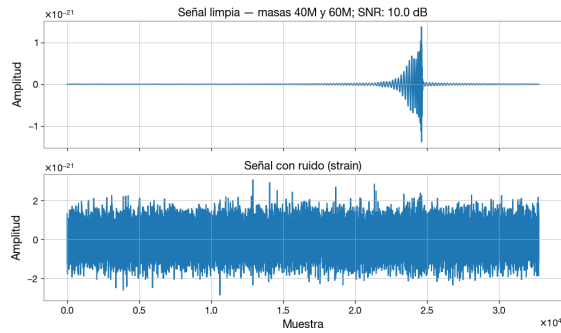
(c) Diferencia de energía para la señal reconstruida.



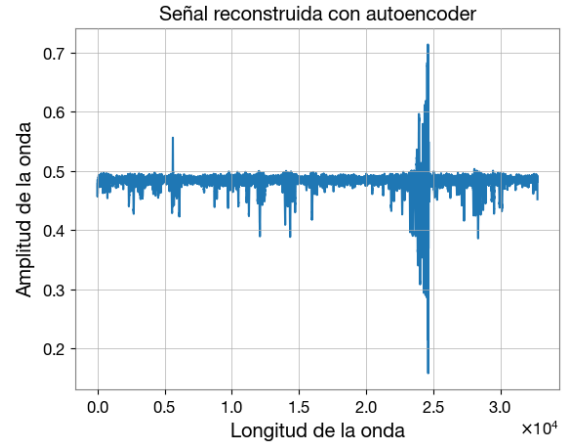
(d) Diferencia de energía para la señal limpia.

Figura 60: Proceso de reconstrucción mediante el autoencoder para un evento de masas de $20M_{\odot}$ y $40M_{\odot}$ con un SNR de 10 dB.

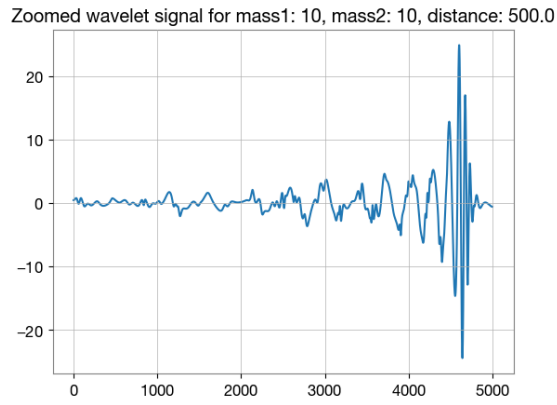
La Figura 60 ilustra el efecto de aumentar las masas a $20, M_{\odot}$ y $40, M_{\odot}$ manteniendo un SNR de 10 dB. En este caso, la reconstrucción producida por el DAE es más precisa, lo que se refleja en una mayor similitud entre las diferencias de energía de la señal reconstruida (Figura 60c) y la señal limpia de referencia (Figura 60d). Este comportamiento se acentúa al incrementar las masas a $40, M_{\odot}$ y $60, M_{\odot}$, donde se obtiene una reconstrucción aún más fiel bajo el mismo nivel de ruido (Figura 61).



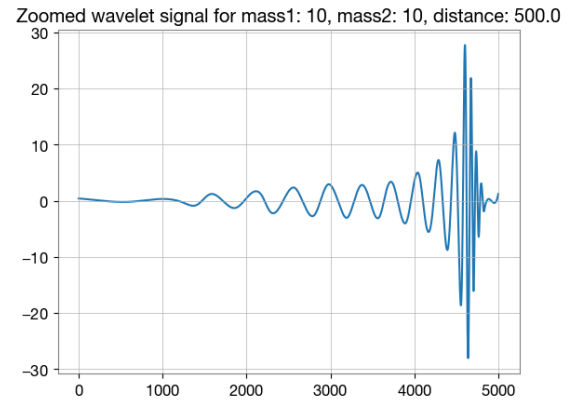
(a) Onda gravitacional simulada limpia y con ruido.



(b) Señal reconstruida con el autoencoder.

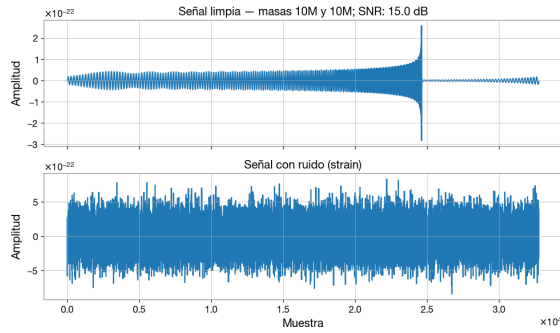


(c) Diferencia de energía para la señal reconstruida.

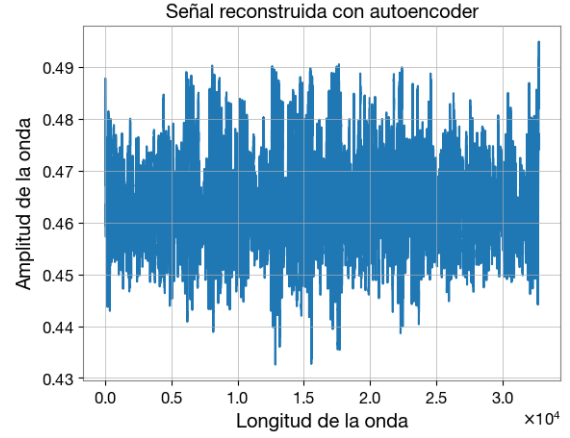


(d) Diferencia de energía para la señal limpia.

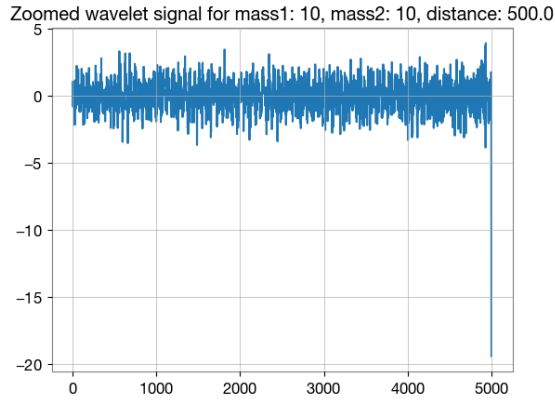
Figura 61: Proceso de reconstrucción mediante el autoencoder para un evento de masas de $40M_{\odot}$ y $60M_{\odot}$ con un SNR de 10 dB.



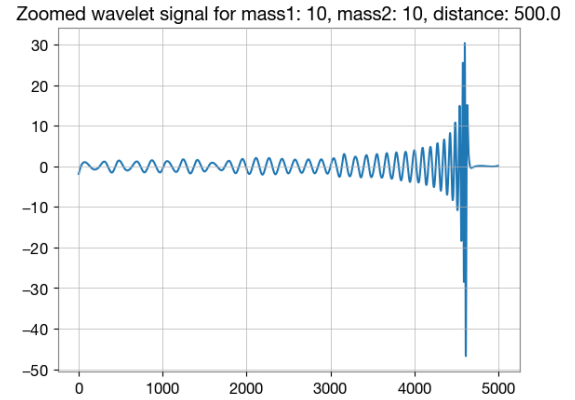
(a) Onda gravitacional simulada limpia y con ruido.



(b) Señal reconstruida con el autoencoder.



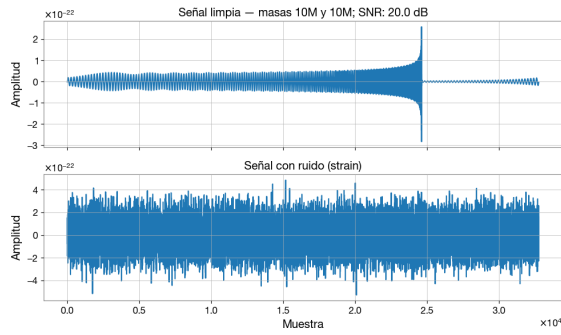
(c) Diferencia de energía para la señal reconstruida.



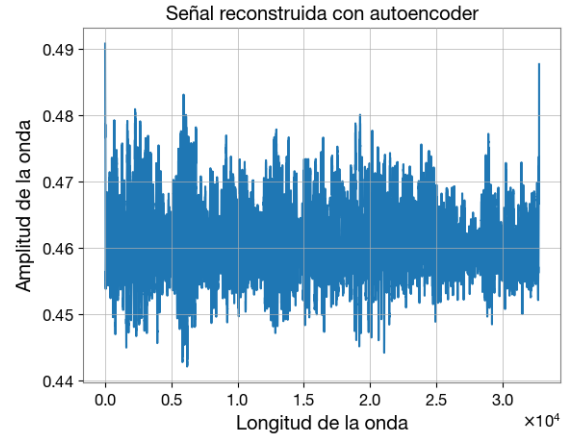
(d) Diferencia de energía para la señal limpia.

Figura 62: Proceso de reconstrucción mediante el autoencoder para un evento de masas de $10M_{\odot}$ y $10M_{\odot}$ con un SNR de 15 dB.

Un comportamiento notable se observa al incrementar el SNR manteniendo masas por debajo de $20M_{\odot}$. Como muestra la Figura 62, con masas de $10M_{\odot}$ y $10M_{\odot}$ y un SNR de 15 dB, el clasificador LSTM logra detectar correctamente la señal, sin embargo, el DAE no es capaz de reconstruirla de manera satisfactoria. Este mismo comportamiento se reproduce al elevar el SNR a 20 dB conservando las mismas masas, como se aprecia en la Figura 63.

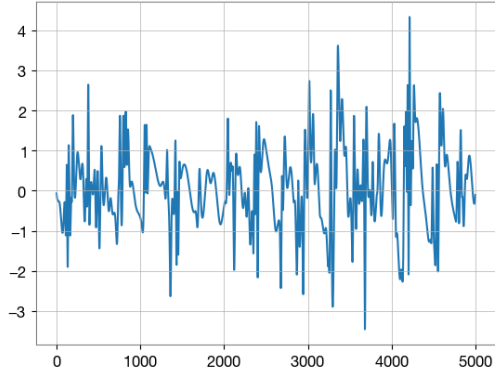


(a) Onda gravitacional simulada limpia y con ruido.



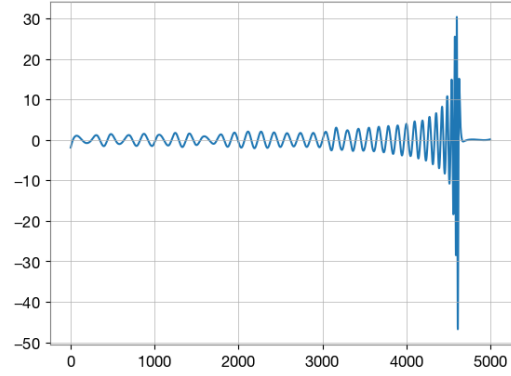
(b) Señal reconstruida con el autoencoder.

Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(c) Diferencia de energía para la señal reconstruida.

Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(d) Diferencia de energía para la señal limpia.

Figura 63: Proceso de reconstrucción mediante el autoencoder para un evento de masas de $10M_{\odot}$ y $10M_{\odot}$ con un SNR de 20 dB.

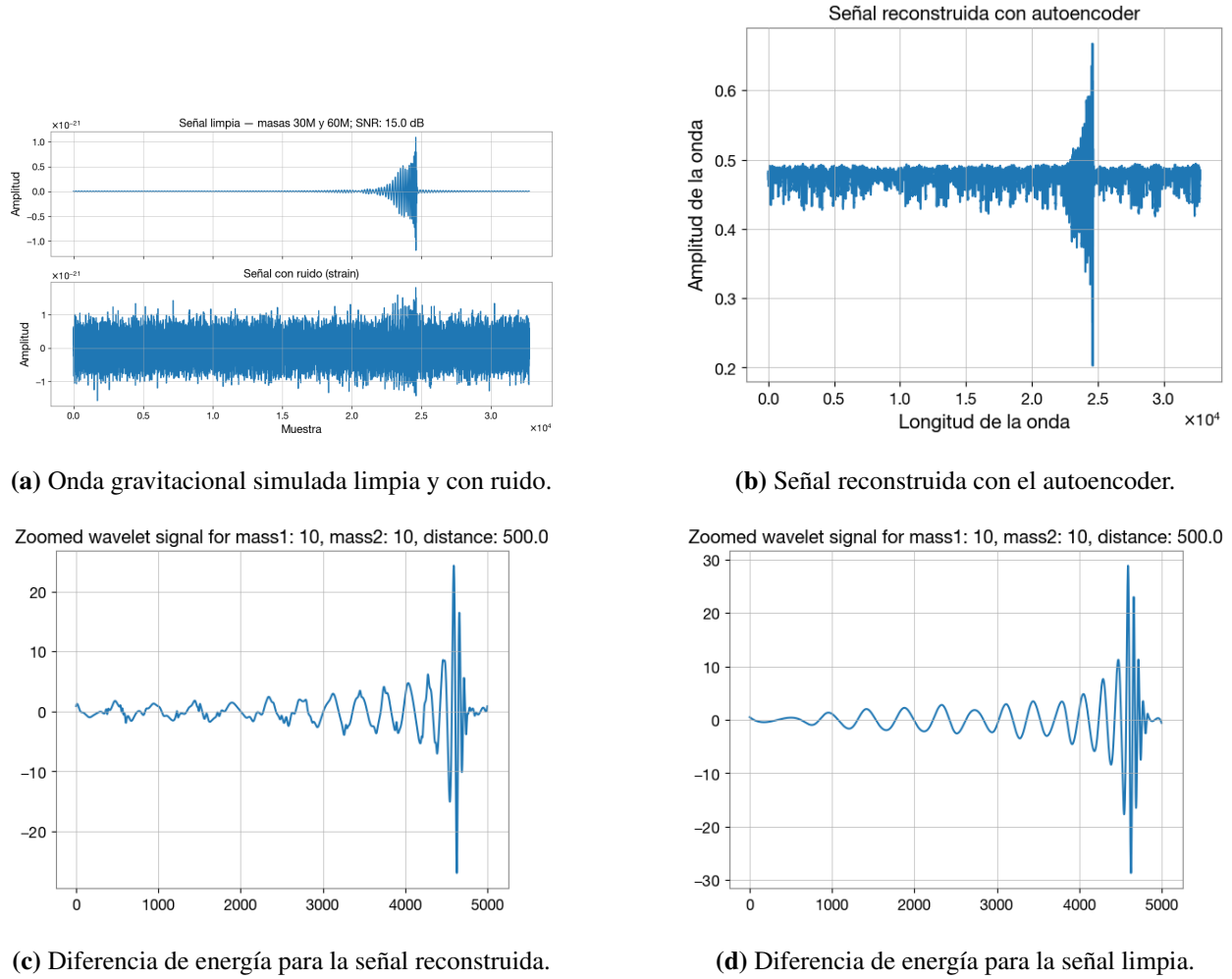
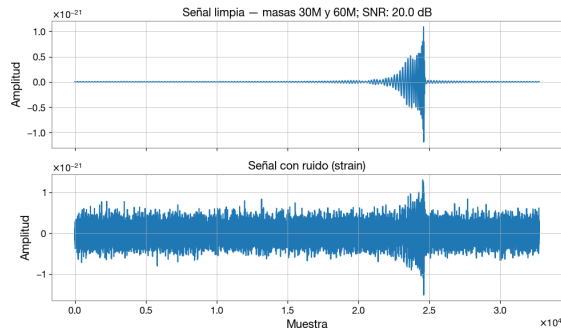


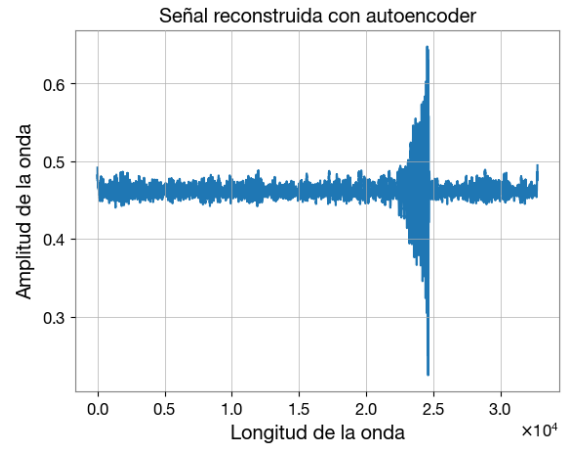
Figura 64: Proceso de reconstrucción mediante el autoencoder para un evento de masas de $30M_{\odot}$ y $60M_{\odot}$ con un SNR de 15 dB.

Sin embargo, una vez que aumentamos las masas de los objetos, este comportamiento desaparece, como se aprecia en la Figura 64 y 65. En el caso de la Figura 65, con masas de $30M_{\odot}$ y $60M_{\odot}$ y un SNR de 20 dB, el proceso del autoencoder es capaz de reconstruir una señal limpia (Figura 65b) a partir de un ruido (Figura 65a). La cual, cuando aplicamos diferencia de energía (Figura 65c), podemos ver como la señal reconstruida es casi igual a la limpia (Figura 65d).

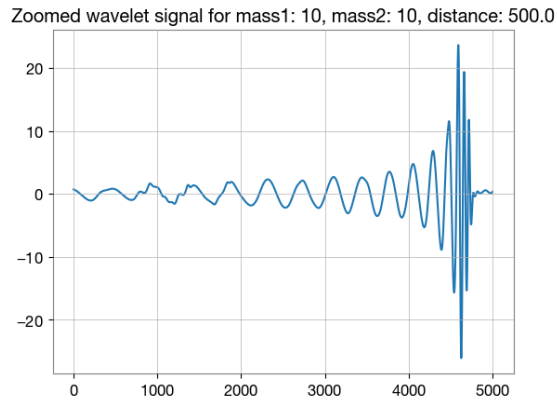
Una vez que el modelo reconstruye una señal y obtenemos su diferencia de energía, podemos pasar a la regresión de la masa del chirp de la onda gravitacional mediante el modelo de CNN.



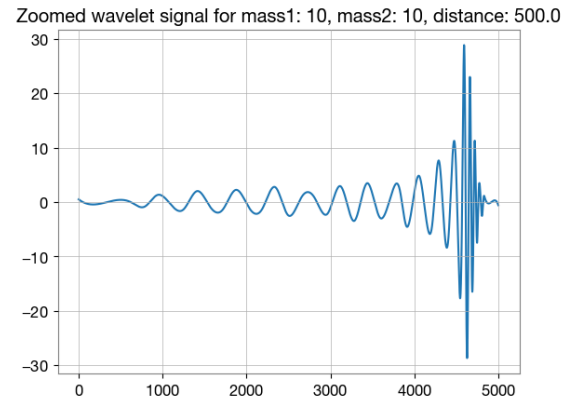
(a) Onda gravitacional simulada limpia y con ruido.



(b) Señal reconstruida con el autoencoder.



(c) Diferencia de energía para la señal reconstruida.



(d) Diferencia de energía para la señal limpia.

Figura 65: Proceso de reconstrucción mediante el autoencoder para un evento de masas de $30M_{\odot}$ y $60M_{\odot}$ con un SNR de 20 dB.



Cuadro 17: Métricas agregadas del modelo de reconstrucción (media $\pm$ desv. est. sobre todas las muestras de prueba).

Métrica	Media	Desv. Est.
RMSE	0.0058	0.0019
MAE	0.0037	0.0009
Correlación	0.9336	0.1041

La Tabla 17 presenta las métricas agregadas del modelo de reconstrucción. El RMSE (error cuadrático medio) cuantifica la diferencia promedio entre la señal reconstruida y la señal limpia en valores normalizados $[0, 1]$, al elevar al cuadrado los residuos, penaliza con mayor peso las desviaciones grandes, por lo que valores cercanos a 0 indican una reconstrucción precisa. El MAE (error absoluto medio) mide la misma discrepancia sin elevar al cuadrado, resultando en una métrica más robusta ante valores atípicos. Finalmente, la correlación de Pearson evalúa el grado de similitud estructural entre ambas señales; un valor próximo a 1 indica que la señal reconstruida reproduce fielmente la forma de onda de la señal limpia.

Los valores de $\text{RMSE} = 0,0058$ ($\sigma = 0,0019$) y $\text{MAE} = 0,0037$ ($\sigma = 0,0009$) reportados en la Tabla 17 evidencian que el modelo produce reconstrucciones con un error promedio inferior al 0,6 % en la escala normalizada, lo que refleja una alta fidelidad entre la señal limpia y la reconstruida. Este resultado es consistente con la correlación de Pearson de 0,9336, que confirma una fuerte similitud estructural entre ambas señales.

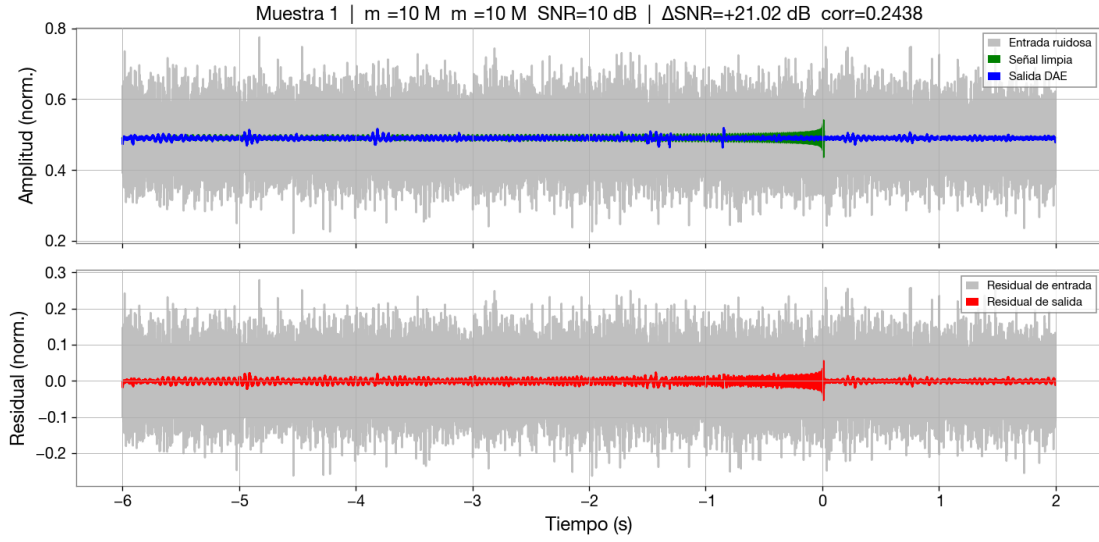


Figura 66: Comparación entre la señal limpia, la reconstruida y la información residual entre las señales para masas de $10M_{\odot}$ y $10M_{\odot}$ y un SNR de 10 dB.

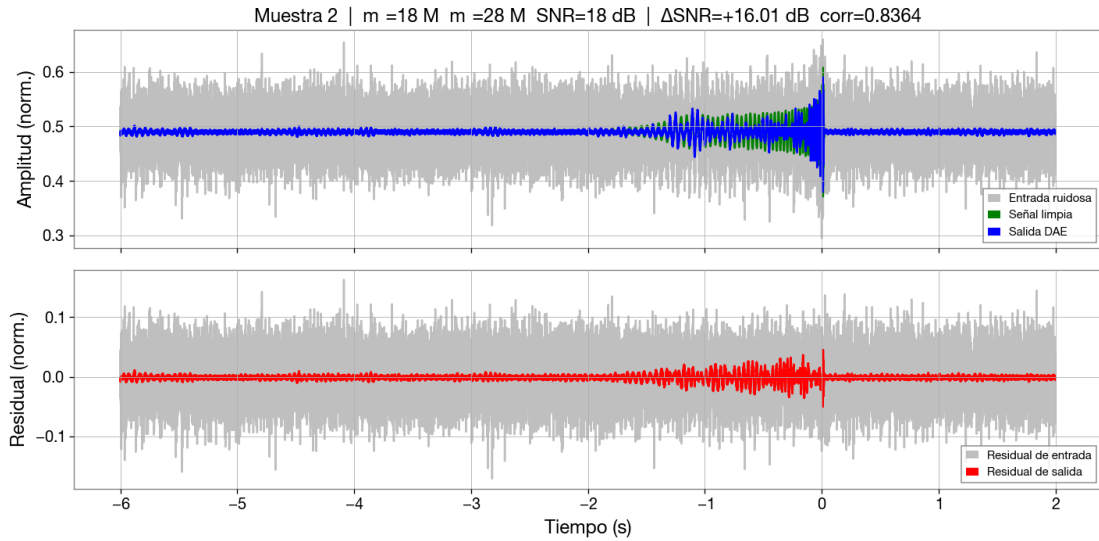


Figura 67: Comparación entre la señal limpia, la reconstruida y la información residual entre las señales para masas de $18M_{\odot}$ y $28M_{\odot}$ y un SNR de 18 dB.

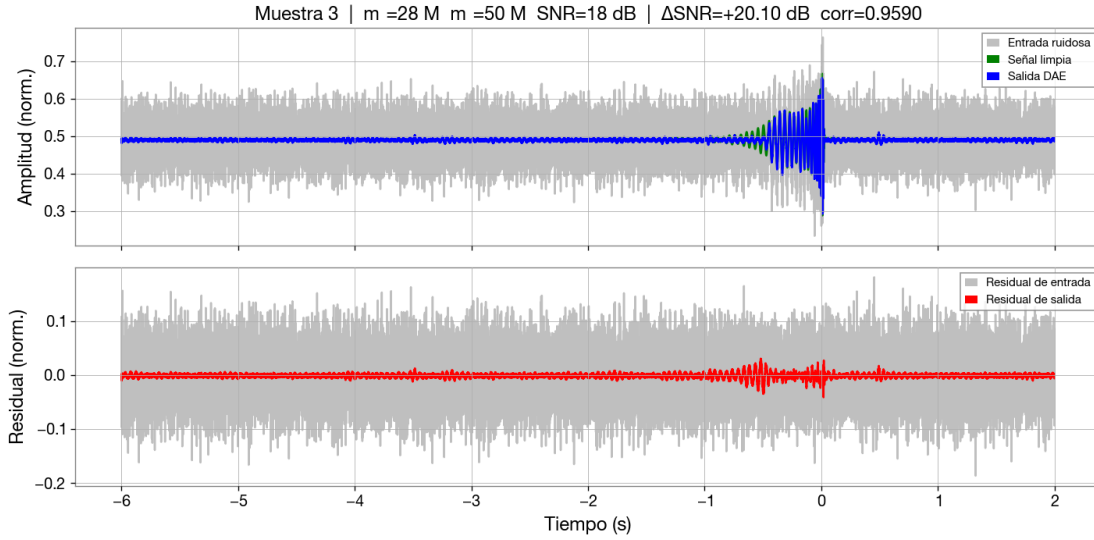


Figura 68: Comparación entre la señal limpia, la reconstruida y la información residual entre las señales para masas de $28M_{\odot}$ y $50M_{\odot}$ y un SNR de 18 dB.

Las Figuras 66, 67 y 68 presentan la comparación entre la señal limpia, la señal reconstruida y el residual para distintas combinaciones de masa y SNR. En cada figura se identifican tres componentes: la señal limpia (verde), la señal reconstruida (azul) y la señal residual (rojo), definida como la diferencia entre ambas; un residual de baja amplitud indica que el modelo capturó la mayor parte de la información de la forma de onda.

La Figura 66 ilustra el caso de masas pequeñas, donde el modelo muestra dificultades para recuperar la forma de onda. La correlación de Pearson obtenida es de 0,2438, valor que refleja una reconstrucción deficiente; consecuentemente, la señal residual conserva un comportamiento muy similar a la señal limpia, lo que indica que prácticamente no se recuperó información relevante.

Las Figuras 67 y 68 contrastan dos escenarios con el mismo SNR de 18 dB pero distintas combinaciones de masa. Para masas de $18M_{\odot}$ y $28M_{\odot}$ (Figura 67) la correlación es de 0,8364, resultado que evidencia una reconstrucción efectiva. Al incrementar las masas a $28M_{\odot}$ y $50M_{\odot}$ (Figura 68), la correlación asciende a 0,9590, lo que demuestra que la masa del sistema es un factor determinante en la capacidad de reconstrucción del modelo.



Modelo CNN de regresión

La etapa final del proceso corresponde a la regresión de la masa del chirp de la onda gravitacional. Para ello, el modelo fue entrenado utilizando las señales de diferencia de energías $[Z(t)_0, \dots, Z(t)_N]$.

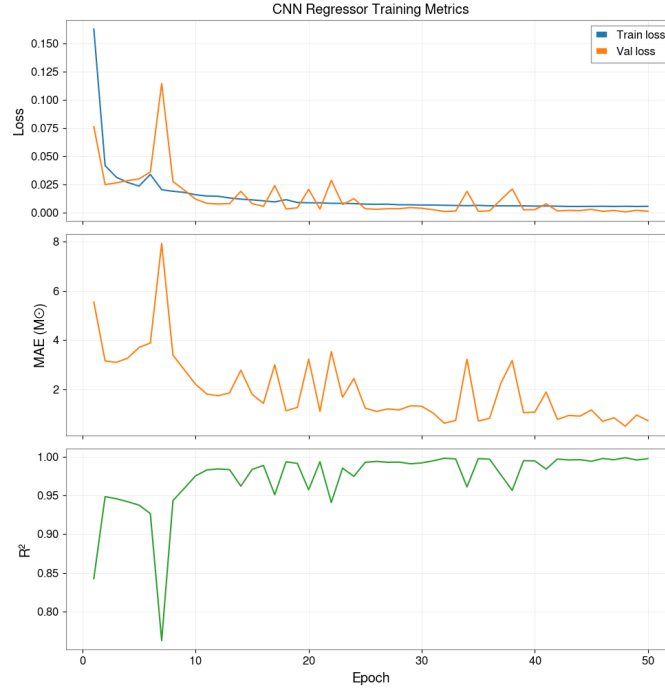


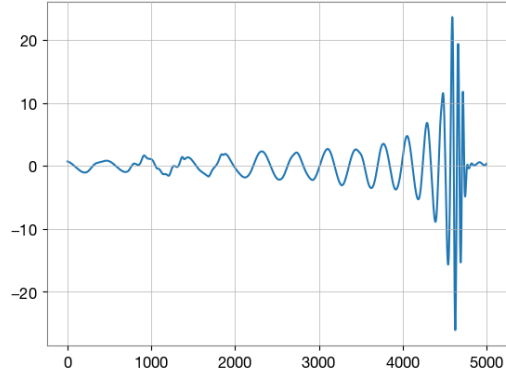
Figura 69: Gráfica de pérdida por época durante el entrenamiento para el conjunto de entrenamiento y validación del modelo de CNN de regresión.

En la Figura 69 se ilustra la evolución de la pérdida durante el entrenamiento. La pérdida para el conjunto de entrenamiento y validación disminuyen con el paso de las épocas, hasta llegar a un MAE de $0,7473M_{\odot}$ y un R^2 de 0,9973

Realizando la regresión de la señal reconstruida (Figura 70a) mediante el modelo de CNN obtenemos los resultados de la tabla 18. También podemos ver el comportamiento del modelo a diferentes niveles de ruido. En la figura 71 se muestra la diferencia de energía para un evento de masas de $30M_{\odot}$ y $60M_{\odot}$ con un SNR de 15 dB, y su regresión mediante el modelo de CNN en la tabla 19. Finalmente en la figura 72 vemos la reconstrucción de una señal de SNR de 10 dB y la regresión de la masa del chirp mediante el modelo de CNN en la tabla 20.

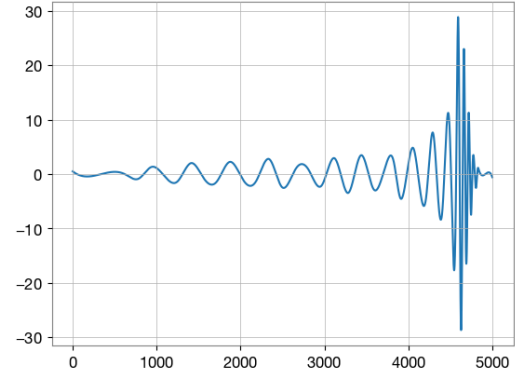


Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(a) Diferencia de energía para la señal reconstruida.

Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(b) Diferencia de energía para la señal limpia.

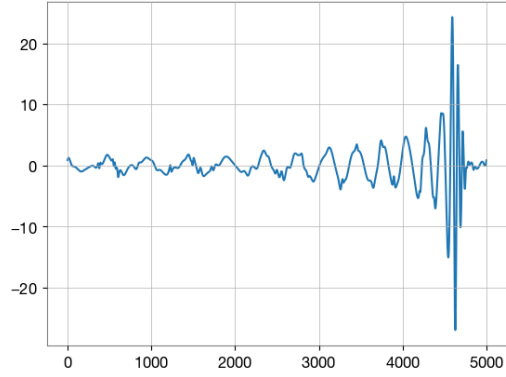
Figura 70: Diferencia de energía para la señal reconstruida y limpia para un evento de masas de $30M_{\odot}$ y $60M_{\odot}$ con un SNR de 20 dB.

Cuadro 18: Resultado de la detección y regresión de la señal reconstruida para dos masas de $30M_{\odot}$ y $60M_{\odot}$ con un SNR de 20 dB.

Parámetro / Resultado	Valor
Masa 1 ($M_{\odot}$)	30
Masa 2 ($M_{\odot}$)	60
SNR	20
Detección (umbral)	Si
Probabilidad de señal GW	1.0
Clase predicha	Señal GW
Masa del chirp predicha	$36,5739M_{\odot}$
Masa del chirp real	$36,5019M_{\odot}$

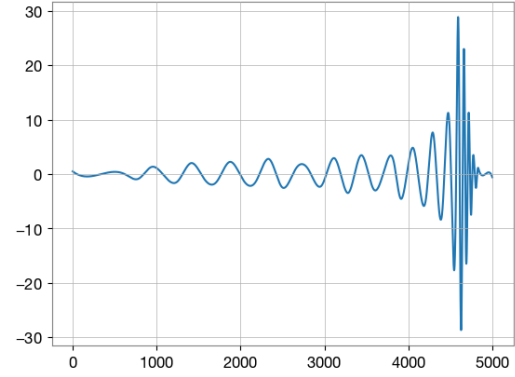


Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(a) Diferencia de energía para la señal reconstruida.

Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(b) Diferencia de energía para la señal limpia.

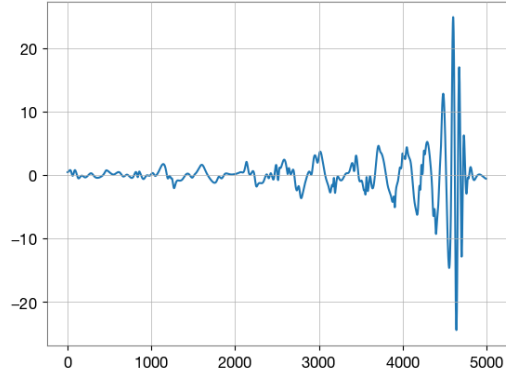
Figura 71: Diferencia de energía para la señal reconstruida y limpia para un evento de masas de $30M_{\odot}$ y $60M_{\odot}$ con un SNR de 15 dB.

Cuadro 19: Resultado de la detección y regresión de la señal reconstruida para dos masas de $30M_{\odot}$ y $60M_{\odot}$ con un SNR de 15 dB.

Parámetro / Resultado	Valor
Masa 1 ($M_{\odot}$)	30
Masa 2 ($M_{\odot}$)	60
SNR	15
Detección (umbral)	Si
Probabilidad de señal GW	1.0
Clase predicha	Señal GW
Masa del chirp predicha	$31,1101M_{\odot}$
Masa del chirp real	$36,5019M_{\odot}$

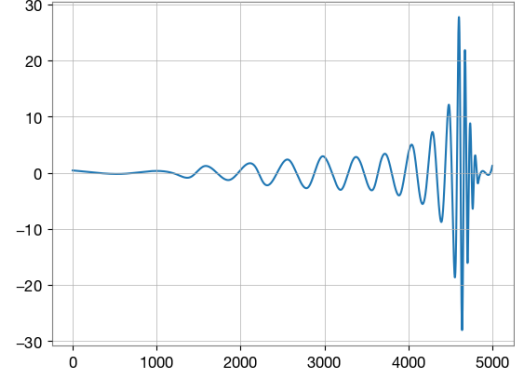


Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(a) Diferencia de energía para la señal reconstruida.

Zoomed wavelet signal for mass1: 10, mass2: 10, distance: 500.0



(b) Diferencia de energía para la señal limpia.

Figura 72: Diferencia de energía para la señal reconstruida y limpia para un evento de masas de $40M_{\odot}$ y $60M_{\odot}$ con un SNR de 10 dB.

Cuadro 20: Resultado de la detección y regresión de la señal reconstruida para dos masas de $40M_{\odot}$ y $60M_{\odot}$ con un SNR de 10 dB.

Parámetro / Resultado	Valor
Masa 1 ($M_{\odot}$)	40
Masa 2 ($M_{\odot}$)	60
SNR	10
Detección (umbral)	Si
Probabilidad de señal GW	1.0
Clase predicha	Señal GW
Masa del chirp predicha	$45,4178M_{\odot}$
Masa del chirp real	$42,4744M_{\odot}$



Discusión

La sección anterior presentó los resultados del proceso completo, desde la construcción de la base de datos hasta la regresión de la masa del chirp a partir de la señal reconstruida. A continuación, se interpretan dichos resultados, se comparan con trabajos previos en la literatura, y se discuten sus implicaciones y limitaciones.

Comparación con el estado del arte

La Tabla 21 resume los trabajos más relevantes en detección, reconstrucción y estimación de parámetros de ondas gravitacionales mediante aprendizaje automático, y los contrasta con los resultados obtenidos en este estudio.

Los trabajos de George & Huerta (2018b) y George & Huerta (2018a) son los primeros en aplicar redes neuronales convolucionales para la detección y estimación de parámetros en datos de LIGO, estableciendo la viabilidad del enfoque. Shen et al. (2019) proponen un DAE recurrente para denoising en un rango de SNR de $[6, 5, 26]$ dB, resultado directamente comparable con el rango $[10, 20]$ dB del DAE U-Net implementado en este estudio. Qiu et al. (2023) y Ussipov et al. (2024) representan el estado del arte más reciente en detección en tiempo real y clasificación de tipos de eventos, respectivamente.

Es importante destacar que la comparación directa entre trabajos es limitada debido a las diferencias en los conjuntos de datos utilizados (parámetros físicos, rangos de masa, rango de SNR), las métricas reportadas y las tareas específicas abordadas. No obstante, los resultados obtenidos en este estudio son comparables en orden de magnitud con los reportados en la literatura para condiciones similares.



Cuadro 21: Comparación con trabajos representativos del estado del arte en análisis de ondas gravitacionales mediante aprendizaje automático. Los campos marcados con – requieren consulta directa de la referencia correspondiente.

Trabajo	Tarea	Modelo	Datos	Métrica	Resultado
George & Huerta (2018b)	Detección	CNN	Simulados	Precisión	100 % ^a
George & Huerta (2018a)	Detección / Regresión de parámetros	CNN	Simulados / LIGO real	Precisión / MPAAE	100 % / 35 % ^a
Shen et al. (2019)	Denoising (DAE)	DAE recurrente	Simulados	Correlación	0,994
Ussipov et al. (2024)	Clasificación	CNN	Simulados	Precisión	96 % ^a
Qiu et al. (2023)	Detección/Clasificación	CNN	Simulados / LIGO real	Precisión / ROC	100 % / 1 ^b
Este trabajo	Detección binaria	LSTM (2 capas)	Simulados	Precisión / ROC-AUC	99,99 % / 0,9999
Este trabajo	Denoising	DAE (U-Net)	Simulados	MAE / Correlación	0,0037 / 0,9336
Este trabajo	Regresión de $\mathcal{M}_c$	CNN (4 capas)	Simulados	MAE / R^2	0,747 $M_\odot$ / 0,9973

^a Este valor es obtenido con señales con un $\text{SNR} \geq 10$.

^b Los resultados se obtienen con un $\text{SNR} \geq 13$.

Base de datos

La base de datos empleada en este estudio fue generada mediante la simulación de ondas gravitacionales utilizando la librería *Bilby*, que permite modelar eventos bajo distintas condiciones físicas, incluyendo masas, SNR, distancia e inclinación orbital. Un aspecto que se hizo evidente desde las primeras etapas de construcción de la base de datos es la alta dimensionalidad de las señales simuladas: cada forma de onda contiene un número elevado de muestras, lo que resulta en una base de datos de tamaño considerable y, en consecuencia, de alto costo computacional para su generación.

Esta dimensionalidad representa la primera limitación del estudio: la construcción de la base de datos requiere un costo computacional que excede los recursos disponibles. Como solución, se restringió el espacio de parámetros explorado. La base de datos resultante, descrita en la sección de metodología, comprende un total de 122,500 señales con variación únicamente en las masas y el SNR, muestreados con incrementos de $2, M_\odot$ y 2 dB respectivamente, y bajo la restricción de que ninguna masa sea el doble de la otra.



La restricción del espacio de parámetros no constituye una limitación exclusiva de este estudio. George & Huerta (2018a) generaron 2,500 señales de eventos BBH con masas en el rango $M \in [5, 75], M_{\odot}$, con incrementos de $1, M_{\odot}$. Shen et al. (2019) emplearon una relación de masas máxima de 10 y un rango $M \in [2, 95], M_{\odot}$ con incrementos de $1, M_{\odot}$, obteniendo 9,861 muestras de entrenamiento y 2,459 de validación. Por su parte, Ussipov et al. (2024) construyeron una base de datos de 480,000 señales de entrenamiento y 16,000 de validación, con masas en $M \in [2, 95], M_{\odot}$, SNR en el rango $[0, 20]$ dB, y una relación de masas máxima de 10.

Como se observa en los trabajos citados, restringir o fijar parámetros en la generación de señales es una práctica común en la literatura. No obstante, es importante señalar que una base de datos construida bajo estas condiciones es específica para los casos contemplados en su diseño. Un estudio de carácter más general requeriría variar el mayor número posible de parámetros físicos, lo que, si bien aumentaría la representatividad del conjunto de datos, implicaría un costo computacional considerablemente mayor.

Análisis preliminar

Previo a la construcción de la base de datos final, se realizó un procesamiento de las señales limpias que consistió en calcular su transformada wavelet y obtener la diferencia de energías de la descomposición de nivel 9. Este paso tiene como finalidad identificar la región de interés $[z(t_0), \dots, z(t_N)]$ dentro de cada señal, la cual se emplea posteriormente para entrenar el modelo de regresión de la masa del chirp.

Cabe destacar que los resultados de la transformada wavelet dependen de la wavelet madre seleccionada. La elección entre wavelets como Morlet, Haar, Daubechies o Symlets no afecta únicamente la descomposición de nivel 9, sino la descomposición en su totalidad, independientemente del nivel analizado.



La selección de la wavelet madre va a variar según las señales que estemos trabajando. En este estudio, se realizaron diferentes experimentos con distintas familias de wavelets, donde la única que entregó resultados significativos fue la de Haar.

En la Figura 73 se observa la descomposición al utilizar la wavelet madre Sym4 y db4. Podemos ver en la Figura 73a como la familia de las wavelets Symlets pierde su resolución frecuencial cuando trabaja con las señales generadas. Lo mismo ocurre en la Figura 73b, con la familia de las wavelets Daubechies.

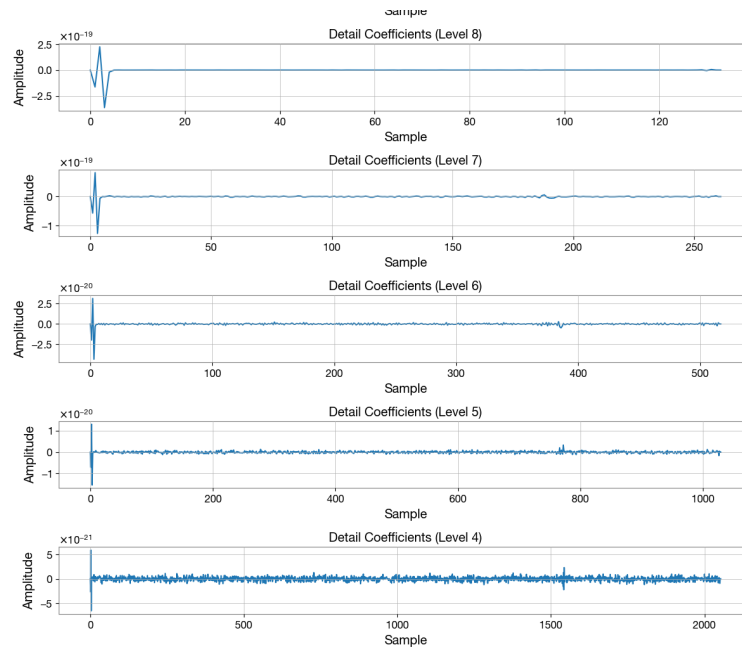
Análisis de la señal

En los estudios que se realizan en LIGO para la detección de eventos gravitacionales, uno de los primeros pasos a realizar es una serie de procesamientos que permiten reducir el ruido y mejorar la calidad de la información contenida en la señal (Abbott et al., 2020). Una parte importante de este estudio es el abordar este problema utilizando una metodología diferente.

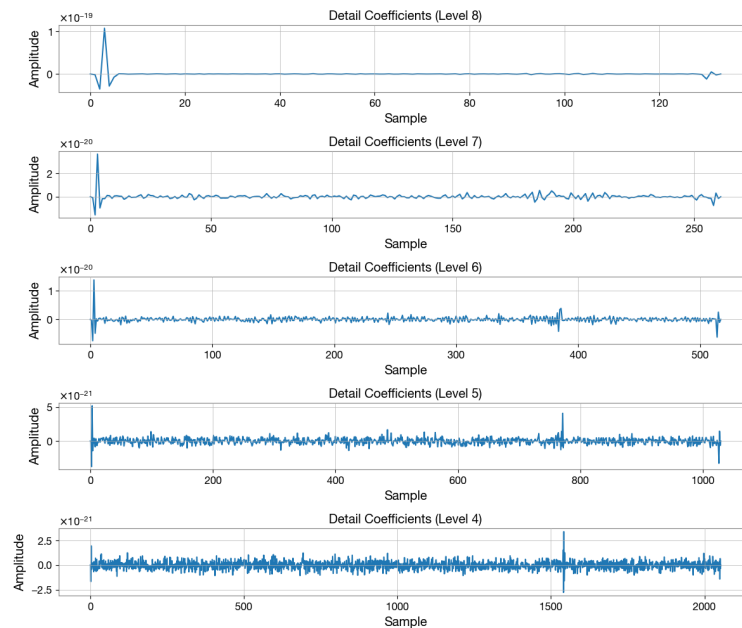
Si bien, en lo que se refiere a la detección de eventos gravitacionales, los modelos implementados en este estudio hacen la mayoría del trabajo, el procesamiento previo de la señal es un elemento importante del proceso, ya que permite mejorar la calidad de la información contenida en la señal y de esta manera, entrenar mejores modelos.

Con este fin, se consideró desde un inicio, la transformada wavelet como una herramienta para procesar la señal. Esto se ve reflejado en su amplio uso en el estudio, especialmente en encontrar la región de interés $[z(t_0), \dots, z(t_N)]$ de la señal limpia. La limitante de la transformada surge al momento de trabajar con señales inundadas por ruido, ya que la descomposición puede perder información relevante de la señal original.

El uso de un filtro FIR previo a la transformada de wavelet permite reducir el ruido de la señal, facilitando la descomposición. De la misma manera, el segundo filtro se enfoca en suprimir los artefactos de alta frecuencia introducidos durante la reconstrucción wavelet.



(a) Wavelets de la familia Symlets (Sym4)



(b) Wavelets de la familia Daubechies (db4)

Figura 73: En la figura 73a se observa la descomposición al utilizar la wavelet madre Sym4. En la figura 73b se observa la descomposición al utilizar la wavelet madre db4.



Es en esta etapa donde nos topamos por primera vez con la particularidad del comportamiento de las masas. Como vimos en la sección de resultados, los modelos, así como el procesamiento de la señal, tenían un rendimiento muy bajo al momento de trabajar con masas menores a 20. Esto, como ya lo comentamos, puede ser debido a que la mayoría de la información está contenida en la masa del chirp, es decir, en la masa resultante del evento (Cutler & Flanagan , 1994).

Masa del chirp

El comportamiento del sistema en función de la masa del chirp plantea interrogantes que merecen ser abordadas en trabajos futuros. En particular, resulta relevante investigar por qué los tres modelos implementados, junto con el preprocesamiento de la señal, presentan un rendimiento reducido para masas del chirp inferiores a $30M_{\odot}$.

Este resultado es interesante debido a que el comportamiento es, en su mayoría, independiente del SNR de la señal. Esto nos plantea un desafío claro para resolver en futuros estudios. El resolver el comportamiento del ruido presente en la señal con masas pequeñas, no solo resolvería el problema del rendimiento de los modelos, y el comportamiento de agujeros negros de masas reducidas, sino que también podría permitir establecer una base para extender el estudio al comportamiento de estrellas de neutrones, cuyas masas no son mayores a $2 M_{\odot}$.

PSO

El algoritmo PSO cumplió dos funciones dentro del sistema. La primera consistió en optimizar los tamaños de ventana w_s y W con el objetivo de maximizar la desviación estándar de la energía de la señal analizada. La segunda fue la reducción de la dimensionalidad de la señal de entrada al clasificador LSTM.

Los resultados de esta optimización fueron significativos: la discriminabilidad de la característica extraída se maximizó y, simultáneamente, la señal de entrada se redujo de 8,000 a 61 muestras



sin comprometer el rendimiento del modelo. Esta reducción dimensional representa una mejora sustancial en la eficiencia computacional del proceso de clasificación.

Modelos de machine learning

Dejando de lado el problema de la masa del chirp, el comportamiento de los modelos presentan un comportamiento sobresaliente. Desde la detección automática de un evento gravitacional hasta la estimación de la masa del chirp. De los modelos implementados, el resultado más alentador fue el del autoencoder.

El autoencoder demostró ser capaz de reconstruir señales gravitacionales en un rango de SNR de 10 a 20 dB, a partir de las cuales el modelo CNN estimó la masa del chirp con un MAE de $0,7473M_{\odot}$. Considerando las restricciones computacionales del estudio, estos resultados abren la posibilidad de desarrollar, en trabajos futuros, modelos de mayor complejidad y precisión.



Conclusión

En este estudio se propuso una metodología integral para la detección, reconstrucción y estimación de la masa del chirp de ondas gravitacionales generadas por eventos de coalescencia de objetos compactos. El sistema está compuesto por tres etapas secuenciales: un clasificador LSTM para la detección del evento, un autoencoder denoising (DAE) para la reconstrucción de la señal, y un modelo CNN para la regresión de la masa del chirp.

El preprocesamiento mediante filtros FIR y la transformada wavelet constituyó una herramienta fundamental para el procesamiento de las señales. Complementariamente, la optimización conjunta de los tamaños de ventana $w_s = 97$ y $W = 61$ mediante el algoritmo PSO permitió maximizar la desviación estándar de la energía por ventana, resaltando la región de interés del evento gravitacional. Un resultado destacable de esta optimización fue la reducción de la dimensionalidad de la señal de entrada al clasificador LSTM de 8,000 a 61 muestras, sin comprometer el rendimiento del modelo. El clasificador LSTM entrenado sobre una base de datos balanceada de 19,600 señales alcanzó una exactitud del 99,99 % y un ROC-AUC de 0,9999, demostrando una capacidad discriminativa prácticamente óptima.

La reconstrucción mediante el DAE produjo resultados satisfactorios para señales con masas superiores a $20M_{\odot}$ presentando un MAE de 0,0037 y una correlación entre las señales de 0,9336 a partir de las cuales el modelo CNN estimó la masa del chirp con un error del $0,7473M_{\odot}$.

La limitación observada para masas del chirp inferiores a $30M_{\odot}$ sugiere que la base de datos de entrenamiento y el proceso de simulación requieren reevaluación en trabajos futuros, particularmente para extender la metodología al estudio de sistemas con estrellas de neutrones.



La metodología propuesta abre diversas líneas de investigación futura. Entre las más relevantes se encuentran: la incorporación del filtro adaptado (*matched filter*) como etapa de preprocesamiento y el análisis de su efecto sobre el rendimiento del sistema; el desarrollo de modelos de regresión de mayor complejidad que permitan estimar, además de la masa del chirp, otros parámetros físicos del evento como los spins de los objetos o su localización en el cielo; y la exploración de arquitecturas más profundas que mejoren la reconstrucción en el régimen de baja masa y bajo SNR.



Referencias

- Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M. R., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., Adhikari, R. X., Adya, V. B., Affeldt, C., Agathos, M., Agatsuma, K., Aggarwal, N., Aguiar, O. D., Aiello, L., Ain, A., Ajith, P., ... Zweizig, J. (2016). Observation of Gravitational Waves from a Binary Black Hole Merger. *Physical Review Letters*, 116(6), 061102. <https://doi.org/10.1103/PhysRevLett.116.061102>
- Abbott, B. P., Abbott, R., Abbott, T. D., Abraham, S., Acernese, F., Ackley, K., Adams, C., Adya, V. B., Affeldt, C., Agathos, M., Agatsuma, K., Aggarwal, N., Aguiar, O. D., Aiello, L., Ain, A., Ajith, P., Alford, T., Allen, G., Allocca, A., ... Zweizig, J. (2020). A guide to LIGO–Virgo detector noise and extraction of transient gravitational-wave signals. *Classical and Quantum Gravity*, 37(5), 055002. <https://doi.org/10.1088/1361-6382/ab685e>
- Abbott, R., Abbott, T. D., Acernese, F., Ackley, K., Adams, C., Adhikari, N., Adhikari, R. X., Adya, V. B., Affeldt, C., Agarwal, D., Agathos, M., Agatsuma, K., Aggarwal, N., Aguiar, O. D., Aiello, L., Ain, A., Ajith, P., Akcay, S., Akutsu, T., ... Zweizig, J. (2023). GWTC-3: Compact Binary Coalescences Observed by LIGO and Virgo during the Second Part of the Third Observing Run. *Physical Review X*, 13(4), 041039. <https://doi.org/10.1103/PhysRevX.13.041039>
- Ambardar, Ashok. (1999). *Analog and digital signal processing*. Brooks/Cole Pub. Co.
- Antonelli, A. (2022). Accurate waveform models for gravitational-wave astrophysics: Synergetic approaches from analytical relativity (No. arXiv:2211.01888). arXiv. <https://doi.org/10.48550/arXiv.2211.01888>
- Ashton, G., Hübner, M., Lasky, P. D., Talbot, C., Ackley, K., Biscoveanu, S., Chu, Q., Divakarla, A., Easter, P. J., Goncharov, B., Vivanco, F. H., Harms, J., Lower, M. E., Meadors, G. D., Melchor, D., Payne, E., Pitkin, M. D., Powell, J., Sarin, N., ... Thrane, E. (2019). Bilby: A User-friendly Bayesian Inference Library for Gravitational-wave Astronomy. *The Astrophysical Journal Supplement Series*, 241(2), 27. <https://doi.org/10.3847/1538-4365/ab06fc>
- Bayley, J., Messenger, C., & Woan, G. (2020). Robust machine learning algorithm to search for continuous gravitational waves. *Physical Review D*, 102(8), 083024.



- <https://doi.org/10.1103/PhysRevD.102.083024>
- Bengio, Y., Louradour, J., Collobert, R., & Weston, J. (2009). Curriculum learning. Proceedings of the 26th Annual International Conference on Machine Learning, 41–48. <https://doi.org/10.1145/1553374.1553380>
- Berti, E., et al. (2025). How Black Hole Spectroscopy Can Put General Relativity to the Test. *Physics Today*, 78(5), 32–37. <https://doi.org/10.1063/pt.fvtp.lpxx>
- Bishop, N. T., & Rezzolla, L. (2016). Extraction of Gravitational Waves in Numerical Relativity. *Living Reviews in Relativity*, 19(1), 2. <https://doi.org/10.1007/s41114-016-0001-9>
- Cutler, C., & Flanagan, É. É. (1994). Gravitational waves from merging compact binaries: How accurately can one extract the binary’s parameters from the inspiral waveform? *Physical Review D*, 49(6), 2658–2697. <https://doi.org/10.1103/PhysRevD.49.2658>
- DAMOUR, T. (2008). INTRODUCTORY LECTURES ON THE EFFECTIVE ONE BODY FORMALISM. *International Journal of Modern Physics A*, 23(08), 1130–1148. <https://doi.org/10.1142/S0217751X08039992>
- d’Inverno, R., & Vickers, J. (2022). *Introducing Einstein’s Relativity*. Oxford University Press Oxford. <https://doi.org/10.1093/oso/9780198862024.001.0001>
- Einstein, A. (1905). Zur Elektrodynamik bewegter Körper. *Annalen Der Physik*, 322(10), 891–921. <https://doi.org/10.1002/andp.19053221004>
- Einstein, A. (1916a). Die Grundlage der allgemeinen Relativitätstheorie. *Annalen Der Physik*, 354(7), 769–822. <https://doi.org/10.1002/andp.19163540702>
- Einstein, A. (1916b). Näherungsweise Integration der Feldgleichungen der Gravitation. *Sitzungsberichte Der Königlich Preussischen Akademie Der Wissenschaften*, 688–696. <https://ui.adsabs.harvard.edu/abs/1916SPAW.....688E/abstract>
- Einstein, A., & Rosen, N. (1937). On gravitational waves. *Journal of the Franklin Institute*, 223(1), 43–54. [https://doi.org/10.1016/S0016-0032\(37\)90583-0](https://doi.org/10.1016/S0016-0032(37)90583-0)
- Farret, F. A., Simoes, M. G., & Brandao, D. I. 2017, *Electronic Instrumentation for Distributed Generation and Power Processes* (1st ed.; CRC Press), <https://www.taylorfrancis.com/books/9781498782425>
- Feng, X., et al. (2015). Artificial Neural Networks Forecasting of PM2.5 Pollution Using Air Mass Trajectory Based Geographic Model and Wavelet Transformation. *Atmospheric Environ-*



- ment, 107, 118–28. <https://doi.org/10.1016/j.atmosenv.2015.02.030>
- Fernández, S., Padilla, Y., Guzmán Obregón, O. A., & Arbella, Y. (2013). Optimización de la Estimación de DOA en Sistemas de Antenas Inteligentes usando criterios de Redes Neuronales. *Ingeniería Electrónica, Automática y Comunicaciones*, 34(1), 70–86. <https://doi.org/10.1234/rielac.v34i1.154>
- Flanagan, É. É., & Hughes, S. A. (2005). The basics of gravitational wave theory. *New Journal of Physics*, 7, 204–204. <https://doi.org/10.1088/1367-2630/7/1/204>
- George, D., & Huerta, E. A. (2018a). Deep Learning for real-time gravitational wave detection and parameter estimation: Results with Advanced LIGO data. *Physics Letters B*, 778, 64–70. <https://doi.org/10.1016/j.physletb.2017.12.053>
- George, D., & Huerta, E. A. (2018b). Deep neural networks to enable real-time multimessenger astrophysics. *Physical Review D*, 97(4), 044039. <https://doi.org/10.1103/PhysRevD.97.044039>
- Géron, A. (2023). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (Third edition). O'Reilly.
- Greiner, W., Neise, L., & Stöcker, H. (1995a). The Laws of Thermodynamics. In *Thermodynamics and Statistical Mechanics* (pp. 33–61). Springer New York. https://doi.org/10.1007/978-1-4612-0827-3_2
- Greiner, W., Neise, L., & Stöcker, H. (1995b). Ensemble Theory and Microcanonical Ensemble. In *Thermodynamics and Statistical Mechanics* (pp. 142–158). Springer New York. https://doi.org/10.1007/978-1-4612-0827-3_6
- Gröchenig, K. (2001). *Foundations of Time-Frequency Analysis*. Birkhäuser Boston. <https://doi.org/10.1007/978-1-4612-0003-1>
- Hampshire, D. P. (2018). A derivation of Maxwell's equations using the Heaviside notation. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2134), 20170447. <https://doi.org/10.1098/rsta.2017.0447>
- Heaviside, O. (1893). A gravitational and electromagnetic analogy. In *Electromagnetic Theory*. Cambridge University Press. <https://doi.org/10.1017/CB09781139058049>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>



- Hulse, R. A., & Taylor, J. H. (1975). Discovery of a pulsar in a binary system. *The Astrophysical Journal*, 195, L51. <https://doi.org/10.1086/181708>
- IBM. (2021). What is a neural network? Retrieved April 26, 2025, from <https://www.ibm.com/think/topics/neural-networks>
- Jeon, H., et al. (2020). Area-Efficient Short-Time Fourier Transform Processor for Time-Frequency Analysis of Non-Stationary Signals. *Applied Sciences*, 10(20), 7208. <https://doi.org/10.3390/app10207208>
- Karim, A. A., et al. (2020). Modified Particle Swarm Optimization With Effective Guides. *IEEE Access*, 8, 188699–725. <https://doi.org/10.1109/ACCESS.2020.3030950>
- Karnam, N. K., Dubey, S. R., Turlapaty, A. C., & Gokaraju, B. (2022). EMGHand-Net: A hybrid CNN and Bi-LSTM architecture for hand activity classification using surface EMG signals. *Biocybernetics and Biomedical Engineering*, 42(1), 325–340. <https://doi.org/10.1016/j.bbe.2022.02.005>
- Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. *Proceedings of ICNN'95 - International Conference on Neural Networks*, 4, 1942–1948. <https://doi.org/10.1109/ICNN.1995.488968>
- Khan, S., Husa, S., Hannam, M., Ohme, F., Pürrer, M., Forteza, X. J., & Bohé, A. (2016). Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era. *Physical Review D*, 93(4), 044007. <https://doi.org/10.1103/PhysRevD.93.044007>
- Kumaresan, R., & Panchal, N. (2010). Encoding Bandpass Signals Using Zero/Level Crossings: A Model-Based Approach. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(1), 17–33. <https://doi.org/10.1109/TASL.2009.2022432>
- LeCun, Y., et al. (1998). Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, 86(11), 2278–324. <https://doi.org/10.1109/5.726791>
- LIGO Lab. (n.d.). FAQ. Retrieved March 23, 2025, from <https://www.ligo.caltech.edu/page/faq>
- LIGO Lab. (n.d.). What are gravitational waves? Retrieved April 24, 2025, from <https://www.ligo.caltech.edu/page/what-are-gw>
- Lovelace, R. V. E., & Tyler, G. L. (2012). On the discovery of the period of the Crab Nebula pulsar.



- The Observatory, 132(3), 186–188.
- Madisetti, V. K. (Ed.). (2010). The digital signal processing handbook: Three volume set (2. ed). CRC Press.
- Marx, E., Benoit, W., Gunny, A., Omer, R., Chatterjee, D., Venterea, R. C., Wills, L., Saleem, M., Moreno, E., Raikman, R., Govorkova, E., Rankin, D., Coughlin, M. W., Harris, P., & Katsavounidis, E. (2024). A machine-learning pipeline for real-time detection of gravitational waves from compact binary coalescences. arXiv. <https://doi.org/10.48550/ARXIV.2403.18661>
- Alex Nitz, Ian Harry, Duncan Brown, Christopher M. Biwer, Josh Willis, Tito Dal Canton, Collin Capano, Thomas Dent, Larne Pekowsky, Gareth S Cabourn Davies, Soumi De, Miriam Cabero, Shichao Wu, Andrew R. Williamson, Bernd Machenschalk, Duncan Macleod, Francesco Pannarale, Prayush Kumar, Steven Reyes, ... Arthur Tolley. (2024). Gwastro/pycbc: V2. 3. 3 release of pycbc [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.10473621>
- Poincaré, H. (1905). Sur la dynamique de l'électron. Proc. Acad. Sci., 140, 1504–1508.
- Polikar, R. (1996). The Wavelet Tutorial. <https://web.iitd.ac.in/~sumeet/WaveletTutorial.pdf>
- Qiu, R., Krastev, P. G., Gill, K., & Berger, E. (2023). Deep learning detection and classification of gravitational waves from neutron star-black hole mergers. Physics Letters B, 840, 137850. <https://doi.org/10.1016/j.physletb.2023.137850>
- Reza, M. H., et al. (2026). A Comprehensive Review of Convolutional Neural Networks: Foundations, Enhancements and Applications. Neural Computing and Applications, 38(4), 56. <https://doi.org/10.1007/s00521-025-11827-w>
- Robinson, J., & Rahmat-Samii, Y. (2004). Particle Swarm Optimization in Electromagnetics. IEEE Transactions on Antennas and Propagation, 52(2), 397–407. <https://doi.org/10.1109/TAP.2004.823969>
- Schutz, B. F. (2018). A First Course in General Relativity (2nd ed.). Cambridge University Press.
- Sejdić, E., Djurović, I., & Stanković, L. (2008). Quantitative Performance Analysis of Scalogram as Instantaneous Frequency Estimator. IEEE Transactions on Signal Processing, 56(8), 3837–3845. <https://doi.org/10.1109/TSP.2008.924856>
- Shannon, C. E. (1948). A Mathematical Theory of Communication. Bell System Technical Journal,



- 27(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Shen, H., George, D., Huerta, Eliu. A., & Zhao, Z. (2019). Denoising Gravitational Waves with Enhanced Deep Recurrent Denoising Auto-encoders. ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 3237–3241. <https://doi.org/10.1109/ICASSP.2019.8683061>
- Smith, S. W. (1997). The scientist and engineer's guide to digital signal processing (1. ed). California Technical Publ.
- SOUNDS OF SPACETIME. (n.d.). The basics of binary coalescence. Retrieved April 24, 2025, from <https://www.soundsofspacetime.org/the-basics-of-binary-coalescence.html>
- Spera, M., Trani, A. A., & Mencagli, M. (2022). Compact Binary Coalescences: Astrophysical Processes and Lessons Learned. *Galaxies*, 10(4), 76. <https://doi.org/10.3390/galaxies10040076>
- Subramanian, S., Krishnamoorthy, B., & Murthy, P. (2025). Dynamic learning rate scheduling based on loss changes leads to faster convergence. arXiv. <https://doi.org/10.48550/ARXIV.2512.14527>
- The Lifesciences Magazine. (n.d.). Power Of Convolutional Neural Networks In Modern AI. Retrieved May 15, 2024, from The Lifesciences Magazine website.
- Ulyanov, D., Vedaldi, A., & Lempitsky, V. (2020). Deep image prior. *International Journal of Computer Vision*, 128(7), 1867–1888. <https://doi.org/10.1007/s11263-020-01303-4>
- UNAM. (n.d.). Transformada Wavelet – acervo para el mejoramiento del aprendizaje de alumnos de ingeniería, en Inteligencia Artificial. Retrieved April 25, 2025, from https://virtual.cuautitlan.unam.mx/intar/?page_id=1108
- Ussipov, N., Zhanabaev, Z., Akhmetali, A., Zaidyn, M., Turlykozhaeva, D., Akniyazova, A., & Namazbayev, T. (2024). Classification of Gravitational Waves from Black Hole-Neutron Star Mergers with Machine Learning. *Journal of Astronomy and Space Sciences*, 41(3), 149–158. <https://doi.org/10.5140/JASS.2024.41.3.149>
- Vadapalli, P. (2025). Basic CNN Architecture: A Detailed Explanation of the 5 Layers in Convolutional Neural Networks. upGrad Blog. Retrieved May 15, 2024.
- Vajente, G., Huang, Y., Isi, M., Driggers, J. C., Kissel, J. S., Szczepańczyk, M. J., & Vitale, S. (2020). Machine-learning nonstationary noise out of gravitational-wave detectors. *Physical*



- Review D, 101(4), 042003. <https://doi.org/10.1103/PhysRevD.101.042003>
- Valens, C. (1999). A Really Friendly Guide to Wavelets. <http://agl.cs.unm.edu/williams/cs530/arfgtw.pdf>
- Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P.-A. (2008). Extracting and composing robust features with denoising autoencoders. Proceedings of the 25th International Conference on Machine Learning - ICML '08, 1096–1103. <https://doi.org/10.1145/1390156.1390294>
- Viswanathan, M. (2019). Digital Modulations Using Python.
- Weber, J. (1970). Anisotropy and Polarization in the Gravitational-Radiation Experiments. *Physical Review Letters*, 25(3), 180–184. <https://doi.org/10.1103/PhysRevLett.25.180>
- Weisberg, J. M., Taylor, J. H., & Fowler, L. A. (1981). Gravitational Waves from an Orbiting Pulsar. *Scientific American*, 245(4), 74–83.
- Weiss, R. (2022). Republication of: Electromagnetically coupled broadband gravitational antenna. *General Relativity and Gravitation*, 54(11), 153. <https://doi.org/10.1007/s10714-022-03021-3>
- White, T. (2020, November 8). Key Concepts: Fourier Transforms and Signal Processing. MWM Research Group. <https://www.mwmresearchgroup.org/blog/key-concepts-fourier-transforms-and-signal-pr>
- Zhang, J., He, T., Sra, S., & Jadbabaie, A. (2019). Why gradient clipping accelerates training: A theoretical justification for adaptivity. *arXiv*. <https://doi.org/10.48550/ARXIV.1905.11881>



Curriculum Vitae

Carlos Patricio Castañeda Bustamante nació en Chihuahua, Chihuahua, México. En 2022 obtuvo el grado de Licenciado en Ingeniería Física por la Universidad Autónoma de Chihuahua, y en 2026 obtuvo el grado de Maestro en Ingeniería en Computación por la misma institución, con orientación hacia la física computacional y la aplicación de aprendizaje automático al análisis de ondas gravitacionales.

Es coautor del artículo “Evaluating machine learning models for near earth objects classification”, en prensa en *Hybrid Intelligent System Design by Fusion of Soft Computing Models* (Springer, <https://doi.org/10.1007/978-3-032-31886-2>). Actualmente prepara un segundo artículo derivado de su tesis de maestría, sobre un modelo de aprendizaje automático de tres etapas para la detección y caracterización de ondas gravitacionales.

De diciembre de 2019 a 2024 trabajó en Creative Circle Films, primero como editor y posteriormente como gerente de proyecto, donde supervisó los procesos de edición, corrección de color y producción, y coordinó equipos de edición y producción audiovisual.

Correo Electrónico: patrickbustamante61@gmail.com

Esta tesis/disertación fue mecanografiada por Carlos Patricio Castañeda Bustamante.