

UNIVERSIDAD NACIONAL DEL LITORAL



DOCTORADO EN INGENIERÍA

**Procesamiento de señales
basado en funciones de forma de onda:
contribuciones algorítmicas y
aplicaciones biomédicas**

Joaquin Victorio Ruiz

FICH
FACULTAD DE INGENIERÍA Y CIENCIAS HÍDRICAS

INTEC
INSTITUTO DE DESARROLLO TECNOLÓGICO PARA LA INDUS-
TRIA QUÍMICA

CIMEC
CENTRO DE INVESTIGACIÓN DE MÉTODOS COMPUTACIONALES

sinc(i)
INSTITUTO DE INVESTIGACIÓN EN SEÑALES, SISTEMAS E IN-
TELIGENCIA COMPUTACIONAL



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas
Instituto de Desarrollo Tecnológico para la Industria Química
Centro de Investigación de Métodos Computacionales
Instituto de Investigación en Señales, Sistemas e Inteligencia
Computacional

**PROCESAMIENTO DE SEÑALES BASADO
EN FUNCIONES DE FORMA DE ONDA:
CONTRIBUCIONES ALGORÍTMICAS Y
APLICACIONES BIOMÉDICAS**

Joaquín Victorio Ruiz

Tesis remitida al Comité Académico del Doctorado
como parte de los requisitos para la obtención
del grado de
DOCTOR EN INGENIERIA
Mención Inteligencia Computacional, Señales y Sistemas
de la
UNIVERSIDAD NACIONAL DEL LITORAL

2025

Secretaría de Posgrado, Facultad de Ingeniería y Ciencias Hídricas, Ciudad
Universitaria, Paraje "El Pozo", S3000, Santa Fe, Argentina



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas
Instituto de Desarrollo Tecnológico para la Industria Química
Centro de Investigación de Métodos Computacionales
Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional

PROCESAMIENTO DE SEÑALES BASADO EN FUNCIONES DE FORMA DE ONDA: CONTRIBUCIONES ALGORÍTMICAS Y APLICACIONES BIOMÉDICAS

Joaquín Victorio Ruiz

Lugar de Trabajo:

IBB

Instituto de Investigación y Desarrollo en

Bioingeniería y Bionformática

Facultad de Ingeniería,

Universidad Nacional de Entre Ríos

Director:

Marcelo Alejandro Colominas

IBB-CONICET-UNER

Co-director:

Leandro Vignolo

sinc(i)-CONICET-UNL

Jurado Evaluador:

Pedro Arini

IAM - FI-UBA

Marcelo Risk

CONICET

Leandro Di Persia

sinc(i)-CONICET-UNL

UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas

Santa Fe, 31 de julio de 2025

Como miembros del Jurado Evaluador de la Tesis de Doctorado en Ingeniería titulada *“Procesamiento de señales basado en funciones de forma de onda: contribuciones algorítmicas y aplicaciones biomédicas”*, desarrollada por el Bioing. Joaquín RUIZ, en el marco de la Mención “Inteligencia Computacional, Señales y Sistemas”, certificamos que hemos evaluado la Tesis y recomendamos que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

La aprobación final de esta disertación estará condicionada a la presentación de la versión digital final de la Tesis ante el Comité Académico del Doctorado en Ingeniería.

Dr. Pedro Arini

Dr. Leandro Di Persia

Dr. Marcelo Risk

Santa Fe, 31 de julio de 2025

Certifico haber leído la Tesis, preparada bajo mi dirección en el marco de la Mención “Inteligencia Computacional, Señales y Sistemas” y recomiendo que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

.....
Dr. Leandro Vignolo
Codirector de Tesis

.....
Dr. Marcelo Alejandro Colominas
Director de Tesis



Universidad Nacional del Litoral
Facultad de Ingeniería y
Ciencias Hídricas

Secretaría de Posgrado

Ciudad Universitaria
C.C. 217
Ruta Nacional N° 168 - Km. 472,4
(3000) Santa Fe
Tel: (54) (0342) 4575 229
Fax: (54) (0342) 4575 224
E-mail: posgrado@fich.unl.edu.ar

Para mis padres, Rosa y Miguel Ángel,

Para mi tía del corazón, Alicia,

Para el amor de mi vida, Diamela.

AGRADECIMIENTO

En primer lugar, quiero agradecer a Marcelo, quien fue mi director de tesis y guía en esta etapa de mi trayectoria académica. Llegar hasta aquí fue posible, en gran medida, gracias a su paciencia, buena predisposición y valioso aporte a las actividades que dieron como resultado este documento de tesis. Asimismo, quiero agradecer a Leandro, mi codirector, y a Gastón, mi director de beca, quienes también supieron acompañarme en este proceso y participar de las valiosas discusiones que permitieron avanzar con los desarrollos aquí plasmados.

Quiero agradecer también a mis compañeros de las cátedras de Calculo Vectorial y Ecuaciones Diferenciales, de quienes aprendí mucho y ayudaron a cultivar mi creciente pasión por la enseñanza y el aprendizaje.

Por otro lado, ya desde el ámbito personal, quiero agradecer a mis compañeros de laboratorio: Iván, Felipe, Gabriel, Juan y Ramiro. Con su calidez humana y amplio conocimiento, hicieron este recorrido mucho más ameno y enriquecedor.

Sin duda, llegar a este punto no habría sido posible sin el apoyo incondicional de mi familia. En particular, mis padres, Rosa y Miguel Ángel, quienes siempre me apoyaron en mis proyectos personales y me inculcaron los valores de la curiosidad, la templanza y la perseverancia. Mi tía Alicia, quien es como una segunda madre para mí, la cual siempre me cuidó y veló por mi bienestar. A mi gran compañera, Diamela, por su paciencia, comprensión y acompañamiento durante cada etapa de nuestra vida juntos.

A mis amigos, tanto dentro como fuera de la universidad, por ser mi cable a tierra durante los altibajos de la carrera. Son muchos para nombrar y, tal vez, olvide a algunos nombres pero a cada uno de ustedes: Gracias.

Finalmente, quiero agradecer a la Universidad Pública, Gratuita y de Calidad, siendo todo esto posible gracias a su existencia.

ÍNDICE GENERAL

ÍNDICE DE TABLAS	VII
ÍNDICE DE FIGURAS	X
Resumen	XI
Abstract	XII
1 Introducción	1
1.1 Sistemas dinámicos y señales no estacionarias	1
1.1.1 Sistemas dinámicos en el ámbito biomédico	3
1.1.2 Señales no estacionarias en el ámbito biomédico	4
1.2 Análisis y procesamiento de señales no estacionarias	5
1.2.1 Análisis espectral	8
1.3 Análisis tiempo-frecuencia	9
1.3.1 Transformada de Fourier de Tiempo Corto	9
1.3.2 Transformada Ondita Continua	11
1.3.3 Representaciones tiempo-frecuencia de señales no estacionarias	12
1.3.4 Posprocesamiento de representaciones tiempo-frecuencia	15
1.4 Modelado fenomenológico de señales no estacionarias	16
1.4.1 Modelo armónico adaptativo	16
1.4.2 Modelo no armónico adaptativo	18
1.4.3 Funciones de forma de onda	19
1.4.4 Discretización y caracterización de formas de onda fijas	20
1.5 Desafíos actuales y propuestas	22
1.6 Objetivos de la tesis	22
1.6.1 Objetivos generales	22
1.6.2 Objetivos específicos	23
1.7 Estructura de la tesis	23
2 Estimación adaptativa del número de armónicos para WSF	25
2.1 Selección de modelos	26
2.2 Selección de modelos de regresión trigonométrica	27
2.2.1 Riesgo sin sesgo	28
2.2.2 Validación Cruzada Generalizada	29
2.2.3 Criterio de Wang	29
2.2.4 Criterio de Kavalieris-Hannan	30
2.3 Algoritmo propuesto para reconstrucción mediante ANHM con orden óptimo	30
2.4 Validación en señales sintéticas	32
2.4.1 Análisis de los criterios de selección de modelos	33
2.4.2 Estimación automática del orden óptimo de la WSF.	33

2.4.3	Estimación del orden del modelo en señales multicomponentes . . .	37
2.5	Aplicación 1: limpieza de ruido en señales de onda de pulso	39
2.5.1	Diseño experimental	40
2.5.2	Resultados de limpieza de ruido	43
2.5.3	Análisis estadístico de los resultados	45
2.6	Aplicación 2: limpieza de ruido en señales cardíacas y respiratorias	46
2.7	Comentarios de fin de capítulo	49
3	Modelo no armónico adaptativo con WSFs variantes en el tiempo	55
3.1	Funciones de forma de onda variables en el tiempo	57
3.2	Métodos de análisis de WSFs variantes en el tiempo.	58
3.2.1	<i>Multiresolution Mode Decomposition</i> (MMD)	58
3.2.2	<i>Shape Adaptive Mode Decomposition</i> (SAMD)	59
3.2.3	Modelo Oscilatorio de Funciones de Forma de Onda	60
3.3	Propuesta para el modelo no armónico adaptativo con formas de onda variante en el tiempo	61
3.3.1	Algoritmo para extracción de WSF variable en el tiempo	62
3.4	Funciones de amplitud armónica	63
3.4.1	Codificación de funciones de forma de onda variantes en el tiempo	63
3.4.2	Sobre la monotonicidad y concavidad de las HAFs	64
3.5	Implementación numérica	64
3.5.1	Efectos de borde y extensión de la señal	65
3.5.2	Estimación del número óptimo de nodos	67
3.5.3	Inicialización del vector de coeficientes	68
3.5.4	Estimación de los coeficientes del modelo	68
3.5.5	Ejemplo de aplicación	69
3.6	Validación en señales sintéticas	70
3.6.1	Caracterización de la WSF variante en el tiempo	71
3.6.2	Limpieza de ruido en señales con WSF variable en el tiempo . . .	72
3.6.3	Descomposición de señales con WSFs variables en el tiempo . . .	75
3.6.4	Detección de instante de cambio en la WSF	76
3.7	Aplicaciones en señales biomédicas	79
3.7.1	Aplicación 1: Limpieza de ruido en señales biomédicas	79
3.7.2	Aplicación 2: Descomposición de señales multicomponente . . .	81
3.7.3	Aplicación 3: Caracterización de transiciones en señales de elec- trocardiografía	82
3.8	Comentarios de fin de capítulo	84
4	Imputación de datos faltantes en señales biomédicas no estacionarias	87
4.1	Imputación de datos faltantes en series temporales	87
4.2	Métodos para imputación de datos faltantes en series temporales	90
4.2.1	Mapa de retardos de Takens	91
4.2.2	Estimación por mínimos cuadrados	92
4.2.3	Descomposición modal dinámica	93
4.2.4	Descomposición modal dinámica extendida	93
4.2.5	Regresión por procesos gaussianos	94
4.2.6	Regresión ARIMA	94
4.2.7	Modelo TBATS	95

4.2.8	Búsqueda de coincidencias no lineales para TFRs ralas	95
4.2.9	Onditas localmente estacionarias	96
4.2.10	Suavizado no lineal de datos con huecos aleatorios y valores atípicos	96
4.3	Selección de los parámetros de imputación	97
4.3.1	Mapa de retardos de Takens	97
4.3.2	Estimación por mínimos cuadrados, DMD y EDMD	98
4.3.3	Regresión por procesos gaussianos	98
4.3.4	Regresión ARIMA	98
4.3.5	Modelo TBATS	99
4.3.6	Búsqueda de coincidencia no lineal para RTF ralas	99
4.3.7	Procesos ondita localmente estacionarios	100
4.3.8	Suavizado no lineal de datos con huecos aleatorios y valores atípicos	100
4.4	Algoritmo propuesto - HaLI	100
4.4.1	Etapa 1	102
4.4.2	Etapa 2	103
4.4.3	Etapa 3	104
4.5	Validación en señales sintéticas	105
4.6	Experimentos en señales reales	110
4.7	Comentarios de fin de capítulo	113
5	Análisis de la señal de voz mediante medidas de variabilidad de la WSF	117
5.1	Análisis de la señal de voz	118
5.2	Modelado no armónico adaptativo aplicado a señales de voz	119
5.3	Medidas de variabilidad de la forma de onda de la voz	120
5.4	Aplicación 1: tipificación automática de voces	122
5.4.1	Conjuntos de datos	124
5.4.2	Detalles del clasificador	125
5.4.3	Medidas de desempeño y optimización de hiperparámetros	125
5.4.4	Resultados y discusión	126
5.4.5	Comparación con el estado del arte	127
5.5	Aplicación 2: detección de enfermedad de Parkinson	128
5.5.1	Conjuntos de datos	129
5.5.2	Detalles del clasificador	130
5.5.3	Medidas de desempeño y optimización de hiperparámetros	130
5.5.4	Resultados y discusión	132
5.5.5	Comparación con el estado del arte	132
5.6	Comentarios de fin de capítulo	134
6	Conclusiones y trabajo a futuro	135
6.1	Conclusiones	135
6.2	Trabajos a futuro	137
	Acrónimos	140
	Notación	141
	LISTA DE PUBLICACIONES	
	BIBLIOGRAFÍA	

ÍNDICE DE TABLAS

2.1	Selección automática del orden del ANHM.	36
4.1	Métodos de imputación inicial considerados en los experimentos.	91
4.2	Resultados de imputación de datos faltantes en señales sintéticas.	109
4.3	Tiempo medio de cálculo de los distintos algoritmos para el experimento con señales sintéticas.	109
4.4	Resultados de imputación de datos faltantes en señales reales.	112
5.1	Distribución de los tipos de voz en ambos conjuntos de datos	125
5.2	Hiperparámetros óptimos de los clasificadores para cada base de datos. . .	125
5.3	Comparación de nuestros resultados con el estado del arte para la tarea de tipificación por voz.	127
5.4	Hiperparámetros óptimos de los 3 clasificadores para la tarea de detección de la enfermedad de Parkinson.	131
5.5	Comparación de nuestra propuesta con el estado del arte en detección de PD.	132

ÍNDICE DE ILUSTRACIONES

1.1	Señales no estacionarias en el ámbito biomédico.	5
1.2	Señal no estacionaria en el dominio temporal.	8
1.3	Señal no estacionaria en el dominio frecuencial.	10
1.4	Señal no estacionaria en el dominio tiempo-frecuencia.	14
1.5	Modelo armónico adaptativo.	17
1.6	Modelo no armónico adaptativo.	19
2.1	Criterios de selección de modelos en señales no estacionarias ($r_0 = 4$). . .	32
2.2	Criterios de selección de modelos en señales no estacionarias ($r_0 = 9$). . .	34
2.3	Espectrogramas de señales no estacionarias contaminadas con ruido. . . .	35
2.4	Espectrograma de una señal no estacionaria multicomponente.	38
2.5	Resultado de la estimación del orden del ANHM en señales multicomponente ($\text{SNR}_{in} = 0$ dB).	39
2.6	Resultado de la estimación del orden del ANHM en señales multicomponente ($\text{SNR}_{in} = 10$ dB).	40
2.7	Resultado de la estimación del orden del ANHM en señales multicomponente ($\text{SNR}_{in} = 20$ dB).	41
2.8	Señal de onda de pulso simulada.	42
2.9	Resultados de limpieza de ruido para señales de onda de pulso.	43
2.10	Resultados de estimación de orden óptimo en señales de onda de pulso. . .	44
2.11	Estudio comparativo del algoritmo propuesto en señales de onda de pulso. .	45
2.12	Espectrogramas de las señales de electrocardiografía de la base de datos Fantasia.	47
2.13	Resultados de reconstrucción de funciones de forma de onda para la base de datos Fantasia.	48
2.14	Resultados de reconstrucción de funciones de forma de onda para los 10 primeros sujetos de la cohorte de adultos mayores.	50
2.15	Resultados de reconstrucción de funciones de forma de onda para los 10 primeros sujetos de la cohorte de jóvenes.	51
2.16	Resultados de reconstrucción de funciones de forma de onda para los 10 últimos sujetos de la cohorte de adultos mayores.	52
2.17	Resultados de reconstrucción de funciones de forma de onda para los 10 últimos sujetos de la cohorte de jóvenes.	53
3.1	Ejemplos de señales biomédicas no estacionarias con forma de onda variable en el tiempo.	56
3.2	Demostración del efecto del interpolador sobre la estimación de las HAFs. . .	65
3.3	Diagrama de flujo del algoritmo adaptativo de extracción de WSFs variables en el tiempo.	66

ÍNDICE DE ILUSTRACIONES

3.4	Demostración de los efectos de borde en la estimación de las HAFs.	67
3.5	Demostración del método de estimación automática del número de nodos óptimos para interpolación de las HAFs.	69
3.6	Resultados del método de estimación adaptativo de WSFs variantes en el tiempo.	71
3.7	Limpieza de ruido para señales monocomponente con WSFs variantes en el tiempo.	73
3.8	Comparación de las formas de onda extraídas con los diferentes métodos.	74
3.9	Limpieza de ruido y descomposición de señales multicomponente con WSFs variantes en el tiempo.	77
3.10	Detección de instante de cambio de la WSF.	78
3.11	Limpieza de ruido en señales de electroencefalografía.	80
3.12	Descomposición de señal de neumografía por impedanciometría.	82
3.13	Reconstrucción de la señal de electrocardiografía durante fibrilación ventricular.	83
3.14	Caracterización de transiciones en la forma de onda de la señal de electrocardiografía.	84
4.1	Ejemplos de señales con datos faltantes en el ámbito biomédico.	89
4.2	Método propuesto para imputación de datos faltantes basado en la interpolación de la descomposición armónica de señales no estacionarias.	101
4.3	Ejemplo de señal sintética con datos faltantes.	105
4.4	Resultados de imputación de datos faltantes en señales sintéticas.	107
4.5	Desempeño de los métodos de imputación inicial para señales sintéticas.	108
4.6	Ejemplos de las diferentes señales biomédicas consideradas para el experimento de imputación de datos faltantes en señales reales.	111
4.7	Resultados de imputación de datos faltantes en señales reales.	113
4.8	Comparación entre la imputación inicial de datos faltantes antes y después del posprocesamiento.	114
4.9	Desempeño de los métodos de imputación inicial para señales reales.	115
5.1	Ejemplo de análisis de una señal de voz utilizando el algoritmo tvWSE.	121
5.2	Diagrama de los pasos de extracción de características y clasificación de la tarea de tipificación de la voz.	123
5.3	Señales de voces según la tipificación de Titze.	124
5.4	Resultados de la tarea de tipificación de voces.	126
5.5	Señales de voz para detección de Enfermedad de Parkinson (PD).	129
5.6	Resultados de detección de enfermedad de Parkinson.	131

PROCESAMIENTO DE SEÑALES BASADO EN FUNCIONES DE FORMA DE ONDA: CONTRIBUCIONES ALGORÍTMICAS Y APLICACIONES BIOMÉDICAS

Resumen

El análisis de sistemas dinámicos puede realizarse mediante el estudio de las señales no estacionarias que estos generan, lo cual es relevante en ingeniería biomédica por su relación con fenómenos fisiológicos y patológicos. Estas señales presentan variaciones temporales complejas, especialmente en el caso de señales multicomponente, donde cada componente tiene variaciones propias de amplitud, frecuencia y forma de onda.

Los modelos matemáticos, especialmente fenomenológicos, resultan útiles para describir estas magnitudes variantes en el tiempo, ofreciendo un marco teórico versátil. Sin embargo, los enfoques tradicionales, como el análisis espectral basado en la transformada de Fourier, no son adecuados para señales no estacionarias. En este contexto, los métodos tiempo-frecuencia (T-F) cobran relevancia ya que permiten representar la evolución temporal de las componentes frecuenciales de la señal. Entre los modelos que aprovechan estas representaciones, el *modelo no armónico adaptativo* (ANHM) destaca por describir señales mediante patrones oscilatorios modulados en amplitud y frecuencia, caracterizados por su *función de forma de onda* (WSF). Aunque la utilidad del ANHM y de las WSF ha sido demostrada en diversas aplicaciones, particularmente en el ámbito biomédico, aun resta explorar adecuadamente aspectos metodológicos para poder aplicar exitosamente estas técnicas al análisis de señales en un contexto más amplio. Algunos de los aspectos más salientes son la determinación automática del número de armónicos óptimos para representar la WSF y la caracterización de WSFs variables en el tiempo.

Para abordar este primer aspecto, en esta tesis se propuso el uso de criterios de selección de modelos trigonométricos para determinar automáticamente el número óptimo de armónicos para la representación de WSF, aplicando estos criterios a señales simuladas y reales, tanto monocomponente como multicomponente. Esto permitió desarrollar un algoritmo basado en el ANHM capaz de adaptarse adecuadamente a señales fisiológicas incluso ante la presencia de ruido e interferencias.

Posteriormente, se diseñó un método para estimar formas de onda variables en el tiempo a partir de una versión generalizada del ANHM, lo cual mejora el desempeño en tareas como eliminación de ruido, descomposición y detección de cambios oscilatorios. Este enfoque mostró ser superior a métodos establecidos en el estado del arte.

Por último, en esta tesis también se abordaron algunas aplicaciones específicas. En primer lugar se desarrolló un algoritmo basado en la descomposición armónica para la imputación de datos faltantes en series temporales, el cual permite mejorar los resultados en señales sintéticas y fisiológicas. Por otro lado, se propuso una serie de medidas de variabilidad de forma de onda para señales no estacionarias de vocales sostenidas obtenidas a partir del modelo ANHM con WSF variante en el tiempo. Estas medidas permitieron obtener características capaces de discriminar señales según la variación ciclo a ciclo de su forma de onda. Estas características se emplearon en la tipificación de voces y la detección de enfermedad de Parkinson, con resultados competitivos frente al estado del arte.

SIGNAL PROCESSING USING
WAVE-SHAPE FUNCTIONS:
ALGORITHMIC CONTRIBUTIONS AND
BIOMEDICAL APPLICATIONS

Abstract

The analysis of dynamic systems can be performed by studying the non-stationary signals they generate, which is relevant in biomedical engineering due to their relationship with physiological and pathological phenomena. These signals exhibit complex temporal variations, especially in the case of multicomponent signals, where each component has its own amplitude, frequency and waveform variations.

Mathematical models, especially phenomenological ones, are useful for describing these time-varying quantities, providing a versatile theoretical framework. However, traditional approaches, such as spectral analysis based on the Fourier transform, are not suitable for non-stationary signals. In this context, time-frequency (T-F) methods become relevant as they allow for the representation of the time evolution of the frequency components of the signal. Among the models that take advantage of these representations, the adaptive non-harmonic model (ANHM) stands out because it describes signals by means of oscillatory patterns modulated in amplitude and frequency, characterized by their wave-shape function (WSF). Although the usefulness of ANHM and WSF has been demonstrated in several applications, particularly in the biomedical field, methodological aspects still need to be adequately explored in order to enable the successful application of these techniques to signal analysis in a broader context. Some of the most salient aspects are the automatic determination of the optimal number of harmonics to represent the WSF and the characterization of time-varying WSFs.

To address this first aspect, this thesis proposes the use of trigonometric model selection criteria to automatically determine the optimal number of harmonics for WSF representation, applying these criteria to simulated and real signals, both single and multicomponent. This enabled the development of an algorithm that is capable of adequately adapting to physiological signals even in the presence of noise and interference.

Subsequently, a method was designed to estimate time-varying waveforms from a generalized version of the ANHM, which improves performance in tasks such as denoising, decomposition and oscillatory change detection. This approach was shown to be superior to established state-of-the-art methods.

Finally, this thesis also addresses a few specific applications. First, an algorithm based on harmonic decomposition was developed for the imputation of missing data in time series, which improved performance in synthetic and physiological signals. On the other hand, a series of waveform variability measures were proposed for non-stationary sustained vowel signals obtained from the ANHM model with time-varying WSF. These measures yielded features capable of discriminating signals according to the cycle-to-cycle variation of their waveform. These features were used in voice typing and Parkinson's disease detection, achieving competitive results compared to the state of the art.

Capítulo 1

Introducción

1.1. Sistemas dinámicos y señales no estacionarias

En el estudio de los fenómenos físicos, es común llevar adelante una representación abstracta de los elementos involucrados en su generación. Estos elementos agrupados conforman *sistemas físicos* cuyo comportamiento se describe mediante formalismos matemáticos. En una primera instancia, se le suelen asignar a estos sistemas características o parámetros que son fijos, es decir, que el comportamiento del sistema queda totalmente determinado desde su concepción y no cambia a lo largo del tiempo. Se sabe, sin embargo, que los sistemas reales no cumplen con esta condición de ser fijos o *estáticos*, sino que debido a diferentes factores tanto externos como internos a los sistemas, las características de éstos cambian, lo cual les confiere la cualidad de ser variantes en el tiempo o *dinámicos*.

El estudio de los sistemas físicos dinámicos conlleva determinar cómo es la evolución temporal de sus características y cómo se relacionan estos cambios con los fenómenos del mundo real. Estos sistemas pueden estudiarse a partir del análisis de las señales que generan, las cuales serán no estacionarias debido a la dinámica de los sistemas. Asimismo, muchos sistemas generan señales oscilatorias, es decir, cuya amplitud fluctúa en torno a una media a lo largo del tiempo. Estos sistemas incluyen, entre otros, sistemas eléctricos, mecánicos y biológicos. La dinámica no lineal del sistema que genera la señal oscilatoria se manifiesta en variaciones temporales de diferentes aspectos de la señal, siendo los más relevantes la frecuencia, la amplitud y la morfología de la onda oscilatoria.

El cambio en la amplitud relativa entre los ciclos de las señales oscilatorias, también conocido como *modulación en amplitud*, es de gran interés cuando es necesario cuantificar la intensidad de la respuesta del sistema ante diferentes condiciones. Dicho cambio en amplitud puede darse de diferentes maneras, aunque en general se considera para las señales oscilatorias que el cambio se produce de manera “suave” y “lenta”. La primera condición establece que no deben darse cambios abruptos de amplitudes entre un ciclo y el otro. La condición de oscilación lenta implica que los cambios en la amplitud ciclo a ciclo de la señal no pueden ser más rápidos que la frecuencia de oscilación de la onda. El estudio del cambio en la amplitud de las oscilaciones ha sido una tarea de sumo interés en el ámbito del procesamiento de señales [6, 7, 8]. Una de las dificultades del análisis de la modulación en amplitud es que resulta difícil independizar los cambios de la amplitud local de cambios en la “tendencia” de la señal. En muchos casos, la señal de interés se encuentra “montada” sobre una variación lenta la cual suele asociarse a fuentes de ruido,

tanto intrínsecas como extrínsecas. Dicha tendencia o “media local” genera elevaciones y depresiones en los ciclos que pueden llegar a confundirse con cambios en la modulación en amplitud.

Tal vez un poco más interesante resulta el estudio de la variación temporal en la frecuencia de oscilación de la señal, conocida como *modulación en frecuencia*. Estos cambios se manifiestan como contracciones o dilataciones de la onda entre un ciclo y el otro. Dicho cambio, a diferencia de la modulación en amplitud, no se verá afectado por fuentes de ruido como tendencias o ruido de alta frecuencia. Además, los cambios en la duración local de la onda suelen estar asociados a fenómenos subyacentes del sistema que aportan información esencial acerca de su funcionamiento. Ejemplos claros de esto son los cambios en la frecuencia local en señales biomédicas, como el electrocardiograma (ECG) [9] o el electroencefalograma (EEG) [10]. Otro ámbito donde suele ser necesario analizar cambios de frecuencia es el estudio de señales musicales generadas por instrumentos [11, 12], como también la señal de voz humana producida durante el canto [13]. Sin buscar elaborar una lista exhaustiva, mencionamos que el estudio de la modulación en frecuencia es de gran relevancia en disciplinas tan diversas como la sismología [14, 15], astronomía [16, 17], dinámica rotacional [18, 19], y econometría [20, 21], entre otras.

Tanto el estudio de la modulación en amplitud como en frecuencia han sido de gran relevancia para el avance de disciplinas como el análisis y procesamiento de señales. Debido a esto, se han desarrollado diversos métodos y procedimientos para estimar estas magnitudes variantes en el tiempo. Un aspecto de las señales oscilatorias que no ha recibido tanta atención, al menos hasta los últimos años, es el análisis de los cambios en la morfología de las oscilaciones de las señales. Los métodos de análisis de señales no estacionarias suelen considerar, en sus formulaciones, señales compuestas por varios componentes. Dichas componentes suelen representarse mediante funciones periódicas simples, como senos y cosenos, moduladas en amplitud y en frecuencia. Sin embargo, se sabe que las señales reales presentan patrones oscilatorios más complicados que las ondas trigonométricas, y dicha *forma de onda* acarrea información importante acerca del estado del sistema generador de las señales. Los métodos tradicionales que contemplan la forma de onda de las señales suelen considerarla como un patrón que se repite de un ciclo al otro, estando ésta afectada por dilataciones-contracciones (modulación en frecuencia) y atenuaciones-amplificaciones (modulación en amplitud). Sin embargo, pueden observarse cambios en la forma que no quedan totalmente explicados por estas modulaciones, por lo cual resulta necesario contar con herramientas de análisis y procesamiento que contemplen no sólo la existencia de un patrón oscilatorio más general para las señales, sino que también pueda capturar la variación a lo largo del tiempo de dicho patrón.

En este capítulo, discutiremos aspectos teóricos y prácticos asociados al estudio de señales no estacionarias teniendo en cuenta su modulación en amplitud, modulación en frecuencia y patrón oscilatorio. Comenzaremos mencionando casos típicos de señales no estacionarias que se suelen encontrar en diferentes situaciones del ámbito biomédico. La condición de no estacionariedad es ubicua en el estudio de señales biomédicas, debido a las interacciones complejas que se dan entre los sistemas fisiológicos que las generan. Seguidamente, discutiremos el marco teórico del análisis de este tipo de señales, para el cual deberemos poder representar de manera conjunta la información temporal y frecuencial de la señales mediante herramientas de análisis tiempo-frecuencia. Seguidamente, mencionaremos la estrategia de modelado fenomenológico para señales no estacionarias, el cual será la base de los desarrollos presentados en esta tesis. En este aspecto, se discutirá

en detalle un modelo adaptativo que aplica el concepto de funciones de forma de onda para caracterizar los patrones oscilatorios de las señales no estacionarias. A partir de analizar este marco teórico, se discutirán algunas limitaciones existentes en la aplicación de los modelos basados en funciones de forma de onda en el estudio de señales no estacionarias del mundo real, lo cual motivará los diferentes aportes novedosos desarrollados en esta tesis. Finalmente, detallaremos los objetivos y la estructura del documento de tesis.

1.1.1. Sistemas dinámicos en el ámbito biomédico

La ingeniería biomédica es una disciplina avocada al estudio de los sistemas fisiológicos de los seres vivos. Más particularmente, casi la totalidad de esta disciplina se enfoca en el estudio del cuerpo humano. El cuerpo humano puede entenderse como un conjunto de sistemas interrelacionados entre sí. Los sistemas que conforman el organismo humano, como por ejemplo el sistema cardiovascular, el sistema respiratorio o el sistema nervioso, entre muchos otros, resultan ser sistemas dinámicos ya que su comportamiento cambia a lo largo del tiempo debido a diferentes procesos fisiológicos o patológicos que se producen en el organismo. Un ejemplo claro de esto son los cambios en la frecuencia cardíaca debido a efectos hormonales o ante la necesidad del corazón de bombear más sangre a las extremidades durante una actividad física de alto esfuerzo. La capacidad de cuantificar estos cambios temporales en la frecuencia de los latidos del corazón y poder hacer un “seguimiento” de este cambio son tareas de gran importancia en la clínica ya que permite diagnosticar enfermedades o detectar de manera preventiva posibles patologías.

El estudio de estos sistemas se lleva a cabo a través del *sensado* de diferentes magnitudes físicas que permiten el monitoreo continuo de las variables de interés. La secuencia de valores que otorgan estos sensores a lo largo del tiempo generan lo que conocemos como *señales*. En las últimas décadas, los sistemas de sensado electrónico se han vuelto mucho más sofisticados gracias a grandes avances tecnológicos como los dispositivos lógicos programables y técnicas de microelectrónica que han permitido fabricar sensores con mejores prestaciones y alta portabilidad. Esto ha provocado una “explosión” en la cantidad de información disponible para estudiar el organismo, llevando al desarrollo de sistemas capaces de monitorear la actividad de los seres vivos en circunstancias cada vez más cercanas a los casos reales.

En paralelo, el desarrollo de la informática ha sido aprovechado por los investigadores en las disciplinas biomédicas para desarrollar algoritmos de análisis y procesamiento de la información capturada por los sensores. Entre otros aspectos, los desarrollos recientes en materia de inteligencia artificial han permitido ampliar las fronteras de conocimiento en un sinnúmero de aplicaciones asociadas a la medicina, gracias a la mejora de los sistemas automáticos de toma de decisión y diagnóstico.

A su vez, estos desarrollos se han nutrido de avances y descubrimientos teóricos en disciplinas de las ciencias exactas, como la física y la matemática. Particularmente, el uso de definiciones y operaciones matemáticas para describir los fenómenos físicos permite definir un marco conceptual riguroso para su estudio. Además, se pueden aprovechar los conocimientos de las disciplinas matemáticas para estudiar estos algoritmos y dar garantías acerca de su funcionamiento y consistencia. En particular, el área del *análisis y procesamiento de señales biomédicas* busca aplicar algoritmos con fundamentos matemáticos sólidos a señales originadas en el cuerpo humano. Particularmente, en el ámbito biomédico es normal encontrarse con sistemas dinámicos y, por consiguiente, señales no

estacionarias cuya variación temporal permite describir y mejorar el entendimiento de los procesos fisiológicos y patológicos del organismo.

1.1.2. Señales no estacionarias en el ámbito biomédico

Las herramientas de análisis y procesamiento de señales no estacionarias son de especial interés en el ámbito biomédico debido a la gran cantidad de registros que pueden obtenerse del cuerpo que presentan características variantes en el tiempo. En la Fig. 1.1 se muestran algunos ejemplos de señales biomédicas no estacionarias. En primer lugar, en la parte superior de la figura se muestra una derivación ECG, donde se han marcado con los picos de mayor amplitud de cada ciclo, conocidos como picos R. La distancia entre dos picos R consecutivos se conoce como intervalo R-R. La variación ciclo a ciclo de la longitud de este intervalo se asocia con la variabilidad del ritmo cardíaco (VRC) y es de gran importancia clínica [9]. Debajo de la señal de ECG se gráfica la serie de intervalos R-R para dicha señal donde vemos que esta distancia es variante en el tiempo, lo cual pone de manifiesto la no estacionariedad de la señal y la dinámica no lineal del sistema biológico subyacente.

Una tarea de procesamiento de gran relevancia en el ámbito biomédico es lo que se conoce como *descomposición de señales*. La descomposición de señales parte de la suposición de que el registro bajo estudio está conformado por dos o más *componentes*, los cuales pueden estar asociados a fenómenos subyacentes distintos. El objetivo de esta tarea es, entonces, separar los diferentes componentes de la señal y obtener señales individuales que puedan ser analizadas de manera independiente. Es común encontrar señales biomédicas que contienen información asociada a dos o más sistemas del organismo. Un ejemplo de esto se muestra en la parte inferior izquierda de la Fig. 1.1 donde se muestra una señal de neumografía por impedanciometría, la cual mide la actividad respiratoria de un paciente mediante un sistema que registra el cambio temporal en la impedancia eléctrica de la cavidad torácica. Como se muestra en la figura, esta señal no sólo contiene información asociada a la actividad respiratoria sino que además presenta un componente asociado a la actividad del corazón, el cual se denomina comúnmente *artefacto cardiogénico*.

Por otro lado, las señales del mundo real suelen estar contaminadas con diferentes tipos de interferencias que reciben el nombre de *ruido* y su presencia suele dificultar o hasta impedir el análisis adecuado de la información de la señal de interés. La *limpieza de ruido* es una serie de técnicas de procesamiento de señales que buscan eliminar estas interferencias presentes en la señal. Los algoritmos de limpieza de ruido buscan eliminar este ruido, preservando la información asociada a la señal fisiológica. En el panel inferior derecho de la Fig. 1.1 se muestra un segmento de un registro de EEG de un neonato durante la manifestación de una convulsión epiléptica. Se nota claramente la aparición de un ritmo lento con una forma de onda particular, el cual se encuentra contaminado por un componente de *ruido* de alta frecuencia asociada a la actividad muscular.

De esta breve presentación resulta clara la gran diversidad de registros de interés para la ingeniería biomédica y las diferentes tareas de procesamiento que deben llevarse a cabo para garantizar la usabilidad de los datos. Comúnmente, se desarrollan herramientas específicas para cada tarea en particular, donde se aprovecha la información conocida *a priori* sobre los fenómenos que generan las señales y las fuentes de ruido que esperamos encontrar en los registros. Sin embargo, suele ser provechoso desarrollar herramientas de análisis y procesamiento que no se ajusten a una tarea específica, sino que permitan

1.2. ANÁLISIS Y PROCESAMIENTO DE SEÑALES NO ESTACIONARIAS

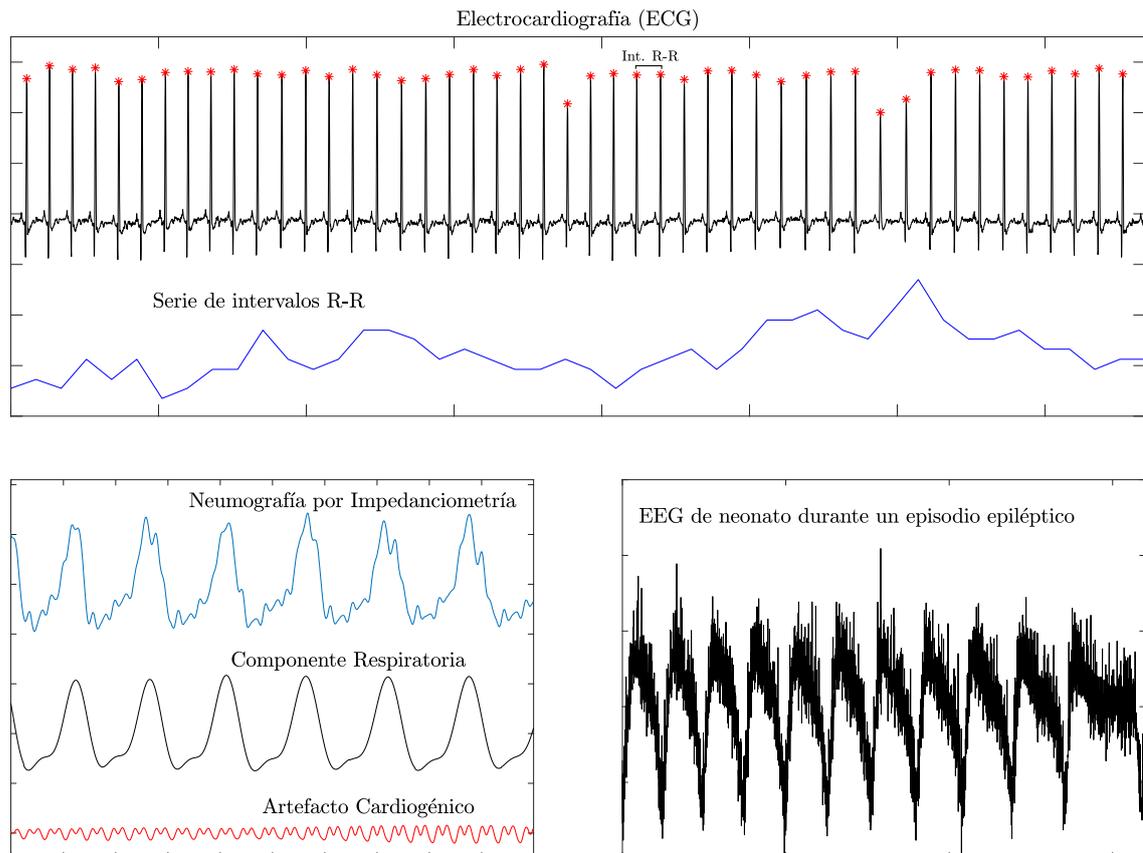


Figura 1.1: Señales no estacionarias en el ámbito biomédico. Primera fila: Señal de electrocardiografía (ECG) en negro, con los picos R indicados con asteriscos rojos, junto a su correspondiente serie de intervalos entre picos R-R (azul). Segunda fila a la izquierda: Señal de neumografía por impedanciometría (azul) junto a su componente respiratorio (negro) y el artefacto cardíaco (rojo) que la conforman. Segunda fila a la derecha: Señal de electroencefalografía (EEG) de un neonato durante un episodio epiléptico, donde se observa un claro componente de ruido de alta frecuencia asociada a la actividad muscular durante el episodio.

definir un marco de análisis más general y puedan ser aplicadas a una gran diversidad de problemas. En vista de esto, nos abocaremos en el resto de este capítulo a describir este tipo de herramientas sin establecer de forma previa las características de las señales que serán estudiadas. Nótese también que dichas herramientas pueden ser aplicadas a señales no estacionarias en disciplinas distintas a la ingeniería biomédica, aunque nos enfocaremos en dicha disciplina al tratar las aplicaciones de los desarrollos propuestos.

1.2. Análisis y procesamiento de señales no estacionarias

Como mencionamos en la sección anterior, la manera más sencilla de definir una señal oscilatoria no estacionaria es mediante una función coseno modulada en amplitud y frecuencia (AM-FM, por las siglas en inglés de *amplitude-modulated and frequency-modulated*) y que depende del tiempo:

$$x(t) = A(t) \cos(2\pi\phi(t)), \quad A(t) > 0, \quad \phi'(t) > 0 \quad \forall t, \quad (1.1)$$

donde $A(t)$ es la modulación en amplitud o *amplitud instantánea* y $\phi(t)$ es la modulación en fase o *fase instantánea*. A partir de esta última, es posible definir la magnitud asociada a la modulación en frecuencia o *frecuencia instantánea*. En su definición más clásica, la frecuencia instantánea se obtiene al derivar a la función de fase con respecto al tiempo: $f(t) = \phi'(t)$. Aunque esta definición resulta sencilla, se encuentra limitada por diferentes aspectos prácticos. En primer lugar, no existe una única forma de elegir las funciones $A(t)$ y $\phi(t)$ asociadas a una señal no estacionaria $x(t)$. Para confirmar esto, vemos que al considerar una función $0 < B(t) < 1$ y escribir la siguiente ecuación

$$x(t) = \frac{A(t)}{B(t)} B(t) \cos(2\pi\phi(t)), \quad (1.2)$$

podemos definir a $x(t)$ como $x(t) = \tilde{A}(t) \cos(2\pi\tilde{\phi}(t))$, siendo $\tilde{A}(t) = A(t)/B(t)$ y $\tilde{\phi}(t) = \frac{1}{2\pi} \arccos(B(t) \cos(2\pi\phi(t)))$. En consecuencia, la definición de la modulación en frecuencia $\tilde{f}(t) = \tilde{\phi}'(t)$ dependerá de la nueva definición para la modulación en fase $\tilde{\phi}(t)$.

Por otro lado, según la teoría de funciones analíticas, la frecuencia instantánea de una señal de la forma $x_a(t) = A_a(t)e^{i2\pi\phi_a(t)}$ puede obtenerse como

$$f_a(t) = \frac{d}{dt} \phi_a(t). \quad (1.3)$$

Donde $\phi_a(t) = (2\pi)^{-1} \angle x_a(t)$, es la fase instantánea de la función analítica. Es posible definir una señal analítica $\overline{x_a(t)}$ a partir de $x(t)$ utilizando la transformada de Hilbert:

$$\mathcal{H}\{x\}(t) = \frac{1}{\pi} \text{PV} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau, \quad (1.4)$$

donde PV indica que se toma el valor principal de Cauchy (PV) de la integral impropia ya que el núcleo de la transformada $h(t) = \frac{1}{\pi t}$ no está definido para $t = 0$. A partir de la transformada de Hilbert, la señal analítica se obtiene como: $x_a(t) = x(t) + i\mathcal{H}\{x\}(t)$ y, bajo ciertas condiciones, podemos considerar $A(t) \approx A_a(t)$ y $\phi(t) \approx \phi_a(t)$.

Estas relaciones son sólo aproximadas debido a dos resultados teóricos importantes. En primer lugar, el teorema de Bedrosian [22] nos dice que cuando los espectros de $A(t)$ y $\cos(2\pi\phi(t))$ son disjuntos y, además, el espectro de $A(t)$ ocupa una región de menor frecuencia, podemos hacer:

$$\mathcal{H}\{A(t) \cos(2\pi\phi(t))\} = A(t) \mathcal{H}\{\cos(2\pi\phi(t))\}. \quad (1.5)$$

Este resultado indica que si $A(t)$ varía más lentamente en el tiempo que $\cos(2\pi\phi(t))$, entonces sería posible estimar $A(t)$ a partir del módulo de $x_a(t)$ de manera exacta, siempre y cuando se cumpla que $\mathcal{H}\{\cos(2\pi\phi(t))\} = \sin(2\pi\phi(t))$. Sin embargo, esta condición no se cumple para cualquier $\phi(t)$.

El teorema de Nuttall [23] otorga una medida del error que se comete al considerar $\mathcal{H}\{\cos(2\pi\phi(t))\} = \sin(2\pi\phi(t))$ y se demuestra que este error está asociado a la energía de los componentes frecuenciales de frecuencia negativa que contiene el espectro de $z_x^q(t) = A(t)e^{i2\pi\phi(t)}$, la versión compleja en cuadratura de $x(t)$, a saber:

$$\int_{-\infty}^{\infty} |\mathcal{H}\{x(t)\} - x^q(t)|^2 dt = 2 \int_{-\infty}^0 |\hat{z}^q(f)|^2 df. \quad (1.6)$$

Aunque el método descrito anteriormente permite obtener una estimación aproximada de $A(t)$, $\phi(t)$ y, consecuentemente, $f(t)$ para una señal como la dada en (1.1), cuando incorporamos un segundo componente a la señal: $x_a(t) = A_1(t)e^{i2\pi\phi_1(t)} + A_2(t)e^{i2\pi\phi_2(t)}$ este análisis no resulta válido ya que no es posible obtener las fases instantáneas de cada componente a partir del cálculo del argumento de la función analítica $x_a(t)$.

En general, las señales no estacionarias suelen estar conformadas por varios componentes AM-FM. Bajo esta condición, es posible pensar a las señales como una *superposición* de componentes simples, o *modos*:

$$x(t) = \sum_{k=1}^K x_k(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t)), \quad A_k(t) > 0, \quad \phi'_k(t) > 0 \quad \forall t, \quad (1.7)$$

donde cada componente tendrá asociado su propia amplitud y fase instantánea. Vemos que si quisiéramos utilizar el método de la función analítica para estimar las magnitudes $A_k(t)$ y $\phi_k(t)$ nos encontramos en una situación similar a la planteada anteriormente para la suma de dos funciones analíticas, por lo cual dicho método ya no podría ser aplicado.

A modo ilustrativo, consideremos una señal no estacionaria $x(t)$ definida sobre el intervalo $[0, 1]$ de la siguiente manera:

$$\begin{aligned} x(t) = & 5 \cos(2\pi 50t) + (3 + \cos(2\pi t)) \cos(2\pi(75t + 25t^2)) + \\ & + (5 - 4t) \cos\left(2\pi\left(200t + \frac{25}{4\pi} \cos(4\pi t)\right)\right). \end{aligned} \quad (1.8)$$

De la definición anterior vemos que la señal $x(t)$ está compuesta por tres componentes. El primer componente presenta una amplitud constante ($A_1(t) = 5$) y una modulación lineal en fase ($\phi_1(t) = 50t$), dando como resultado una señal cuya frecuencia de oscilación es igual a una constante $\phi'_1(t) = 50$. Este primer componente presenta entonces las características de una señal estacionaria (sin modulación en amplitud o en frecuencia) y se conoce comúnmente como un *tono puro*. El segundo componente presenta una modulación tanto en amplitud ($A_2(t) = 3 + \cos(2\pi t)$) como en fase ($\phi_2(t) = 75t + 25t^2$). Vemos que la frecuencia de esta oscilación crece linealmente con el tiempo según la ley: $\phi'_2(t) = 75 + 50t$. Este tipo de señales no estacionarias se conocen como *chirps* y, en particular, cuando la frecuencia varía linealmente con el tiempo se lo denomina *chirp lineal*. Por último, vemos que el tercer componente presenta una modulación lineal en amplitud ($A_3(t) = 5 + 4t$) y una modulación no lineal en fase ($\phi_3(t) = 200t + 25/(4\pi) \cos(4\pi t)$). Dicha modulación en fase da como resultado una modulación en frecuencia con forma sinusoidal: $\phi'_3(t) = 200 - 25 \sin(4\pi t)$, por lo cual este componente recibe el nombre de *chirp sinusoidal*.

En la Fig. 1.2 se muestra la gráfica de la función $x(t)$ definida en (1.8). Además, se muestran tres segmentos de dicha señal en la segunda fila de la figura, donde se observa que la forma de la oscilación cambia para diferentes ventanas temporales. Esto pone de manifiesto lo dificultoso que resulta el análisis de este tipo de señales en el dominio temporal. En este dominio, no es posible determinar el número de componentes de la

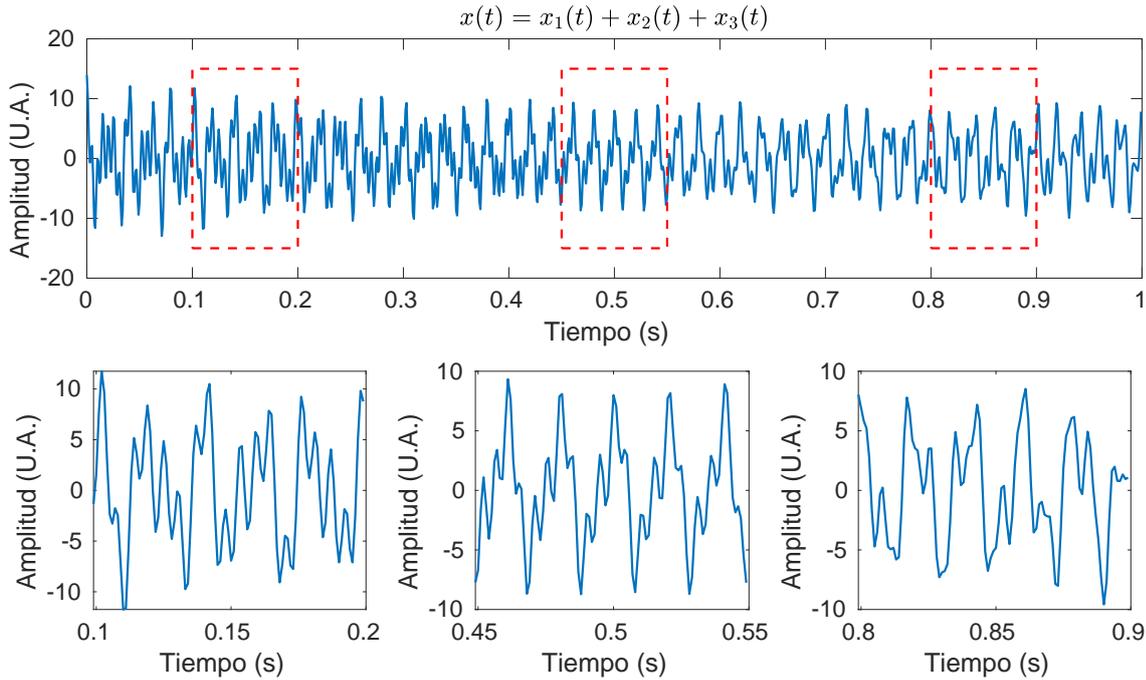


Figura 1.2: Señal no estacionaria en el dominio temporal. Primera fila: Señal no estacionaria $x(t)$ compuesta por tres componentes AM-FM, con tres segmentos de 100 milisegundos resaltados con recuadros rojos. Segunda fila: Versiones ampliadas de los segmentos de señal resaltados en rojo.

señal y menos aún a qué tipo de modulaciones están sometidas cada una de ellas. Resulta necesario, entonces, un nuevo enfoque para analizar este tipo de señales.

1.2.1. Análisis espectral

Una manera efectiva de estudiar señales como la considerada en (1.8) es mediante transformaciones que permitan representar la información de la señal en un nuevo dominio. Quizás la primera herramienta de este tipo que suele utilizarse es la transformada de Fourier (FT, de las siglas en inglés de *Fourier transform*), la cual representa la información de la señal en un dominio frecuencial. Matemáticamente, a la transformada de Fourier de una señal $x(t)$ la denotamos como $\hat{x}(f)$ y se obtiene mediante la ecuación:

$$\hat{x}(f) = \mathcal{F}\{x(u)\} = \int_{-\infty}^{\infty} x(u)e^{-i2\pi fu} du, \quad (1.9)$$

donde $i = \sqrt{-1}$ es la unidad imaginaria y f es la variable frecuencial, medida en Hercios ($\text{Hz} = \text{s}^{-1}$). Esta transformada puede interpretarse como un producto interno de funciones: $\hat{x}(f) = \langle x(t), e^{i2\pi ft} \rangle = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft} dt$, por lo que la amplitud de $\hat{x}(f)$ mide el grado de semejanza entre la señal $x(t)$ y una oscilación compleja con frecuencia angular $2\pi f$, medida en rad/s . Además, esta operación puede ser vista como una *descomposición* de la señal sobre una base de funciones ortonormales $\{e^{i2\pi ft}\}$, la cual se conoce como base de Fourier.

Al tratarse de una función de valores complejos, $\hat{x}(f)$ suele ser estudiada a través de

1.3. ANÁLISIS TIEMPO-FRECUENCIA

su módulo $|\hat{x}(f)|$ o su argumento $\angle\hat{x}(f)$, dando como resultado los llamados *espectros* de amplitud y de fase, respectivamente. Por otro lado, podemos definir la *energía* total de una señal como $E = \int_{-\infty}^{\infty} |x(t)|^2 dt$. De la definición anterior y de (1.9) se obtiene la siguiente identidad que se conoce como relación de Parseval:

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |\hat{x}(f)|^2 df, \quad (1.10)$$

la cual vincula la energía total de la señal con la integral de $|\hat{x}(f)|^2$ respecto a la variable frecuencial f . Esta relación permite interpretar al cuadrado del módulo de $\hat{x}(f)$ como una función dependiente de f , la cual se denomina *densidad espectral de potencia* (PSD, por las siglas en inglés de *power spectral density*) y se utiliza comúnmente para estudiar cómo se distribuye la energía de la señal para diferentes valores de f .

Podemos ver en el panel inferior de la Fig. 1.3 el espectro de potencia correspondiente a la señal (1.8), donde se revela un poco más de información acerca de la señal. Del mismo, podemos ver que la señal está compuesta por al menos tres componentes con rangos de frecuencias bien definidos. Vemos que el único componente que queda correctamente descrito mediante esta representación es el tono puro $5 \cos(2\pi 50t)$ dado que se manifiesta como un pico estrecho en torno a la frecuencia de oscilación del modo $f = 50$ y cuya amplitud está asociada a la amplitud constante $A_1 = 5$ de este componente. Para el caso de los chirps, su información en el dominio frecuencial se encuentra dispersa en un rango acotado de frecuencia pero no es posible determinar el comportamiento de cada chirp a partir de esta representación frecuencial. Esto pone de manifiesto la principal limitación de la transformada de Fourier para el análisis de señales con componentes no estacionarias. Se requiere, entonces, de una herramienta que sea adecuada para analizar este tipo de señales.

1.3. Análisis tiempo-frecuencia

Debido a las limitaciones mencionadas en la sección anterior, resulta necesario entonces plantear un nuevo enfoque de análisis para señales no estacionarias donde sea posible representar la información temporal y frecuencial de la señal al mismo tiempo. En este aspecto, surge como una alternativa prometedora la utilización de las llamadas representaciones tiempo-frecuencia (TFR, por las siglas en inglés de *time-frequency representation*) las cuales son representaciones bidimensionales que se definen en un plano cuyos ejes coordenados están asociados a las variables independientes tiempo y frecuencia. Estas representaciones se obtienen a partir de aplicar transformaciones sobre la señal temporal $x(t)$ o su análogo frecuencial $\hat{x}(f)$, siendo las más utilizadas la Transformada de Fourier de Tiempo Corto y la Transformada Ondita Continua.

1.3.1. Transformada de Fourier de Tiempo Corto

Un primer enfoque para generar una TFR se basa en un análisis local de Fourier, donde se afecta a la señal bajo estudio por una función $g(t)$, conocida como la *ventana de análisis*. Dicha ventana debe presentar ciertas propiedades matemáticas, siendo la más saliente que sea una función simétrica respecto a un instante de tiempo t y que su amplitud tienda rápidamente a cero al alejarse de su centro. Este análisis de Fourier por ventanas

1.3. ANÁLISIS TIEMPO-FRECUENCIA

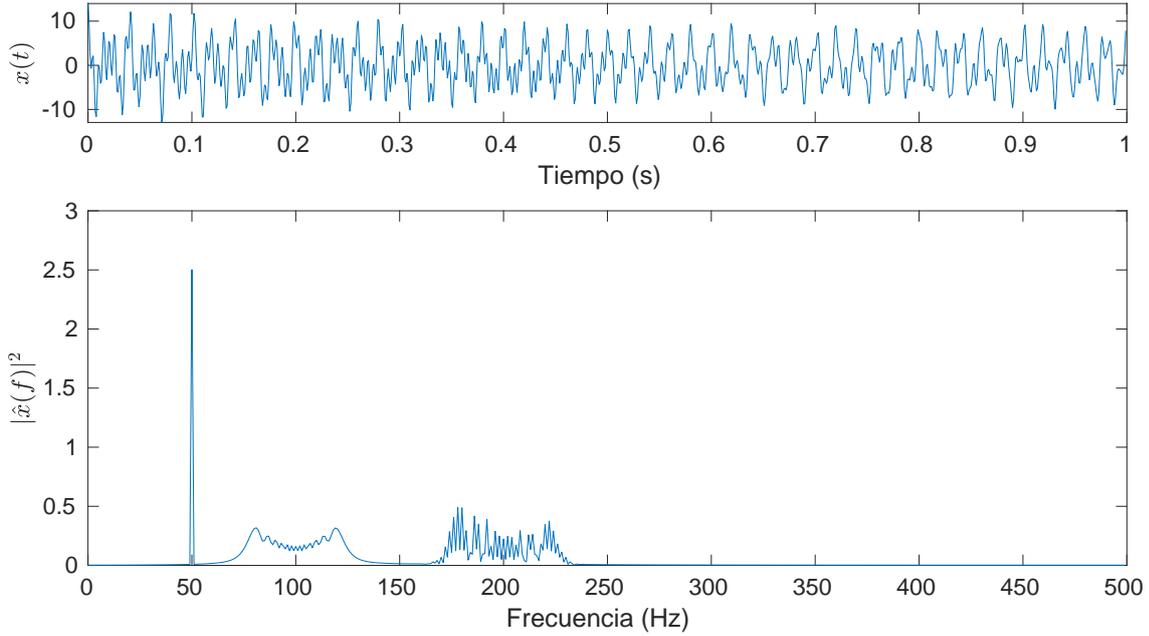


Figura 1.3: Señal no estacionaria en el dominio frecuencial. Primera fila: Señal no estacionaria $x(t)$ compuesta por tres componentes AM-FM. Segunda fila: Espectro de potencia $|\hat{x}(f)|^2$ para la señal multicomponente.

motiva la definición de la transformada de Fourier de tiempo corto (STFT, por las siglas en inglés de *short-time Fourier transform*).

Definición 1.3.1 Transformada de Fourier de Tiempo Corto (STFT)

Dada una función real $g(t)$ normalizada en $L^2(\mathbb{R})$ y par, la transformada de Fourier de tiempo corto de la señal x usando la función de ventana g se define como

$$F_x^g(t, f) = \int_{-\infty}^{\infty} x(u)g(u - t)e^{-i2\pi f(u-t)} du. \quad (1.11)$$

Además, la representación en un plano $t-f$ del módulo al cuadrado de la STFT, $|F_x^g(t, f)|^2$, se denomina el espectrograma de x .

Aquí aclaramos que la definición anterior es una definición *modificada* de la STFT tradicional, donde el núcleo de la transformación está afectado por un desfase $e^{i2\pi ft}$. Esta definición ha recibido especial atención recientemente [24, 25, 26] ya que permite interpretar la STFT como una *convolución* entre la señal $x(t)$ y el núcleo tiempo-frecuencia $g_f(t) = g(t)e^{i2\pi ft}$, es decir, $F_x^g(t, f) = (x \star g_f)(t)$. Por propiedad de la transformada de Fourier de la convolución, vemos que

$$F_x^g(t, f) = \mathcal{F}^{-1} \{ \hat{x}(u)\hat{g}(u - f) \}. \quad (1.12)$$

Esto implica que podemos pensar a la STFT como el resultado de aplicar un banco de filtros $\hat{g}(u - f)$, cada uno centrado en la frecuencia f y cuya respuesta en frecuencia está dada por la transformada de Fourier de la ventana $g(t)$. Al utilizar la misma ventana para todos los valores de f , todos los filtros del banco contarán con la misma amplitud y el

1.3. ANÁLISIS TIEMPO-FRECUENCIA

mismo ancho de banda. De esta manera, el cómputo de la STFT puede implementarse de manera eficiente en el dominio frecuencial mediante (1.12) si podemos encontrar $\hat{g}(f)$ de manera analítica. En particular, para una ventana gaussiana $g_\sigma(t) = e^{-\sigma t^2}$ se demuestra que su transformada de Fourier también es una función gaussiana de la forma

$$\hat{g}_\sigma(f) = \sqrt{\frac{\pi}{\sigma}} e^{-\pi^2 f^2 / \sigma}. \quad (1.13)$$

Esto nos permite implementar la STFT con ventana gaussiana de la siguiente manera:

$$F_x^g(t, f) = \sqrt{\frac{\pi}{\sigma}} \mathcal{F}^{-1} \left\{ \hat{x}(u) e^{-\pi^2 (u-f)^2 / \sigma} \right\}. \quad (1.14)$$

Es posible recuperar la señal original $x(t)$ mediante la inversión de la transformada de Fourier de tiempo corto [27]:

$$x(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_x^g(u, f) g(t-u) e^{i2\pi f(t-u)} du df. \quad (1.15)$$

Alternativamente, es posible recuperar la señal $x(t)$ mediante una operación de reconstrucción “vertical” [28]:

$$x(t) = \frac{2}{g(0)} \operatorname{Re} \left\{ \int_{-\infty}^{\infty} F_x^g(t, f) df \right\}, \quad (1.16)$$

donde Re implica tomar la parte real de la integral ya que su resultado será una función a valores complejos.

1.3.2. Transformada Ondita Continua

Otra manera de generar representaciones bidimensionales a partir de señales unidimensionales se basa la teoría de las onditas o *wavelets* [29]. Una ondita $\psi(t)$ es una función oscilatoria de media cero

$$\int_{\mathbb{R}} \psi(t) dt = 0, \quad (1.17)$$

y con soporte compacto: $\psi(t) = 0$ para $t \in \mathbb{R} \setminus \operatorname{supp}(\psi)$. Esta ondita nos permite definir la transformada ondita continua (CWT, por las siglas en inglés de *continuous wavelet transform*) de la siguiente manera:

Definición 1.3.2 Transformada Ondita Continua (CWT)

Dada $x(t) \in \mathbb{R}$ y una función ψ perteneciente a la clase de Schwartz, la transformada ondita continua (CWT) de x usando la ondita madre ψ se define como

$$W_x^\psi(t, s) = \int_{-\infty}^{\infty} x(u) \frac{1}{s} \psi^* \left(\frac{u-t}{s} \right) du, \quad s > 0 \quad (1.18)$$

Además, la representación en un plano $t-s$ del módulo al cuadrado de la CWT, $|W_x^\psi(t, s)|^2$, se denomina el *escalograma* de x .

1.3. ANÁLISIS TIEMPO-FRECUENCIA

Notamos aquí que nuestra definición para la CWT difiere de las definiciones tradicionales de Daubechies [29] y Mallat [30] en un factor de \sqrt{s} . De la definición anterior, se observa que la CWT se obtiene de calcular el producto interno entre la señal $x(t)$ y versiones desplazadas en el tiempo y escaladas de $\psi(t)$. Esto nos permite interpretar a la CWT como el conjunto de coeficientes que se obtienen al descomponer la señal $x(t)$ sobre un diccionario con átomos $\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi(\frac{u-t}{s})$. Otra interpretación de este operador se basa en observar que, al igual que la STFT, podemos pensar a la CWT como una convolución temporal entre la señal x y la onda desplazada y escalada $\psi_{u,s}(t)$: $W_x^\psi(t, s) = (x \star \bar{\psi}_s)(t)$, siendo $\bar{\psi}_s(t) = \frac{1}{s}\psi(\frac{-t}{s})$. De esto surge la posibilidad de calcular la CWT en el dominio frecuencial mediante la siguiente fórmula:

$$W_x^\psi(t, s) = \mathcal{F}^{-1}\{\hat{x}(f)\hat{\psi}(sf)\}. \quad (1.19)$$

Al igual que la STFT, la CWT puede interpretarse como un banco de filtros con la diferencia de que ahora los filtros no tendrán todos el mismo ancho de banda sino que éste dependerá de los valores de la variable de escala s , resultando entonces en un banco de filtros con factor de calidad Q constante (el factor de calidad Q es una medida que relaciona el valor de la frecuencia central de un filtro pasa-banda con su ancho de banda).

Cabe aclarar que la CWT genera una representación de la señal en un espacio *tiempo-escala*, a diferencia de la STFT que permite obtener una representación en un plano tiempo-frecuencia. Sin embargo, es posible establecer una relación entre la variable de escala s y la variable frecuencial f . Oscilaciones rápidas (alta frecuencia) se asocian con escalas más pequeñas, mientras que oscilaciones más lentas (baja frecuencia) se corresponden a escalas más grandes. De esta manera, es posible definir una relación entre s y f de la siguiente manera:

$$f = \frac{f_p}{s}, \quad (1.20)$$

donde f_p es la frecuencia para la cual se produce el pico máximo de $\hat{\psi}(f) = |\mathcal{F}\{\psi(t)\}|$.

La señal $x(t)$ puede recuperarse a partir de $W_x^\psi(t, s)$ mediante la inversión:

$$x(t) = \frac{2}{C_\psi} \operatorname{Re} \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_x^\psi(u, s) \psi\left(\frac{t-u}{s}\right) du \frac{ds}{s^2} \right\}, \quad (1.21)$$

siendo $C_\psi = \int_0^{\infty} |\hat{\psi}(f)|^2 \frac{df}{f} < \infty$ la denominada “condición de admisibilidad” [31]. Al igual que con la STFT, es posible aplicar una fórmula de reconstrucción vertical para recuperar $x(t)$:

$$x(t) = \frac{2}{\tilde{C}_\psi} \operatorname{Re} \left\{ \int_{-\infty}^{\infty} W_x^\psi(t, s) \frac{ds}{s} \right\}. \quad (1.22)$$

siendo ahora la condición $\tilde{C}_\psi = \int \hat{\psi}^*(f) \frac{df}{f} < \infty$.

1.3.3. Representaciones tiempo-frecuencia de señales no estacionarias

La variación temporal del contenido frecuencial de las señales modeladas según (1.7) puede estudiarse a partir del espectrograma o el escalograma, ya que en ambos casos

podemos obtener una representación en el plano tiempo-frecuencia. Para el caso de la STFT de una señal como en (1.7) obtenemos [32]:

$$F_x^g(t, f) = \int_{-\infty}^{\infty} x(u)g(u-t)e^{-i2\pi f(u-t)}du = \sum_{k=1}^K \int_{-\infty}^{\infty} x_k(u)g(u-t)e^{-i2\pi f(u-t)}du = \approx \sum_{k=1}^K x_k(t)\hat{g}(f - \phi'_k(t)), \quad (1.23)$$

siendo válida la aproximación del segundo renglón siempre y cuando $|\phi'_k(t) - \phi'_l(t)| > \Delta$ para $k \neq l$ y $t > 0$ y, además, se debe cumplir que $|\phi''_k(t)| < \epsilon$ para todo k y para algún ϵ lo suficientemente pequeño. La primera condición establece que las frecuencias instantáneas de cada una de los componentes de la señal ocupan regiones separadas en el plano tiempo-frecuencia o, como se dice coloquialmente, “no se cruzan”. La condición sobre la segunda derivada de las fases instantáneas define un conjunto de señales conocidas como *tonos levemente perturbados*. De la ecuación (1.23) se desprende que: $|F_x^g(t, f)|^2 = \sum_k |x_k(t)|^2 |\hat{g}(f - \phi'_k(t))|^2$.

Podemos aplicar la transformada de Fourier de tiempo corto sobre la señal dada por (1.8) y obtener una representación tiempo-frecuencia al calcular su espectrograma $|F_x^g(t, f)|^2$. Este espectrograma se muestra en la fila inferior de la Fig. 1.4. De analizar esta representación, podemos ver que la energía de cada uno de los componentes de $x(t)$ se concentra en torno a la región de valores de f cercanos a la frecuencia instantánea $\phi'_k(t)$ de dicho componente. Esto provoca que el espectrograma de la señal presente una serie de crestas o *ridges*, estando cada cresta asociada a un componente individual. Estas crestas permiten estimar tanto las frecuencias instantáneas como las amplitudes de los componentes mediante un procedimiento de *extracción de crestas*.

La tarea de extracción de crestas en representaciones tiempo-frecuencia ha recibido mucha atención a lo largo de los años [33, 34, 35, 36, 37]. Diferentes enfoques y algoritmos han sido propuestos para esta tarea. El objetivo de esta tarea es encontrar un conjunto de crestas $c_k(t) \in \mathcal{C}$, que definen las trayectorias en plano tiempo-frecuencia asociadas a las frecuencias instantáneas de los diferentes componentes de la señal: $c_k(t) \approx \phi'_k(t)$ para $k = 1, 2, \dots, K$.

En su formulación más general, el problema de extracción de cresta se define como la maximización de un funcional que dependerá de los valores de frecuencia que conforman las crestas y de la amplitud de la representación en dichos valores de frecuencia:

$$\{c_k^*(t)\}_{k=1}^K = \operatorname{argmax}_{c_k(t) \in \mathcal{C}} E(c_k(t), |F_x^g(t, c_k(t))|^2) \quad (1.24)$$

La forma que tome el funcional E determinará el tipo de curvas $c_k(t)$ que se podrán obtener al resolver el problema de optimización, así como también la estrategia de resolución más adecuada. En una primera instancia, resulta intuitivo seleccionar las trayectorias en el plano tiempo-frecuencia donde la amplitud del espectrograma sea máxima

$$\{c_k^*(t)\}_{k=1}^K = \operatorname{argmax}_{c_k(t) \in \mathcal{C}} \int |F_x^g(t, c_k(t))|^2 dt. \quad (1.25)$$

Una manera computacionalmente eficiente de resolver el problema (1.25) es mediante la utilización de un algoritmo *voraz*, con el cual se resuelven problemas de optimización

1.3. ANÁLISIS TIEMPO-FRECUENCIA

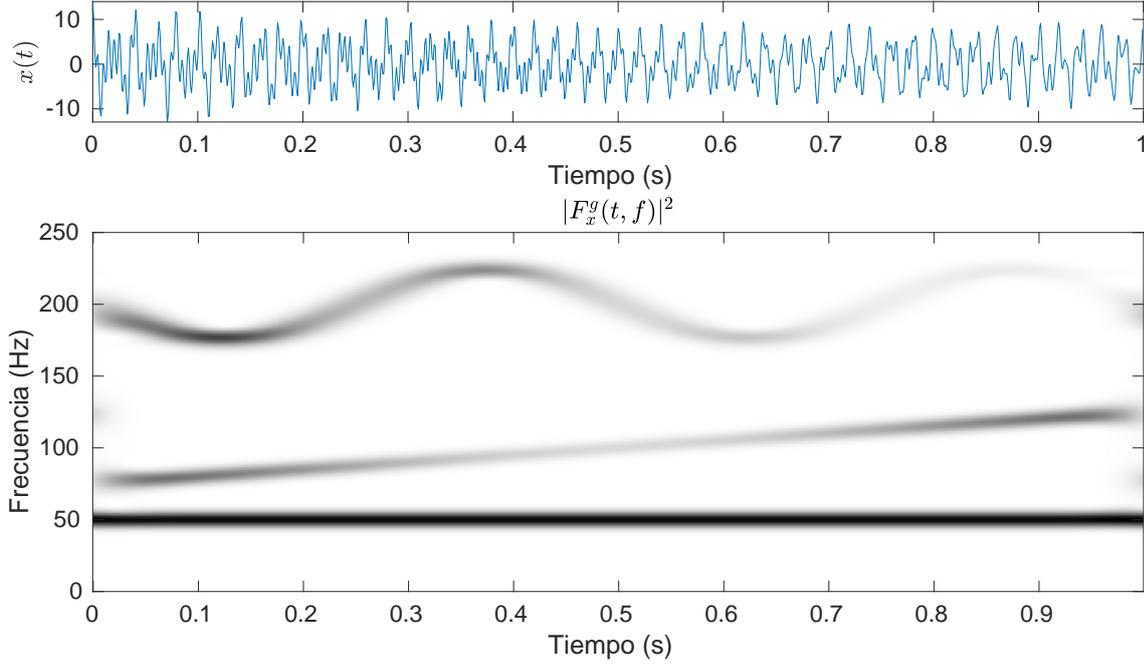


Figura 1.4: Señal no estacionaria en el dominio tiempo-frecuencia. Primera fila: Señal no estacionaria $x(t)$ compuesta por tres componentes AM-FM. Segunda fila: Espectrograma de la señal $x(t)$ donde se evidencia la presencia de *crestas* en el plano tiempo-frecuencia en torno a las curvas asociadas a la frecuencia instantánea de cada componente.

independientes para cada instante de tiempo y para cada componente de la señal. La curva óptima $c_k^*(t)$ para cada componente se obtiene mediante la conexión de los puntos óptimos encontrados para cada tiempo t . Alternativamente, es posible considerar que el óptimo para cada instante depende de una serie de instantes consecutivos antes o después del instante de interés. En este caso, y si la cantidad de puntos es finita, es posible resolver el problema (1.25) mediante un algoritmo de programación dinámica [34]. Independientemente de la estrategia de resolución, esta propuesta falla ante la presencia de ruido de alto nivel, el cual se manifiesta como “picos” aislados en el espectrograma que pueden confundir al algoritmo y generar curvas incorrectas debido a la presencia de “saltos” en la frecuencia. Esta misma situación puede producirse en el caso de tener más de un componente en la señal, ya que esto genera varias crestas en la representación y, si la amplitud de los componentes es similar entre sí, el algoritmo podría saltar de un componente a otro en diferentes instantes de tiempo.

Una manera de superar esta dificultad es incorporar términos de regularización al funcional E que promuevan la suavidad de la curva extraída:

$$\{c_k^*(t)\}_{k=1}^K = \operatorname{argmax}_{c_k(t) \in \mathcal{C}} \left(\int |F_x^g(t, c_k(t))|^2 dt - \lambda \int |c_k'(t)|^2 dt - \mu \int |c_k''(t)|^2 dt \right), \quad (1.26)$$

donde los coeficientes λ y μ controlan qué tan suave será la curva extraída. Por otro lado, han surgido estrategias alternativas que restringen el rango de búsqueda en frecuencia para instantes de tiempo consecutivos, evitando así los saltos [36]. A esta propuesta se le

suele incorporar un procedimiento de múltiples pasadas con inicialización aleatoria [35], en la cual se extraen varias crestas para un mismo componente. Luego se aplica un criterio de selección para obtener una cresta adecuada, siendo el criterio más habitual el de seleccionar la cresta de mayor energía acumulada en el tiempo.

Independientemente de cómo se obtengan, las crestas $c_k^*(t)$ permiten tener una estimación de la frecuencia instantánea de los componentes de la señal. Es común que también necesitemos una estimación de la modulación en amplitud y de modulación en fase para poder reconstruir el modo aislado como $x_k(t) = A_k(t) \cos(2\pi\phi_k(t))$. Esto puede lograrse mediante la siguiente operación

$$\hat{y}_k(t) = \frac{2}{g(0)} \int_{|f-c_k^*(t)| < \Delta} F(t, f) df. \quad (1.27)$$

Esto no es más que una reconstrucción vertical de la STFT, restringiendo el dominio de reconstrucción a una *franja* de frecuencias de ancho 2Δ en torno a $c_k^*(t)$. La función $\hat{y}_k(t)$ es una versión “complejizada” del modo real $x_k(t)$, a partir del cual podemos estimar la modulación en amplitud y fase de una manera análoga al método de las funciones analíticas: $A_k(t) \approx \hat{A}_k(t) = |\hat{y}_k(t)|$ y $\phi_k(t) \approx \hat{\phi}_k(t) = \angle \hat{y}_k(t)$. Claramente, el ancho de la franja de reconstrucción Δ debe ser elegido de tal manera que no se solapen las regiones de reconstrucción de los diferentes componentes. Una forma de elegir este parámetro para el caso de STFT con una ventana gaussiana es observando que el desvío estándar de la ventana en frecuencia es: $\text{std}[\hat{g}(f)] = \text{std}\left[\sqrt{\frac{\pi}{\sigma}} e^{-\pi^2 f^2 / \sigma}\right] = \sqrt{\sigma} / (\sqrt{2\pi})$. Sabiendo que a una distancia de 3 veces el desvío estándar desde el pico de la gaussiana su amplitud disminuyó más del 98 % de su máximo, definimos $\Delta = 3\sqrt{\sigma} / (\sqrt{2\pi})$.

1.3.4. Posprocesamiento de representaciones tiempo-frecuencia

En los últimos años, se han comenzado a estudiar operaciones y transformaciones sobre las representaciones tiempo-frecuencia clásicas (como STFT y CWT) con el objetivo de obtener nuevas TFRs con mejores propiedades según la tarea que se busca realizar. Un ejemplo de esto son los métodos de concentración de representaciones tiempo-frecuencia, cuyo objetivo es conseguir TFRs con una mayor resolución que aquella que es posible según el principio de incertidumbre de Heisenberg-Gabor. Este tipo de métodos comenzaron a estudiar en la década de 1970 a partir de los desarrollos de Kodera [38, 39], quien propuso una TFR obtenida a partir del cálculo de los centros de gravedad de la distribución de energía tiempo-frecuencia de la señal en las regiones del plano donde la energía sea mayor a cero. Esta idea fue posteriormente retomada por Auger y Flandrin, quienes propusieron el método de *reasignación* [40] para mejorar la legibilidad de las TFR mediante una propuesta análoga a la de Kodera. El resultado de esta técnica son representaciones cuya información se encuentra más concentrada en torno a las crestas mediante una reubicación de los coeficientes de la TFR tanto en tiempo como en frecuencia. Aunque estas representaciones presentan ventajas a la hora de estimar la frecuencia instantánea de los componentes de la señal, no admiten una fórmula de reconstrucción para recuperar la señal original ni los componentes individuales.

Notando esta limitante, Daubechies y Maes propusieron la técnica de *synchrosqueezing* [41], la cual realiza una reasignación de los coeficientes solamente en la dirección frecuencial manteniendo de esa manera la “sincronía” entre la TFR original y la concentrada. Esta TFR obtenida mediante *synchrosqueezing* admite una fórmula de recons-

trucción vertical y además también permite recuperar los componentes individuales de la señal mediante una reconstrucción por franjas en torno a las crestas [42].

1.4. Modelado fenomenológico de señales no estacionarias

Debido a la alta complejidad de los sistemas y procesos que dan origen a las señales, es común incorporar información de los fenómenos físicos y del proceso de registro de las señales para mejorar el desempeño de las estrategias de análisis y procesamiento [43]. En este contexto, los modelos *fenomenológicos* tienen como objetivo modelar el valor observado de la señal, sin establecer consideraciones acerca de cómo interactúan las variables del sistema físico que genera dicha observaciones. Esto les confiere una gran versatilidad a la hora de lograr describir fenómenos muy diversos de la vida real, y permiten de esta manera desarrollar herramientas de análisis y procesamiento generales que puedan desempeñarse satisfactoriamente en diferentes ámbitos. La idea de representar una señal real mediante una superposición de componentes AM-FM es un ejemplo de un modelo fenomenológico para señales no estacionarias y las herramientas de análisis que hemos discutido hasta ahora son fácilmente aplicables a este tipo de señales. En esta sección, discutiremos más en detalle el modelo propuesto en (1.7) para posteriormente desarrollar un nuevo modelo más versátil que incluye la información asociada a la morfología de la onda oscilatoria en su formulación.

1.4.1. Modelo armónico adaptativo

Para comenzar, repetimos el modelo (1.7) y definimos una serie de condiciones adicionales:

$$x(t) = \sum_{k=1}^K A_k(t) \cos(2\pi\phi_k(t)), \text{ con } A_k(t), \phi'_k(t) > 0 \forall t. \quad (1.28)$$

Dada la modulación en fase del k -ésimo componente podemos definir la *frecuencia instantánea* $f_k(t)$ de ese componente como la primera derivada de la fase, $f_k(t) = \phi'_k(t)$. Además, definimos las condiciones de variabilidad lenta para $A_k(t)$ y $\phi'_k(t)$:

$$|A'_k(t)|/A_k(t) < \varepsilon_1, \quad |\phi''_k(t)| < \varepsilon_2, \quad \forall t.$$

Recordamos de la Sec. 1.3.3 que la condición sobre el módulo de la derivada segunda de $\phi_k(t)$ define la clase de tonos levemente perturbados. Cuando $K > 1$ también imponemos una condición de separabilidad: $|\phi'_k(t) - \phi'_{k-1}(t)| > d$, para todo t . Nótese que aquí también estamos suponiendo que los componentes se ordenan por sus frecuencias instantáneas de forma creciente.

Este modelo se conoce como *modelo armónico adaptativo* (AHM, por las siglas en inglés de *adaptive harmonic model*) ya que cada uno de sus componentes se define como una función armónica (cosenoidal) modulada en amplitud y en frecuencia. Las condiciones de variación lenta para $A_k(t)$ y $\phi'_k(t)$ confieren a la señal de una estacionariedad *local*, permitiendo su estudio en ventanas de corta duración como si fuese estacionaria. En la Fig. 1.5 se muestra la señal multicomponente (1.8) y cada una de las señales AM-FM que la componen por separado.

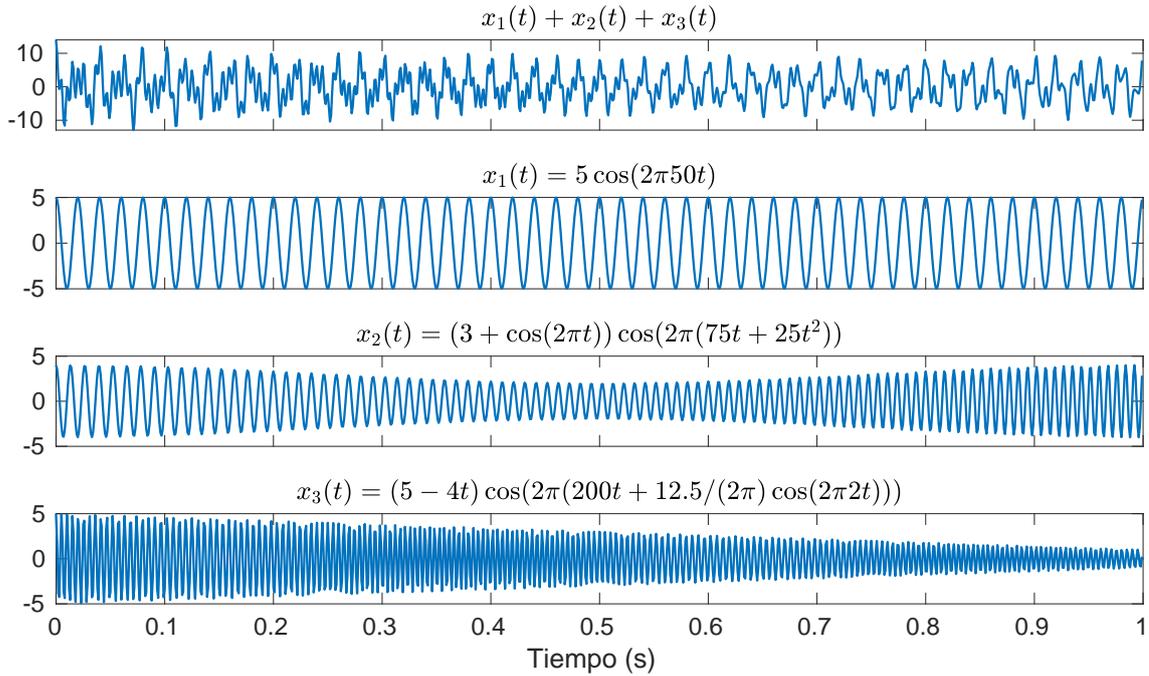


Figura 1.5: Modelo armónico adaptativo. Primera fila: Señal multicomponente que sigue el modelo armónico adaptativo. Segunda fila: Primer componente $x_1(t)$ (tono puro). Tercera fila: Segundo componente $x_2(t)$ (chirp lineal). Cuarta fila: tercer componente $x_3(t)$ (chirp sinusoidal).

Claramente, este modelo permite definir una familia muy amplia de señales y se ha aplicado en diversos ámbitos, como el estudio de maquinas rotativas [44, 45, 46], análisis de señales biomédicas [47, 48, 49], procesamiento de señales de radar [50], entre otros. Sin embargo, una limitante de este modelo es que asocia cada componente con una oscilación simple tipo coseno. Es sabido que las señales reales presentan patrones oscilatorios más complejos. Además, la señal puede contener dos o más componentes, cada uno con su propio patrón oscilatorio. Un ejemplo claro de esto son las señales de electrocardiografía, las cuales suelen verse afectadas por un componente de baja frecuencia asociada al proceso respiratorio. Sería esperable poder obtener, a partir de dicha señal, dos componentes bien definidos, uno asociado a la actividad cardíaca y otro vinculado con el fenómeno respiratorio. Una primera alternativa para superar esta limitante es combinar varios modos de la forma $A_k(t) \cos(2\pi\phi_k(t))$ de tal manera de que el “componente resultante” aproxime adecuadamente la oscilación del componente buscado. Sin embargo, no es claro cómo se deben elegir los modos correspondientes ni tampoco cuántos modos serán necesarios para reconstruir cada componente. Adicionalmente, este modelo requiere que la condición de separabilidad se cumpla para todos los pares de modos $\{k, k + 1\}$. Como veremos más adelante, es común que esta condición no se cumpla, particularmente para componentes en el rango de altas frecuencias.

1.4.2. Modelo no armónico adaptativo

Con el objetivo de superar las limitantes del AHM, HT Wu propuso el siguiente modelo alternativo [51]:

$$x(t) = \sum_{k=1}^K A_k(t) s_k(2\pi\phi_k(t)). \quad (1.29)$$

Aquí, para cada componente, la función coseno (*función armónica*) se reemplaza por una función $s_k(\cdot)$, la cual es 2π -periódica con un patrón oscilatorio más general. Debido a que los componentes ya no corresponden a funciones armónicas moduladas, este modelo recibe el nombre de *modelo no armónico adaptativo* (ANHM, por las siglas en inglés de *adaptive non-harmonic model*). Las funciones $s_k(t)$ reciben el nombre de *funciones de forma de onda* (WSF, por las siglas en inglés de *wave-shape function*) y, al tratarse de función periódicas y de potencia finita, pueden ser descritas mediante su descomposición en serie de Fourier:

$$x(t) = \sum_{k=1}^K \sum_{\ell=-\infty}^{\infty} A_k(t) \hat{s}_k(\ell) e^{i2\pi\ell\phi_k(t)}, \quad (1.30)$$

donde $\hat{s}_k(\ell)$ son los coeficientes de expansión de la serie de Fourier, en su forma exponencial compleja, para la WSF $s_k(t)$. En la Fig. 1.6 vemos casos de señales biomédicas que se ajustan (al menos parcialmente) al ANHM, particularmente al caso *monocomponente* ($K = 1$):

$$x(t) = A(t) s(2\pi\phi(t)). \quad (1.31)$$

En este caso, la expansión de s mediante su serie de Fourier da como resultado:

$$\begin{aligned} x(t) &= A(t) \sum_{\ell=-\infty}^{\infty} \hat{s}(\ell) e^{2\pi i \ell \phi(t)} \\ &= A(t) \sum_{\ell=0}^{\infty} \alpha_\ell \cos(2\pi\ell\phi(t)) + \beta_\ell \sin(2\pi\ell\phi(t)), \end{aligned} \quad (1.32)$$

donde obtenemos $\alpha_\ell = 2 \operatorname{Re}\{\hat{s}(\ell)\}$ y $\beta_\ell = -2 \operatorname{Im}\{\hat{s}(\ell)\}$, al tomar la parte real y la parte imaginaria de $\hat{s}(\ell)$, respectivamente. Sabemos que la suma de una función seno y coseno con el mismo argumento se puede reescribir como un coseno desfasado, por lo que el modelo se convierte en:

$$x(t) = A(t) \sum_{\ell=0}^{\infty} a_\ell \cos(2\pi\ell\phi(t) + \varphi_\ell), \quad (1.33)$$

con $\varphi_\ell = \tan^{-1}(\beta_\ell/\alpha_\ell)$ y $a_\ell = \sqrt{\alpha_\ell^2 + \beta_\ell^2}$. Observamos aquí que este modelo termina siendo un caso particular del modelo presentado en (1.28) donde la fase instantánea del ℓ -ésimo componente es ℓ veces la fase del primer componente (también conocida como *componente fundamental*) más un desfase constante $\varphi_\ell/(2\pi\ell)$. En este contexto, denotamos cada componente como un *armónico* siendo el armónico fundamental $a_1 \cos(2\pi\phi(t) + \varphi_1)$.

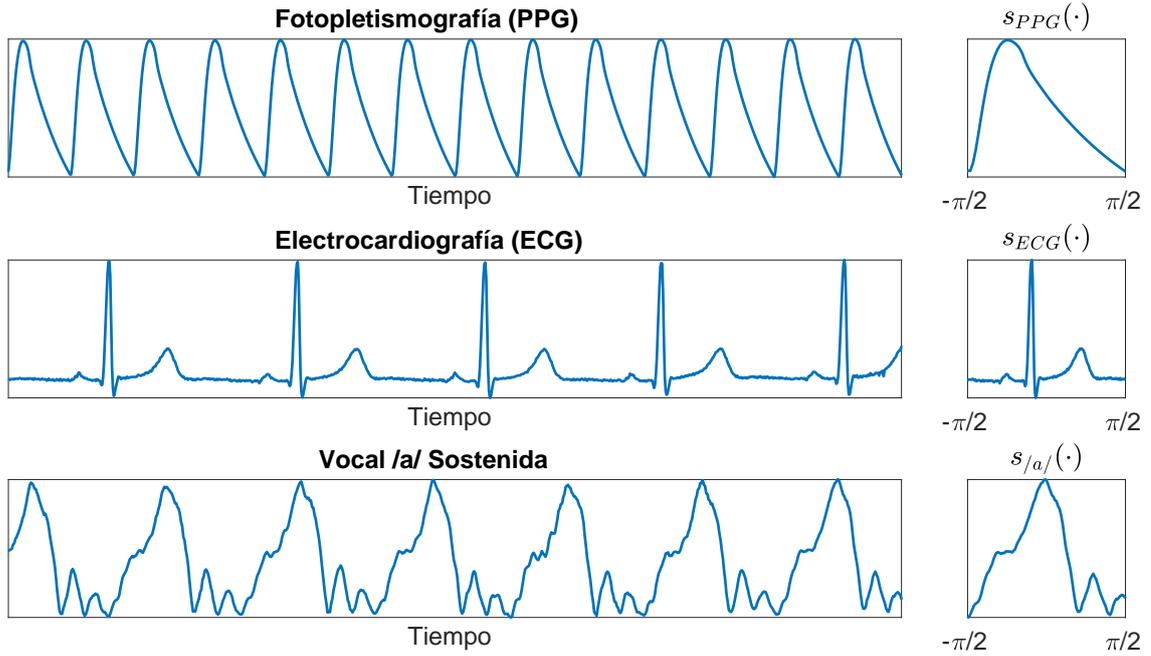


Figura 1.6: Modelo no armónico adaptativo. Señales biomédicas que se ajustan al caso monocomponente del ANHM, junto con la forma de onda característica de cada señal. Primera fila: Señal de fotopletismografía (PPG) medida a través de un oxímetro de pulso. Segunda fila: Señal de ECG. Tercera fila: Señal de voz de una emisión de una vocal /a/ sostenida.

1.4.3. Funciones de forma de onda

HT Wu estableció una serie de condiciones para el modelo en (1.32), principalmente relacionadas con la estructura de la serie de Fourier de $s(t)$:

Definición 1.4.1 *Clase de señales de forma de onda [51]*

Dados $\delta \geq 0$, $r_0 \in \mathbb{N}$ y $\theta \geq 0$, la clase de funciones de forma de onda $\mathcal{S}^{\delta, r_0, \theta}$ es el conjunto de funciones periódicas $s(t)$ con media cero, es decir $\hat{s}(0) = 0$, y norma-2 unitaria para las cuales se cumplen las siguientes condiciones:

$$(C1) \quad \forall \ell \in \mathbb{N} - \{1\}, \quad |\hat{s}(\ell)| \leq \delta |\hat{s}(1)|$$

$$(C2) \quad \sum_{\ell > r_0} |\ell \hat{s}(\ell)| \leq \theta$$

La condición C1 implica que el coeficiente asociado a la frecuencia fundamental no puede ser cero. Además, si $\delta < 1$, entonces el primer armónico es *dominante* y el componente fundamental tendrá la cresta más alta en el plano TF. Podemos entonces estimar la frecuencia instantánea detectando la cresta más energética en el espectrograma. Aunque esta condición puede resultar conveniente, algunas señales biomédicas habituales, como el ECG o la voz, no necesariamente la cumplen.

La condición C2 nos dice que la sucesión de coeficientes de Fourier $\{\hat{s}(\ell)\}$ tiende lo suficientemente rápido a cero, lo cual permite aproximar la forma de onda (que, en

general, puede tener un número infinito de armónicos) por una versión de ancho de banda limitado con sólo r_0 armónicos, con precisión θ .

En el caso más general, cuando la señal este compuesta por varias WSFs, cada una de ellas tendrá su propia representación de ancho de banda limitado resultando en la siguiente representación para la señal completa:

$$x(t) \approx \sum_{k=1}^K A_k(t) \sum_{\ell=1}^{r_k} (\alpha_{k,\ell} \cos(2\pi\ell\phi_k(t)) + \beta_{k,\ell} \sin(2\pi\ell\phi_k(t))), \quad (1.34)$$

donde el límite superior para la sumatoria más interna, r_k , determina el número de componentes armónicos considerados para representar la WSF de la k -ésima componente de la señal. Además, se inician las sumatorias a partir de $\ell = 1$ ya que consideramos modos de media cero y sin tendencia local.

Cabe destacar que la idea de considerar una señal no estacionaria como una superposición de modos AM-FM es el fundamento del método de descomposición modal empírica (EMD, por las siglas en inglés de *empirical mode decomposition*) [52]. Mediante un procedimiento iterativo y guiado por los datos, EMD permite descomponer una señal no estacionaria en un conjunto de funciones oscilatorias AM-FM conocidas como *funciones modales intrínsecas* (IMF, por las siglas en inglés de *intrinsic mode function*). Basándose en la idea de las IMF, HT Wu define el siguiente conjunto de funciones “tipo” IMF (IMT, por las siglas en inglés de *intrinsic-mode type function*):

Definición 1.4.2 Clase de funciones IMT [51]

Dados $\varepsilon, \delta, r_0 > 0$ y $\epsilon \ll 0$, la clase de funciones $\mathcal{C}_\varepsilon^{\delta, r_0}$ es el conjunto de funciones $f(t) : \mathbb{R} \rightarrow \mathbb{R}$, $f \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$ con la siguiente forma:

$$f(t) = A(t)s(2\pi\phi(t)),$$

donde $s \in \mathcal{S}^{\delta, r_0, \varepsilon}$ y tal que $A(t)$ y $\phi(t)$ satisfacen las siguientes condiciones:

$$\begin{aligned} A(t) &\in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R}), \quad \phi(t) \in C^2(\mathbb{R}), \\ \inf_{t \in \mathbb{R}} A(t) &> 0, \quad \inf_{t \in \mathbb{R}} \phi'(t) > 0, \quad \sup_{t \in \mathbb{R}} \phi'(t) < \infty, \\ |A'(t)| &\leq \varepsilon, \quad |\phi''(t)| \leq \varepsilon \text{ para todo } t \in \mathbb{R}, \\ M'' &= \sup_{t \in \mathbb{R}} |\phi''(t)| < \infty. \end{aligned}$$

Según esta definición, el ANHM dado en (1.29) puede pensarse como una *superposición* de IMTs, donde se agrega la condición de separabilidad: $|\phi'_k(t) - \phi'_{k+1}(t)| > \Delta$ sobre las frecuencias instantáneas de los componentes fundamentales para cada IMT.

1.4.4. Discretización y caracterización de formas de onda fijas

Aunque la teoría de las funciones de forma de onda se desarrolló originalmente para tiempo continuo, es posible aplicarla a una versión discretizada de $x(t)$, y es en este contexto donde describiremos los algoritmos desarrollados en esta tesis. La discretización de la señal se realiza de la siguiente manera: dada una señal de tiempo continuo $x(t)$ y una duración total de la señal T , podemos definir una señal discreta $\mathbf{x}(n)$ de longitud N

1.5. DESAFÍOS ACTUALES Y PROPUESTAS

mediante el muestreo de $x(t)$: $\mathbf{x}(n) = x(n\Delta t)$, para $0 \leq n \leq N-1$, donde $\Delta t = T/N = 1/f_s$ es el intervalo de muestreo y f_s se conoce como la tasa o frecuencia de muestreo. Si la señal discretizada se corresponde con el modelo (1.29), entonces las amplitudes y fases instantáneas también serán discretas: $\mathbf{A}_k(n) = A_k(n\Delta t)$, $\phi_k(n) = \phi_k(n\Delta t)$. Sin pérdida de generalidad, supondremos también que todas las señales estudiadas $\mathbf{x}(n)$ son de media cero.

Dada la señal discreta, debemos primeramente estimar las amplitudes y fases instantáneas $\mathbf{A}_k(n)$ y $\phi_k(n)$, respectivamente. Para esto, podemos aplicar los algoritmos de extracción de crestas y reconstrucción de modos explicados en la Sec. 1.3.3. Aquí hay que tener en cuenta que solamente será necesario obtener la estimación de la modulación en amplitud y fase del componente fundamental asociado a cada WSF. Una vez hecho esto, tenemos que determinar las funciones de forma de onda. Para ello, resolvemos el siguiente problema de regresión por mínimos cuadrados para obtener los coeficientes de expansión de Fourier de las diferentes WSFs que componen a la señal:

$$\hat{\boldsymbol{\gamma}} = \arg \min_{\boldsymbol{\gamma} \in \mathbb{R}^H} \|\mathbf{x}(n) - \mathbf{C}\boldsymbol{\gamma}\|_2^2, \quad (1.35)$$

donde $H = 2 \sum_{k=1}^K r_k$ y \mathbf{C} es un diccionario *ad hoc* de la forma:

$$\mathbf{C} = [\mathbf{C}_1 \ \dots \ \mathbf{C}_K \ \mathbf{D}_1 \ \dots \ \mathbf{D}_K], \quad (1.36)$$

donde las submatrices $\mathbf{C}_k \in \mathbb{R}^{N \times r_k}$ y $\mathbf{D}_k \in \mathbb{R}^{N \times r_k}$ están conformadas por los vectores columna: $\mathbf{c}_{k,\ell} = \mathbf{A}_k(n) \cos(2\pi\ell\phi_k(n))$ y $\mathbf{d}_{k,\ell} = \mathbf{A}_k(n) \sin(2\pi\ell\phi_k(n))$, respectivamente, para $\ell = 1, \dots, r_k$ y $k = 1, \dots, K$. El objetivo de este problema de mínimos cuadrados es encontrar un vector de coeficientes de la forma

$$\boldsymbol{\gamma} = [\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_K, \boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_r]^T, \quad (1.37)$$

siendo $\boldsymbol{\alpha}_k = [\alpha_{k,1}, \dots, \alpha_{k,r_k}]^T \in \mathbb{R}^{r_k \times 1}$ y $\boldsymbol{\beta}_k = [\beta_{k,1}, \dots, \beta_{k,r_k}]^T \in \mathbb{R}^{r_k \times 1}$ los vectores de coeficientes de Fourier asociados a la k -ésima WSF. Este problema de mínimos cuadrados tiene una solución dada por las ecuaciones normales:

$$\hat{\boldsymbol{\gamma}} = (\mathbf{C}^T \mathbf{C})^{-1} \mathbf{C}^T \mathbf{x}. \quad (1.38)$$

Finalmente, podemos reconstruir la señal con la fórmula de síntesis

$$\hat{\mathbf{x}}(n) = \mathbf{C}\hat{\boldsymbol{\gamma}}. \quad (1.39)$$

o bien, reconstruir el modo asociado a una WSF particular mediante la formula

$$\hat{\mathbf{x}}_k = [\mathbf{C}_k \ \mathbf{D}_k] \hat{\boldsymbol{\gamma}}_k, \quad (1.40)$$

con $\hat{\boldsymbol{\gamma}}_k = [\hat{\alpha}_{k,1}, \dots, \hat{\alpha}_{k,r_k}, \hat{\beta}_{k,1}, \dots, \hat{\beta}_{k,r_k}]^T$. Nótese que necesitamos fijar los valores de los números de armónicos r_k para cada WSF de antemano.

1.5. Desafíos actuales y propuestas

El modelo no armónico adaptativo ha sido utilizado ampliamente en el ámbito del procesamiento de señales, particularmente en aplicaciones biomédicas [53, 54, 55, 56, 57, 58]. En estos trabajos, la estimación de las funciones de forma de onda se realizó fijando el número de armónicos r previamente a aplicar y resolver el modelo para la reconstrucción de las WSFs. Esta consideración es bastante común en la práctica, donde la elección del número de armónicos se realiza basándose en el conocimiento previo del usuario sobre las señales bajo análisis. Aunque este procedimiento da buenos resultados cuando las características de las señales de diferentes sujetos son similares entre sí, resulta pertinente proponer un criterio objetivo para determinar, de manera automática, el valor óptimo de este parámetro a partir de la señal bajo estudio, sin necesidad de definirlo en base al conocimiento previo del proceso.

Por otro lado, un aspecto clave de este modelo que no hemos discutido es que supone una forma de onda *fija*, es decir, que el patrón oscilatorio de la señal se preserva a lo largo del tiempo y las variaciones entre ciclos son explicadas totalmente por la modulación en amplitud y la modulación en fase de la señal. Es claro que esta condición no se cumplirá para una gran variedad de señales reales, particularmente en casos de señales biomédicas asociadas a procesos patológicos. Aunque se han propuesto alternativas al ANHM que consideran la posibilidad de una forma de onda variable en el tiempo [59, 60], veremos en los capítulos siguientes que dichos modelos presentan limitaciones y es posible obtener un mejor desempeño en tareas de procesamiento como limpieza de ruido, segmentación y descomposición de señales multicomponente a partir de la formulación de un nuevo procedimiento de estimación de la forma de onda variante en el tiempo.

Finalmente, aunque el ANHM y sus variantes se han aplicado a diversas señales del ámbito biomédico, existen todavía muchas áreas de aplicación sin explorar. En particular, identificamos dos tareas donde el ANHM podría ser una alternativa interesante a las herramientas tradicionales, a saber, la imputación de datos perdidos en señales biomédicas y el estudio de la voz en base a la variabilidad temporal de su forma de onda.

Los aportes originales de esta tesis abordan estos desafíos y definen los objetivos que se detallan en la siguiente sección.

1.6. Objetivos de la tesis

A partir de lo expuesto anteriormente, se definen los siguientes objetivos para esta tesis:

1.6.1. Objetivos generales

1. Diseñar, desarrollar y validar algoritmos de procesamiento de señales no estacionarias basados en la aplicación del modelo no armónico adaptativo.
2. Aplicar las herramientas algorítmicas desarrolladas a diferentes tareas del ámbito biomédico.

El cumplimiento de estos objetivos generales se logrará a través de los objetivos específicos detallados a continuación.

1.6.2. Objetivos específicos

1. Proponer un procedimiento de estimación adaptativa del número de armónicos necesarios para reconstruir la WSF de una señal no estacionaria, mediante un método guiado por los datos.
2. Desarrollar un método adaptativo para caracterizar la forma de onda variante en el tiempo de señales no estacionarias, tanto para el caso monocomponente como para el caso multicomponente.
3. Diseñar e implementar un algoritmo de imputación de datos perdidos para el estudio de señales unidimensionales no estacionarias que tenga en cuenta las variaciones debido a las modulaciones en amplitud y frecuencia, y la variabilidad de la forma de onda.
4. Validar los algoritmos propuestos mediante su aplicación en tareas de procesamiento de señales sintéticas y reales.
5. Aplicar el modelo no armónico adaptativo con forma de onda variante en el tiempo al estudio de la voz humana, haciendo foco en situaciones donde la variabilidad en la morfología de la onda permita inducir información relevante acerca del estado del paciente.

1.7. Estructura de la tesis

El resto de este documento se organiza en varios capítulos donde se presentan las propuestas originales de esta tesis, las cuales permitieron alcanzar los objetivos planteados. En el Cap. 2 se describe un enfoque para la estimación adaptativa del número óptimo de armónicos para reconstruir la WSF de una señal no estacionaria, utilizando criterios de selección de modelos de regresión trigonométrica. El Cap. 3 introduce una propuesta para estimar la WSF variante en el tiempo de señales no estacionarias, basada en una variante novedosa del ANHM, que emplea un algoritmo adaptativo de ajuste de curvas.

En el Cap. 4, se detalla un algoritmo de imputación de datos faltantes en señales biomédicas no estacionarias, fundamentado en la descomposición de la señal en base a su representación usando el ANHM. En el Cap. 5 se presenta la aplicación del algoritmo propuesto en el Cap. 3 para extraer medidas asociadas a la variabilidad de la forma de onda en señales de voz, con el objetivo de utilizarlas como características para entrenar y validar clasificadores en las tareas de tipificación de voces y detección de la enfermedad de Parkinson.

Finalmente, en el Cap. 6 se presentan las conclusiones de esta tesis y se proponen posibles líneas de investigación futuras derivadas de los desarrollos expuestos en los capítulos anteriores.

Capítulo 2

Estimación adaptativa del número de armónicos para función de forma de onda

Como discutimos en el capítulo anterior, el modelo no armónico adaptativo resulta una herramienta muy potente para el análisis y procesamiento de señales no estacionarias cuya morfología puede ser caracterizada mediante una función de forma de onda [61]. La WSF de la señal puede ser descrita mediante los coeficientes de su expansión en serie de Fourier y, bajo los criterios establecidos en la Sec. 1.4.3, es posible utilizar representaciones de ancho de banda limitado para describir a estas señales y sus funciones de forma de onda. Surge, en este contexto, la necesidad de elegir el número de armónicos a considerar para la representación de la WSF. Este parámetro debe ser determinado de manera previa a la obtención del vector de coeficientes óptimos de Fourier $\hat{\gamma}$. Es lógico pensar que la cantidad de armónicos necesarios para reconstruir una función de forma de onda dependerá de la señal bajo estudio. Por lo tanto, resulta atractivo disponer de una estrategia adaptativa o *guiada por los datos* para determinar el número de armónicos más adecuado a partir de la información disponible de la señal.

Un problema similar a éste es el *conteo* de señales, donde se busca determinar de manera guiada por los datos la cantidad de señales contenidas en las mediciones realizadas. Para este fin, diferentes propuestas han sido exploradas, como por ejemplo la utilización de la entropía de Rényi [62] o el conteo basado en el umbralado de representaciones tiempo-frecuencia binarizadas [63]. Aunque estos métodos presentan resultados adecuados en el análisis de señales multicomponentes con modulación no lineal, ningún método aprovecha la estructura armónica de las funciones de forma de onda.

En este capítulo, describiremos en detalle el problema de estimación del número óptimo de armónicos para representar la función de forma de onda de una señal no estacionaria. Este problema será planteado como una selección de modelos. La resolución de este problema sirve como una etapa previa al planteo de un algoritmo de estimación de la forma de onda de señales no estacionarias que cumpla con la condición de ser guiado por los datos, es decir, cuya aplicación dependa lo menos posible de parámetros establecidos a priori por el operador. Esta condición resulta deseable para conferirle al modelo la capacidad de ser altamente adaptable al problema específico que se busca resolver.

2.1. Selección de modelos

En el contexto del modelado matemático de sistemas, la selección de modelos implica la elección de los parámetros del modelo que resulte óptimo según algún criterio para describir la información del sistema. Este criterio es, en términos concretos, una función que depende del modelo y de sus parámetros. De esta forma, podemos definir una familia de modelos \mathcal{M}_θ , parametrizados mediante un conjunto de parámetros $\theta \in \Theta$, donde Θ es el espacio de los parámetros. El conjunto óptimo de parámetros $\hat{\theta}$ se obtiene mediante la minimización del criterio de selección $\Phi(\theta)$:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \Phi(\theta) \quad (2.1)$$

Generalmente, el criterio elegido $\Phi(\theta)$ debe ser capaz de establecer una relación de compromiso entre la bondad en el ajuste y la complejidad del modelo. Estudios publicados en la bibliografía demuestran que modelos con mayor número de parámetros (complejidad) son capaces de ajustarse mejor a los datos observados [64, 65, 66]. Sin embargo, esto puede llevar a un *sobreajuste* del modelo debido a que los datos suelen contener interferencias e información indeseada y, al aumentar la complejidad del modelo, no sólo estaremos ajustándonos a la información de interés sino también a las fuentes de ruido.

Bajo esta premisa, Akaike [67] propuso un criterio de selección basándose en la teoría de la información aplicada a modelos estadísticos. Este criterio se conoce como el criterio de la información de Akaike (AIC, por las siglas en inglés de *Akaike's information criterion*) y se ha utilizado en diferentes disciplinas. En particular, el criterio de Akaike puede expresarse como:

$$\text{AIC}(k) = -2 \ln \hat{L} + 2k, \quad (2.2)$$

donde k es el número de parámetros del modelo y \hat{L} se conoce como el valor máximo de la función de verosimilitud para el modelo. La función de verosimilitud $L(\theta|x)$ mide la probabilidad de que el modelo con parámetros θ sea el correcto, dado los datos observados. Comúnmente, modelos de mayor complejidad están asociados a mayores valores de verosimilitud, lo cual hace que disminuya el valor del término $-2 \ln \hat{L}$ para estos modelos más complejos. Al agregar un término de penalización ($2k$), el AIC permite seleccionar un modelo de menor complejidad manteniendo una correspondencia adecuada con los datos.

En línea con el AIC, se han propuesto diversos criterios de selección de modelos como el criterio de información bayesiano (BIC, del inglés *Bayesian information criterion*) [68] o la mínima longitud de descripción (MDL, del inglés *minimum description length*) [69]. Estos siguen la misma estructura del AIC pero modifican sus términos de penalización para obtener criterios con ciertas propiedades deseables, como la consistencia estadística y menor propensión a sobrestimar la complejidad del modelo.

Aunque los criterios descritos hasta ahora pueden ser aplicados en un amplio rango de problemas, veremos que suele ser beneficioso definir criterios específicos que se adecuen a la familia de modelos con la cual estudiaremos nuestras señales. En la siguiente sección, presentaremos la familia de modelos anidados¹ de regresión trigonométrica, relacionare-

¹Se dice que la familia de modelos $\{\mathcal{M}_k\}_{k \in K}$ es anidada si los modelos más pequeños son casos particulares de los modelos más grandes. Para una familia de modelos anidados, la tarea de selección de modelos suele denominarse como el problema de *selección de orden*.

mos dichos modelos con el ANHM y discutiremos una serie de criterios desarrollados específicamente para la tarea de selección de orden en modelos de este tipo.

2.2. Selección de modelos de regresión trigonométrica

Un modelo de regresión trigonométrica tiene la forma

$$x(t) = \mu + \sum_{\ell=1}^{r_0} (\alpha_{\ell} \cos(\lambda_{\ell} t) + \beta_{\ell} \sin(\lambda_{\ell} t)) + u(t) + n(t), \quad (2.3)$$

donde el término $u(t)$ agrupa a los componentes no sinusoidales de la señal, como desvíos de la tendencia local o componentes autorregresivos. La constante μ puede ser interpretada como el valor medio de la señal para el caso que $\mathbb{E}[u(t)] = 0$. Finalmente, $n(t)$ representa a las fuentes de ruido que contaminan la señal y se suele describir mediante una distribución de probabilidad gaussiana con media igual a cero y una varianza σ^2 que, a priori, puede ser desconocida.

Es posible establecer una analogía entre este modelo de regresión y el ANHM monocomponente. Vemos que se puede alcanzar la ecuación (1.32) si se reemplaza la fase lineal $\lambda_{\ell} t$ en los términos trigonométricos de (2.3) por $2\pi\ell\phi(t)$ y los coeficientes de amplitud por $A(t)\alpha_{\ell}$ y $A(t)\beta_{\ell}$, donde α_{ℓ} y β_{ℓ} se obtienen al resolver el problema de mínimos cuadrados en (1.35) para $K = 1$. También podemos considerar, sin pérdida de generalidad, $\mu = 0$, es decir, señales con media cero. Bajo estas condiciones, el único parámetro a determinar en este modelo es la cantidad de términos trigonométricos r_0 .

A la hora de establecer un criterio de selección de modelos en el contexto de la regresión trigonométrica, la propuesta más simple resulta en minimizar el error cuadrático medio (MSE, del inglés *mean squared error*) entre los datos y la salida del modelo:

$$\text{MSE}(r) = \frac{1}{N} \sum_{n=0}^{N-1} \left(\mathbf{x}(n) - \mathbf{A}(n) \sum_{\ell \leq r} a_{\ell} \cos(2\pi\ell\phi(n)) + b_{\ell} \sin(2\pi\ell\phi(n)) \right)^2. \quad (2.4)$$

Podemos observar fácilmente que este criterio decrece monótonamente con el valor de r , por lo cual el modelo elegido siempre sería el de mayor r . Ante la ausencia de ruido, el MSE alcanzará el valor de cero cuando el orden del modelo sea el correcto (r_0), debido a que los términos de la sumatoria en (2.3) se anularían para $\ell > r_0$. Ante la presencia de ruido, sin embargo, estos términos ya no serán iguales a cero lo cual provoca una degradación de la señal estimada. Por lo tanto, se vuelve necesario encontrar un criterio que permita establecer una relación de compromiso entre el MSE y el orden del modelo, determinado por el parámetro r .

Una primera propuesta fue realizada por Wax y Kailath [70], quienes presentaron una estrategia para determinar el número de señales que componen los datos basándose en los criterios AIC y MDL:

$$\begin{aligned} \text{AIC}(r) &= -2\hat{L}(r) + 2r \\ \text{MDL}(r) &= -\hat{L}(r) + \frac{1}{2}r \log(N). \end{aligned}$$

Donde la complejidad del modelo depende del número de componentes r que conforman los datos. Para estimar el logaritmo de la máxima verosimilitud \hat{L} en función

2.2. SELECCIÓN DE MODELOS DE REGRESIÓN TRIGONOMÉTRICA

de r se construye una matriz de covarianza \mathbf{R} a partir de los datos y luego se aplica el teorema espectral para obtener la eigendescomposición de dicha matriz. A partir de esta descomposición, la máxima verosimilitud se calcula de la siguiente manera:

$$\hat{L}(r) = \log \left(\frac{\prod_{i=r+1}^N \lambda_i^{1/(N-r)}}{\frac{1}{N-r} \sum_{i=r+1}^N \lambda_i} \right)^{(N-r)N}, \quad (2.5)$$

donde $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_N$ son los eigenvalores de \mathbf{R} . Nótese aquí que el término dentro del paréntesis no es otra cosa que la razón entre la media geométrica y la media aritmética de los $N - r$ menores eigenvalores. En su trabajo, los autores mencionan que el número de señales componentes puede ser estimado de forma consistente con cualquier criterio de la forma $-\hat{L}(r) + \alpha(N)r$ que satisfaga las condiciones $\alpha(N) \rightarrow \infty$ y $\alpha(N)/N \rightarrow 0$. De esto, se deduce que AIC no es un estimador consistente mientras que MDL sí lo es.

Una de las principales limitantes de la propuesta anterior, es que su desempeño se deteriora notablemente en la presencia de señales no estacionarias con frecuencia variante en el tiempo. Sin embargo, la idea de establecer una relación de compromiso entre el tamaño del modelo y el error de aproximación sirve de base de una serie de criterios de selección de modelos aplicados al contexto de modelos de regresión trigonométrica. En el resto de esta sección, presentaremos diferentes criterios para determinar el número de términos de un modelo de regresión trigonométrica de la forma (2.3).

2.2.1. Riesgo sin sesgo

La definición de este criterio surge de observar que el error cuadrático medio es un estimador sesgado del riesgo de predicción para el modelo con parámetro r [71]. Es posible definir un estimador sin sesgo del riesgo de predicción haciendo [71, Ec. 2.22]:

$$\Phi_R(r) = \text{MSE}(r) + 2\sigma^2(2r + 1)/N, \quad (2.6)$$

donde el segundo término de la suma corrige el sesgo inherente al MSE y σ^2 es un estimador de la varianza del ruido. En general, este valor de varianza no se conoce a priori, por lo cual resulta necesario obtener un estimador $\tilde{\sigma}^2$ para poder aplicar este criterio. En el contexto del estudio de señales no estacionarias, Donoho propuso un estimador robusto para la varianza del ruido basándose en el análisis de operadores basados en onditas [72]. Recientemente, este estimador ha sido utilizado para la tarea de limpieza de ruido a partir del umbralado de la STFT [73]. A partir de analizar al espectrograma como una distribución de probabilidad, se obtiene el siguiente estimador del ruido:

$$\tilde{\sigma}^2 = \left(\frac{\text{median}(|\text{Re}(\mathbf{F}_x^g(n, k))|)}{0,6745 \|\mathbf{g}\|_2} \right)^2, \quad (2.7)$$

donde \mathbf{F}_x^g es una versión discretizada de la STFT de la señal discreta $\mathbf{x}(n)$ utilizando una ventana $\mathbf{g}(n)$ y siendo $\|\mathbf{g}\|_2$ la norma-2 de dicha ventana.

2.2.2. Validación Cruzada Generalizada

Otra forma de obtener un estimador no sesgado del riesgo de predicción a partir de $\text{MSE}(r)$ es mediante la validación cruzada generalizada (GCV, del inglés *generalized cross-validation*) [74], la cual se define de la siguiente manera:

$$\Phi_G(r) = \frac{\text{MSE}(r)}{\left(1 - \frac{2r}{N} - \frac{1}{N}\right)^2}, \quad (2.8)$$

Vemos que en este caso la corrección del sesgo se realiza al escalar el MSE por un factor $\left(1 - \frac{2r}{N} - \frac{1}{N}\right)^2$ el cual es siempre menor a 1. Este criterio, a diferencia de $\Phi_R(r)$, no requiere de un estimador de la varianza del ruido para su cálculo.

Tanto $\Phi_G(r)$ como $\Phi_R(r)$ fueron utilizados para estimar el número de componentes en modelos de regresión compuestos por términos trigonométricos y polinómicos [75]. En dicho trabajo, los autores fijaron el orden de la porción polinómica y utilizaron los criterios de selección de modelos $\Phi_R(r)$ y $\Phi_G(r)$ para estimar el número de componentes trigonométricos en series temporales estacionarias.

2.2.3. Criterio de Wang

En [76], Wang propuso un criterio de selección de modelos para regresión trigonométrica de la siguiente forma:

$$\Phi_W(r, c) = \log(\text{MSE}(r)) + cr(\log N)/N, \quad (2.9)$$

donde N es la longitud de la señal, r es el orden del modelo y c es una constante real positiva.

En su trabajo, Wang establece la siguiente condición para el parámetro c :

$$c > \frac{2 \max_f \{\hat{n}(f)\}}{(2\pi)^{-1} \int_{-\pi}^{\pi} \hat{n}(f) df}, \quad (2.10)$$

donde $\hat{n}(f)$ es una estimación del espectro de potencia del ruido. En general, es difícil saber cómo es la estructura espectral del ruido que contamina la señal. En el caso particular de señales contaminadas con ruido blanco estacionario, Quinn [77] demostró que cualquier criterio de la forma $\Phi_g(r) = N \log(\text{MSE}(r)) + 2rg(N)$ es un estimador consistente del orden del modelo si se cumple la siguiente condición:

$$\lim_{N \rightarrow \infty} g(N)/\log N > 1, \quad (2.11)$$

donde $g(N)$ representa el término de penalización que debe cumplir $N^{-1}g(N) \rightarrow 0$ cuando $N \rightarrow \infty$. Si comparamos $\Phi_g(r)$ con (2.9) notamos que $2g(N) = c \log N$ y, por lo tanto, la condición establece que $\Phi_W(r, c)$ será un estimador consistente de r siempre que se cumpla: $c > 2$. En los experimentos que discutiremos más adelante, se establecieron los siguientes valores para c : $\{2, 1; 5; 8; 12\}$.

2.2.4. Criterio de Kavalieris-Hannan

A partir del criterio de Wang y de sus propios desarrollos previos, Kavalieris y Hannan [78] propusieron el siguiente criterio para estimar el orden de un modelo de regresión trigonométrica:

$$\Phi_K(r, h) = \log\{\tilde{\sigma}_r^2(h)\} + (5r + h)(\log N)/N. \quad (2.12)$$

Aquí, $\tilde{\sigma}_r^2(h)$ es la varianza de predicción lineal de un modelo autorregresivo de orden h que se ajusta a los residuos $u_r(t) = x(t) - x_r(t)$, siendo $x_r(t)$ la reconstrucción de la señal según el modelo (2.3) utilizando r componentes trigonométricos. Esta varianza del modelo ajustado a los residuos cumple la misma función que $MSE(r)$ en los modelos anteriores.

El rango de valores para el parámetro h se define como $1 \leq h \leq H$. Según trabajos previos de los autores [79], si $\mathbf{u}_r(n)$ puede ser modelada razonablemente mediante una función de transferencia racional (modelo ARMA) entonces podemos establecer: $H = (\log(N))^\alpha$ con $0 < \alpha < \infty$. Sin pérdida de generalidad, seleccionamos $\alpha = 2$.

2.3. Algoritmo propuesto para reconstrucción mediante ANHM con orden óptimo

Los criterios que describimos en la sección anterior fueron descritos en el contexto de modelos de regresión trigonométrica según (2.3), es decir, que considera señales *no moduladas*. Nuestro objetivo es validar la aplicabilidad de estos mismos criterios para el caso de señales AM-FM, como las modeladas según (1.32). Para esto, proponemos un algoritmo de análisis y reconstrucción de señales no estacionarias basado en el ANHM, con estimación automática del número de armónicos de la WSF a partir de criterios de selección de modelos de regresión trigonométrica. Dicho procedimiento se describe en el Algo. 1 y la validación del algoritmo se realizará a través de experimentos sobre señales sintéticas generadas a partir de (1.33) y señales reales del ámbito biomédico. Ponemos especial énfasis en los pasos 6 a 12, donde llevamos adelante el procedimiento de estimación de orden óptimo para la WSF utilizando los criterios de selección de modelos de regresión trigonométrica. Nótese que este algoritmo se aplica a señales que siguen el ANHM *monocomponente* (1.31). Veremos más adelante cómo se puede extender el mismo para el caso de señales multicomponente.

En los siguientes experimentos y simulaciones, se especifica el rango de valores admisibles para el parámetro r como el intervalo $1 \leq r \leq r_{max}$. La cota superior r_{max} se calcula como: $r_{max} = \left\lfloor \frac{f_s/2}{f_{max}} \right\rfloor$, donde f_s es la frecuencia de muestreo y $f_{max} = \max_t f(t)$ es el máximo valor de la frecuencia instantánea del componente fundamental $f(t) = \phi'(t)$ para todo t .

Para todas las señales bajo estudio, se calcularon los cuatro criterios de selección de orden que se detallaron en la sección anterior y se eligió el valor óptimo r^* de la siguiente manera

$$r^* = \underset{r \in [1, r_{max}]}{\operatorname{argmin}} \Phi_X(r); \quad (2.13)$$

para $X \in \{G, R, W, K\}$.

2.3. ALGORITMO PROPUESTO PARA RECONSTRUCCIÓN MEDIANTE ANHM CON ORDEN ÓPTIMO

Algoritmo 1 Análisis de señal y reconstrucción utilizando el ANHM con orden óptimo para la WSF

Entrada: Señal $\mathbf{x}(n)$. Ventana de análisis para la STFT $\mathbf{g}(n)$. Número de bins de frecuencia K . Parámetros de extracción de crestas λ , μ y máximo salto admisible en frecuencia I . Ancho de banda de reconstrucción Δ . Máximo orden admisible r_{max} . Criterios de estimación de orden X para $X \in \{G, R, W, K\}$.

1: Dada una señal $\mathbf{x}(n)$ de la forma (1.32) calcule su STFT en tiempo discreto como

$$\mathbf{F}_{\mathbf{x}}^{\mathbf{g}}(n, m) = \sum_{\nu=0}^{N-1} \mathbf{x}(\nu) \mathbf{g}(\nu - n) e^{-i2\pi \frac{m}{M} \frac{N}{T} (\nu - n)}, \quad (2.14)$$

con $m = 0, 1, \dots, M - 1$ y $n = 0, 1, \dots, N - 1$.

2: Dada la TFR $\mathbf{F}_{\mathbf{x}}^{\mathbf{g}}$, encontrar la cresta asociada al componente fundamental $\mathbf{c}(n) \approx \phi'(n)$. Para esto, resolver el siguiente problema de optimización:

$$\mathbf{c}^*(n) = \arg \max_{\mathbf{c}(n)} \left(\sum_{n=0}^{N-1} |\mathbf{F}_{\mathbf{x}}(n, \mathbf{c}(n))|^2 - \sum_{n=1}^{N-1} \lambda |\mathbf{c}(n) - \mathbf{c}(n-1)|^2 - \sum_{n=2}^{N-1} \mu |\mathbf{c}(n) - 2\mathbf{c}(n-1) + \mathbf{c}(n-2)|^2 \right).$$

Este problema se resuelve mediante el algoritmo voraz discutido en la Sec. 1.3.3.

3: Reconstruir el componente fundamental de $\mathbf{x}(n)$ integrando la STFT alrededor de la cresta $\mathbf{c}^*(n)$:

$$\mathbf{y}(n) = \frac{1}{\mathbf{g}(0)} \sum_{|m - \mathbf{c}^*(n)| < \Delta} \mathbf{F}_{\mathbf{x}}^{\mathbf{g}}(n, m)$$

4: Calcular las estimaciones para la amplitud $\tilde{\mathbf{A}}(n) = |\mathbf{y}(n)|$ y la modulación de fase $\tilde{\phi}(n) = \angle \mathbf{y}(n)$.

5: **para** $r = 1$ hasta r_{max} **hacer**

6: Construir el diccionario \mathbf{C}_r según (1.36) considerando $K = 1$, formado por los vectores columna: $\mathbf{c}_\ell = \mathbf{A}(n) \cos(2\pi \ell \phi(n))$ y $\mathbf{d}_\ell = \mathbf{A}(n) \sin(2\pi \ell \phi(n))$, para $\ell = 1, \dots, r$.

7: Encontrar el vector de coeficientes óptimos $\hat{\gamma}_r$ reemplazando \mathbf{C}_r en (1.38).

8: Calcular $\mathbf{x}_r = \mathbf{C}_r \hat{\gamma}_r$, la aproximación de orden r de \mathbf{x} .

9: Calcular el valor del criterio de selección $\Phi_X(r)$.

10: **fin para**

11: Encontrar el orden óptimo del modelo r^* utilizando (2.13).

12: Construir \mathbf{C}_{r^*} como se describió anteriormente, resolver para $\hat{\gamma}_{r^*}$ utilizando (1.38) y luego calcular $\mathbf{x}_{r^*}(n) = \mathbf{C}_{r^*} \hat{\gamma}_{r^*}$.

Salida: Señal reconstruida $\mathbf{x}_{r^*}(n)$, vector de coeficientes de Fourier $\hat{\gamma}_{r^*}$ y orden óptimo r^* para el modelo ANHM.

2.4. VALIDACIÓN EN SEÑALES SINTÉTICAS

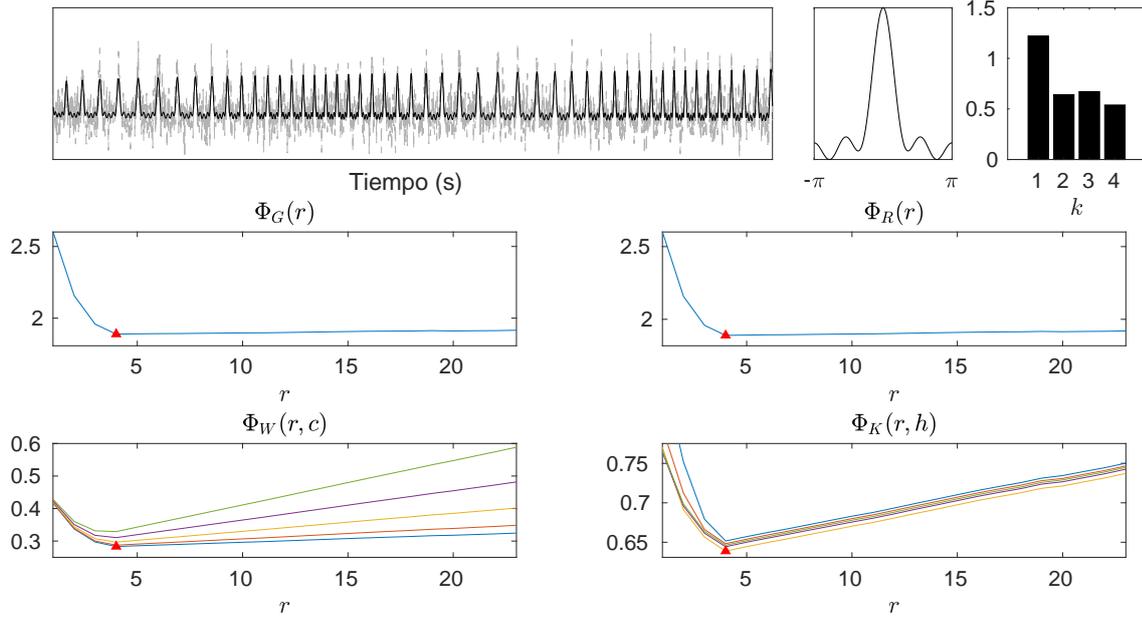


Figura 2.1: Criterios de selección de modelos en señales no estacionarias ($r_0 = 4$). Primera fila: Señal no estacionaria $x(t)$ contaminada con un alto nivel de ruido en gris con la señal original (sin ruido) superpuesta en negro. Función de forma de onda para la señal $x(t)$ y módulo de sus coeficientes de Fourier $|\hat{s}(k)|$. Segunda fila: Criterios de selección $\Phi_G(r)$ y $\Phi_R(r)$ calculados a partir de $x(t)$. Tercera fila: Criterios de selección $\Phi_W(r, c)$ y $\Phi_K(r, h)$ calculados a partir de $x(t)$.

2.4. Validación en señales sintéticas

Comenzaremos describiendo una serie de experimentos realizados sobre señales sintetizadas a partir del ANHM y evaluaremos el desempeño de nuestra propuesta para la estimación del orden óptimo en estas señales ante la presencia de ruido.

En todos los casos, las señales se discretizaron muestreándolas a una frecuencia de $f_s = 3000$ muestras por segundo. Calculamos la STFT como se muestra en Algo. 1 utilizando una ventana gaussiana discreta $\mathbf{g}(n) = e^{-2\sigma n^2}$ con $\sigma = 10^{-4}$ e intervalo de frecuencia discreta $0 \leq k \leq f_s/2$ con paso de frecuencia de f_s/N . En el algoritmo de extracción de la cresta, consideramos $\lambda = \mu = 0,1$ y fijamos el salto máximo del índice de frecuencia en $I = 10$. Para el procedimiento de estimación de amplitud y fase basado en la integración alrededor de la cresta fundamental, utilizamos $\Delta = 50$ índices de frecuencia.

Antes de continuar hacemos la observación de que, dado que conocemos los valores exactos de $A(t)$ y $\phi(t)$ de nuestra señal simulada, podríamos utilizarlos directamente para reconstruir la señal en lugar de estimarlos mediante el procedimiento de extracción de crestas. La presencia de ruido provoca errores en las estimaciones, lo que degrada la calidad de las señales reconstruidas y sería esperable que esto repercuta en los resultados del experimento. No obstante, optamos por estimar $A(t)$ y $\phi(t)$ utilizando Algo. 1 con el fin de simular un contexto más realista, dónde es posible evaluar el procedimiento de reconstrucción en su conjunto.

2.4.1. Análisis de los criterios de selección de modelos

Empezamos analizando una señal modulada en amplitud con modulación de frecuencia no lineal y patrón de oscilación no sinusoidal:

$$x(t) = \left(1 + 0,05\sqrt{t}\right) s \left(2\pi \left(70t + \frac{15}{2\pi} \cos(2\pi t)\right)\right) + n(t), \quad (2.15)$$

con $t \in [0, 1]$, y donde el término $n(t)$ representa ruido blanco gaussiano de media cero. La frecuencia fundamental instantánea de la señal es: $\phi'(t) = 70 - 15 \sin(2\pi t)$. Consideramos, en este primer ejemplo, dos funciones de forma de onda de ancho de banda limitado, es decir, s tales que $\hat{s}(\ell) = 0$ para $\ell > r_0$. El nivel de ruido se cuantifica mediante la relación señal a ruido (SNR, del inglés *signal-to-noise ratio*) como: $\text{SNR} = 20 \log(\|x(t)\|_2 / \|n(t)\|_2)$, siendo \log el logaritmo natural. Nótese que $\|n(t)\|_2$ es el desvío estándar del ruido.

En la Fig. 2.1 se muestra una señal cuya forma de onda está compuesta por un número reducido de armónicos ($r_0 = 4$) y que a su vez está contaminada con un alto nivel de ruido (SNR = 0 dB). Los valores de los criterios de selección de modelos en función del orden r para esta realización de la señal se muestran en la segunda y tercera fila de la Fig. 2.1. Notamos que para $\Phi_G(r)$, $\Phi_R(r)$ se observa una caída sostenida del valor del criterio hasta el valor $r^* = 4$, donde se alcanza el mínimo del criterio. Para el caso de $\Phi_W(r, c)$ y $\Phi_K(r, h)$ también se alcanza el mínimo r^* para el orden correcto pero el descenso del criterio es más suave en el caso de $\Phi_W(r)$ para ordenes menores al óptimo. Además, para $r > r^*$ se evidencia el efecto del factor de penalización para diferentes valores de c y h para $\Phi_W(r, c)$ y $\Phi_K(r, h)$, respectivamente. Vemos que en el caso de $\Phi_W(r, c)$ la constante c afecta la pendiente de la porción lineal del criterio, mientras que para $\Phi_K(r, h)$ la constante h afecta la ordenada al origen de la curva. Por lo tanto, la elección del coeficiente c para el criterio $\Phi_W(r, c)$ es de suma importancia ya que podemos incurrir en una subestimación del orden si c es grande, o bien, sobrestimar el orden si el c elegido es pequeño.

En la Fig. 2.2 mostramos un ejemplo similar al anterior, pero ahora con un mayor número de armónicos para la WSF ($r_0 = 9$). Vemos que para este caso los criterios $\Phi_G(r)$ y $\Phi_R(r)$ estiman correctamente el orden del modelo, mientras que $\Phi_W(r, c)$ y $\Phi_K(r, c)$ tienden a subestimar el orden para el modelo, arrojando los valores $r^* = 7$ y $r^* = 5$, respectivamente. Notamos además que para todos los criterios el cambio de dirección de la curva producido en el “codo” ya no es tan abrupto sino que se trata de una curva más suave. Cabe aclarar que la condición recién descrita contempla el peor caso para el desempeño de los criterios de selección. Como veremos en la siguiente sección, bajo condiciones de ruido moderadas, todos los criterios considerados tienden a desempeñarse correctamente bajo el supuesto de frecuencia variable en el tiempo.

2.4.2. Estimación automática del orden óptimo de la función de forma de onda en señales AM-FM

Las simulaciones descritas en el apartado anterior corresponden a sólo dos realizaciones de señales ruidosas con un número fijo de componentes armónicos. Es necesario evaluar la robustez de los criterios en un conjunto de señales más amplio, donde se contemple la mayor cantidad de fuentes de variabilidad posibles. Las principales causas de

2.4. VALIDACIÓN EN SEÑALES SINTÉTICAS

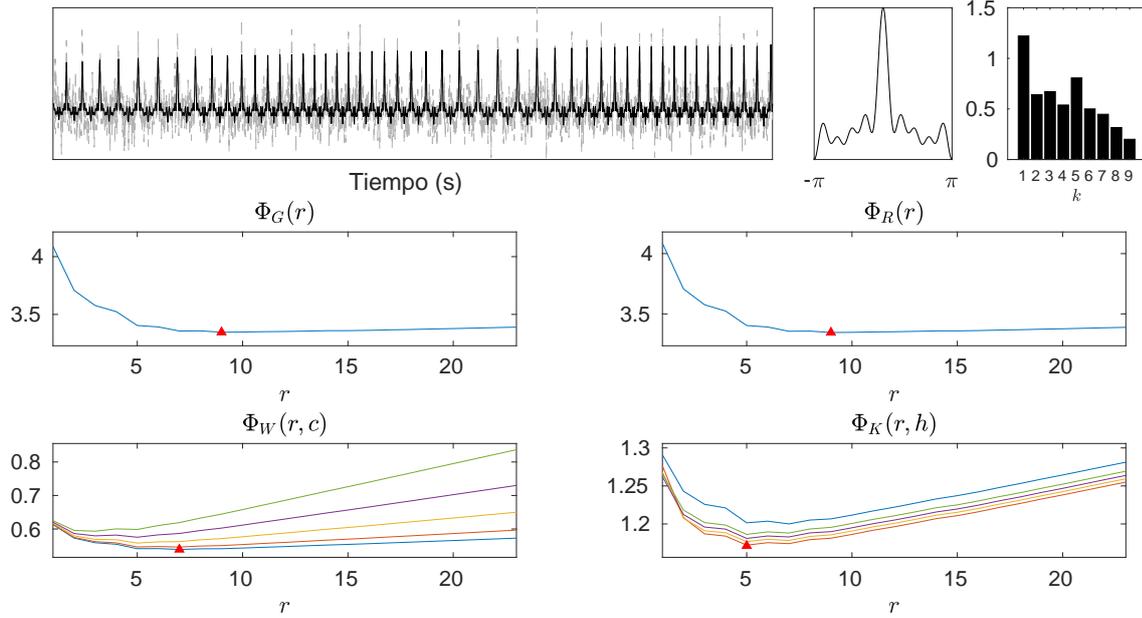


Figura 2.2: Criterios de selección de modelos en señales no estacionarias ($r_0 = 9$). Primera fila: Señal no estacionaria $x(t)$ contaminada con un alto nivel de ruido en gris con la señal original (sin ruido) superpuesta en negro. Función de forma de onda para la señal $x(t)$ y módulo de sus coeficientes de Fourier $|\hat{s}(k)|$. Segunda fila: Criterios de selección $\Phi_G(r)$ y $\Phi_R(r)$ calculados a partir de $x(t)$. Tercera fila: Criterios de selección $\Phi_W(r, c)$ y $\Phi_K(r, h)$ calculados a partir de $x(t)$.

variación en el procedimiento de selección del modelo se deben al nivel de ruido y a la amplitud relativa de los armónicos, principalmente los de alta frecuencia. Los criterios pueden subestimar el orden si la amplitud relativa de los armónicos de orden superior es pequeña en comparación con el nivel de ruido. Para minimizar el efecto de estos factores, generamos 1,000 realizaciones y aleatorizamos la amplitud del segundo al último armónico en el rango $[0,1, 0,9]$ manteniendo el armónico fundamental como dominante (es decir, $|\hat{s}(1)| > |\hat{s}(k)|$ para $k = 2, \dots, r_0$). De este modo, sintetizamos diferentes funciones de forma de onda s , construyendo señales de la forma $x(t) = A(t)s(2\pi\phi(t))$, con $A(t) = 1 + 0,05\sqrt{t}$. También consideramos diferentes tipos de modulaciones en frecuencia:

- $\phi_{NM}(t) = 100t$ (Sin modulación).
- $\phi_{LM}(t) = \frac{15}{2}t^2 + 50t$ (Modulación Lineal).
- $\phi_{SM}(t) = \frac{15}{2\pi} \cos(2\pi t) + 70t$ (Modulación Sinusoidal).

Consideramos en total, cuatro niveles de SNR: 0, 5, 10 y 15 dB. Fijamos los valores posibles de r_0 en 1, 3, 4, 6 y 9. Utilizamos los mismos parámetros para la STFT y el procedimiento de extracción de crestas que en la sección anterior. En total, se generaron 1,000 realizaciones para cada combinación de SNR, r_0 y $\phi(t)$; y se calcularon todos los criterios para cada caso. Mostramos los espectrogramas para señales ruidosas de cada tipo de modulación en fase en la Fig. 2.3 utilizando diferentes números de armónicos. Es

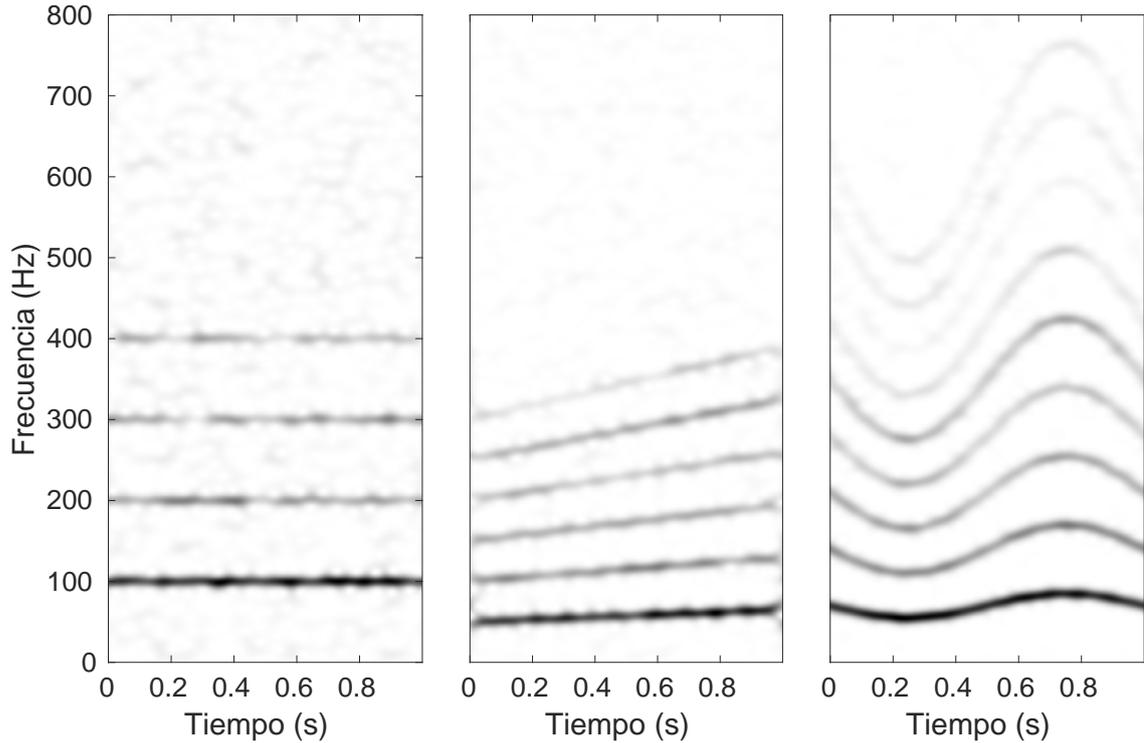


Figura 2.3: Espectrogramas de señales no estacionarias contaminadas con ruido. Izquierda: Espectrograma de $x_{NM}(t)$ contaminado con ruido aditivo de 0 dB y $r_0 = 4$. Centro: Espectrograma de $x_{LM}(t)$ contaminado con ruido aditivo de 5 dB y $r_0 = 6$. Derecha: Espectrograma de $x_{SM}(t)$ contaminado con ruido aditivo de 10 dB y $r_0 = 9$.

evidente que a medida que aumenta el nivel de ruido, los armónicos de orden superior que forman crestas de menor amplitud en el plano T-F quedan oscurecidos por el ruido, lo que dificulta su detección. Las tasas de acierto de los diferentes criterios se muestran en la Tabla 2.1 para las diferentes condiciones de cantidad de armónicos, ruido y modulación en fase. En particular, para los criterios $\Phi_W(r, c)$ y $\Phi_K(r, h)$ el orden óptimo se elige a partir de la coordenada r del mínimo global en 2D del criterio considerando los rangos de los parámetros c y h establecidos en las secciones correspondientes.

En primera instancia, analizamos los resultados para la señal sin modulación en frecuencia. En este caso, cuando la SNR es baja (0 dB), $\Phi_G(r)$ y $\Phi_R(r)$ funcionan bien como estimadores de orden para $r_0 \leq 6$, presentando una tasa de acierto mayor al 75 %. Vemos que el peor resultado en este rango se da para $r_0 = 1$, indicando una tendencia a la sobrestimación del orden óptimo para estos criterios en este caso. Los criterios $\Phi_W(r, c)$ y $\Phi_K(r, h)$ presentan un mejor desempeño para este rango de valores, particularmente para número de armónicos $r_0 \leq 4$. Cuando la señal está compuesta por un gran número de componentes armónicos ($r_0 = 9$), vemos que los criterios no son capaces de estimar correctamente el orden del modelo cuando la SNR es baja. A medida que aumenta la SNR, el desempeño de los criterios mejora. Los criterios $\Phi_G(r)$ y $\Phi_R(r)$ tienden a desempeñarse mejor cuando la señal está compuesta por muchos armónicos ($r_0 = 6$ y $r_0 = 9$), mientras que en contraposición los criterios $\Phi_W(r, c)$ y $\Phi_K(r, h)$ funcionan mejor para un número bajo de armónicos ($r_0 = 1$ y $r_0 = 3$).

Respecto a la modulación en frecuencia, vemos que el comportamiento de los criterios

2.4. VALIDACIÓN EN SEÑALES SINTÉTICAS

$A(t)s(2\pi\phi_{NM}(t))$										
	$\Phi_G(r)$					$\Phi_R(r)$				
	1	3	4	6	9	1	3	4	6	9
0 dB	0.75	0.81	0.83	0.82	0.58	0.80	0.84	0.85	0.82	0.61
5 dB	0.78	0.82	0.82	0.89	0.96	0.83	0.87	0.86	0.89	0.97
10 dB	0.78	0.79	0.83	0.90	0.99	0.83	0.84	0.87	0.91	0.98
15 dB	0.77	0.84	0.84	0.90	0.99	0.82	0.89	0.89	0.91	0.99
	$\Phi_W(r, c)$					$\Phi_K(r, h)$				
	1	3	4	6	9	1	3	4	6	9
0 dB	1	0.97	0.94	0.79	0.11	1	0.89	0.84	0.33	0
5 dB	1	1	0.99	0.99	0.86	1	0.99	0.98	0.93	0.59
10 dB	1	1	1	1	0.99	1	1	1	0.99	0.98
15 dB	0.99	1	1	1	1	1	1	1	1	0.99
$A(t)s(2\pi\phi_{LM}(t))$										
	$\Phi_G(r)$					$\Phi_R(r)$				
	1	3	4	6	9	1	3	4	6	9
0 dB	0.76	0.79	0.82	0.85	0.51	0.81	0.83	0.84	0.85	0.53
5 dB	0.80	0.81	0.85	0.90	0.94	0.84	0.84	0.90	0.91	0.93
10 dB	0.82	0.84	0.86	0.93	0.98	0.85	0.88	0.90	0.94	0.97
15 dB	0.87	0.88	0.92	0.97	0.99	0.86	0.88	0.91	0.96	0.98
	$\Phi_W(r, c)$					$\Phi_K(r, h)$				
	1	3	4	6	9	1	3	4	6	9
0 dB	1	0.97	0.95	0.79	0.10	1	0.92	0.82	0.35	0
5 dB	1	0.99	0.99	0.99	0.87	1	0.99	0.98	0.93	0.65
10 dB	1	1	1	1	0.99	1	1	1	0.99	0.95
15 dB	1	1	1	1	1	1	1	1	1	0.99
$A(t)s(2\pi\phi_{SM}(t))$										
	$\Phi_G(r)$					$\Phi_R(r)$				
	1	3	4	6	9	1	3	4	6	9
0 dB	0.77	0.80	0.83	0.85	0.50	0.81	0.85	0.87	0.87	0.50
5 dB	0.79	0.80	0.83	0.89	0.93	0.83	0.87	0.89	0.93	0.94
10 dB	0.80	0.80	0.83	0.90	0.97	0.83	0.88	0.90	0.95	0.99
15 dB	0.79	0.81	0.85	0.90	0.98	0.84	0.90	0.92	0.96	1
	$\Phi_W(r, c)$					$\Phi_K(r, h)$				
	1	3	4	6	9	1	3	4	6	9
0 dB	1	0.98	0.95	0.80	0.12	1	0.91	0.82	0.42	0
5 dB	1	1	1	0.99	0.88	1	0.99	0.98	0.92	0.70
10 dB	1	1	1	1	0.99	1	1	1	0.99	0.95
15 dB	1	1	1	0.99	1	1	1	1	1	1

Tabla 2.1: Selección automático del orden del ANHM. En cada celda se representa la tasa de acierto (sobre $N = 1,000$ señales) del criterio para selección el orden correcto del modelo r_0 para los diferentes tipos de modulación en fase y valores de relación señal a ruido. Referencias: $\Phi_G(r)$ - Validación cruzada generalizada. $\Phi_R(r)$ - Riesgo sin sesgo. $\Phi_W(r, c)$ - Criterio de Wang. $\Phi_K(r, h)$ - Criterio de Kavalieris-Hannan.

descrito anteriormente se sigue respetando tanto para $x_{LM}(t) = A(t)s(2\pi\phi_{LM}(t))$ cómo para $x_{SM}(t) = A(t)s(2\pi\phi_{SM}(t))$. Los criterios $\Phi_G(r)$ y $\Phi_R(r)$ tienen una ventaja sobre $\Phi_W(r, c)$ y $\Phi_K(r, h)$ en situaciones donde el nivel de ruido es alto (SNR = 0 dB) o moderadamente alto (SNR = 5 dB). Para los niveles de ruido moderadamente bajo (SNR = 10 dB) y bajo (SNR = 15 dB) todos los criterios tienden a desempeñarse adecuadamente, siendo $\Phi_W(r, c)$ y $\Phi_K(r, h)$ levemente superiores a $\Phi_G(r)$ y $\Phi_R(r)$.

En general, todos los criterios permiten estimar correctamente el valor real del número de armónicos r_0 en la mayoría de los casos. El inconveniente principal se produce en la situación de señales con un gran número de componentes armónicos inmersas en niveles de ruido elevados. Además, notamos que en el caso de señales con bajo número de componentes armónicos los criterios $\Phi_W(r, c)$ y $\Phi_K(r, h)$ tienden a desempeñarse mejor que el resto, mientras que $\Phi_G(r)$ y $\Phi_R(r)$ son preferibles en situaciones donde la señal presenta un elevado número de armónicos.

2.4.3. Estimación del orden del modelo en señales multicomponentes

En esta sección, extendemos la estimación automática del número de componentes armónicos en el ANHM al caso multicomponente, es decir, cuando $K > 1$ en (1.29). En particular, consideramos la superposición de dos señales sintéticas con oscilación no sinusoidal y modulación en frecuencia variante en el tiempo. Para este caso, es necesario determinar dos parámetros desconocidos: r_1 y r_2 , el número de armónicos para la primera y la segunda WSF, respectivamente. Para este fin, definimos versiones modificadas de los criterios de selección del orden óptimo sustituyendo la variable escalar de orden del modelo r por una variable bidimensional (r_1, r_2) . La definición de los criterios “bidimensionales” modificados es la siguiente:

$$\begin{aligned}\Phi_G(r_1, r_2) &= N^2 \text{MSE}(r_1, r_2) / (N - 2(r_1 + r_2) - 1)^2, \\ \Phi_R(r_1, r_2) &= \text{MSE}(r_1, r_2) + 2\hat{\sigma}^2(2(r_1 + r_2) + 1) / N, \\ \Phi_W(r_1, r_2, c) &= \log(\text{MSE}(r_1, r_2)) + c(r_1 + r_2)(\log N) / N, \\ \Phi_K(r_1, r_2, h) &= \log\{\hat{\sigma}_{r_1, r_2}^2(h)\} + (5(r_1 + r_2) + h)(\log N) / N,\end{aligned}$$

donde los errores de aproximación se calculan a partir de la reconstrucción de la señal $\mathbf{x}_{r_1, r_2}(n)$ sumando los primeros r_1 términos de la primer componente y los primeros r_2 términos del segundo componente. La estimación de los coeficientes de Fourier se realiza de la misma manera que antes, sólo que ahora la matriz \mathbf{C}_{r_1, r_2} incluirá las funciones de amplitudes instantáneas ($\mathbf{A}_1(n)$ y $\mathbf{A}_2(n)$) y fases instantáneas ($\phi_1(n)$ $\phi_2(n)$) de cada componente. Posteriormente, se calculan los criterios para cada dupla (r_1, r_2) en el dominio rectangular $1 \leq r_1 \leq r_{1, max}$, $1 \leq r_2 \leq r_{2, max}$ y se seleccionan las coordenadas (r_1^*, r_2^*) del mínimo global del criterio como los valores óptimos para los ordenes de cada función de forma de onda.

En estas simulaciones, generamos una señal multicomponente $s(t) = s_1(t) + s_2(t)$, con

$$s_1(t) = \sum_{\ell=1}^{r_1} a_{1, \ell} \cos(2\pi\ell(120t + 15t^2)) \quad (2.16)$$

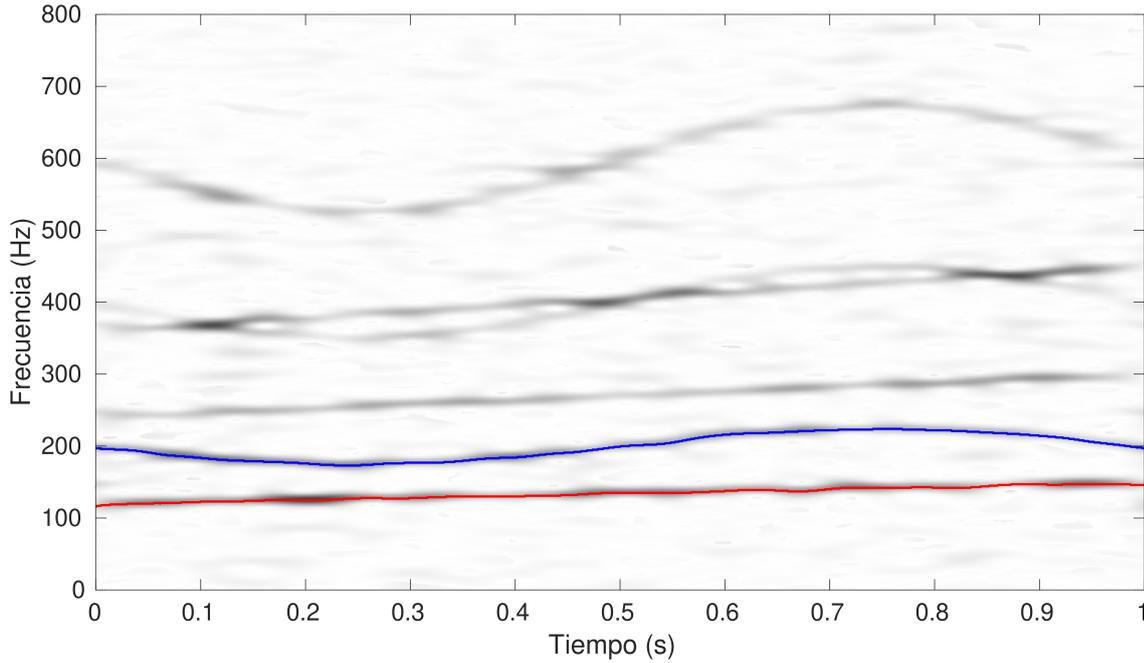


Figura 2.4: Espectrograma de una señal no estacionaria multicomponente. Espectrograma de la señal $s(t) = s_1(t) + s_2(t)$ contaminada con ruido aditivo de 0 dB. Las crestas fundamentales de los componentes $s_1(t)$ y $s_2(t)$ se resaltan en azul y rojo, respectivamente.

siendo una suma de chirps lineales, y

$$s_2(t) = \sum_{\ell=1}^{r_2} a_{2,\ell} \cos \left(2\pi\ell \left(200t + \frac{25}{2\pi} \cos(2\pi t) \right) \right) \quad (2.17)$$

siendo una suma de chirps sinusoidales. Estas señales se discretizaron a una frecuencia de muestreo de 3000 Hz en el intervalo $[0, 1]$. Se fijó $r_1 = 4$, $r_2 = 3$, $a_{1,1} = a_{2,1} = 1$, y se aleatorizaron las amplitudes de los coeficientes del segundo al último armónico para cada realización en el intervalo $[0,1; 0,9]$ (al hacer esto, se mantuvo al armónico fundamental como dominante). Se utilizaron los parámetros mencionados al inicio de la Sec. 2.4 para calcular la STFT, extraer las crestas fundamentales y reconstruir los componentes de la señal.

Respecto a los criterios, fijamos $c = 2,1$ para $\Phi_W(r_1, r_2, c)$ y $h = (\log N)^2$ para $\Phi_K(r_1, r_2, h)$. Se agregó ruido a $s(t)$ en 3 niveles diferentes de SNR: 0, 10 y 20 dB. Los criterios fueron calculados para 1,000 realizaciones de señales ruidosas para cada nivel de ruido y elegimos el par de órdenes óptimo (r_1^*, r_2^*) . En la Fig. 2.4 se muestra el espectrograma de una realización de esta señal multicomponente con SNR de 0 dB. Queda claro que los armónicos de orden superior de ambos componentes se solapan en el plano tiempo-frecuencia para este tipo de señales pero que aun es posible determinar las crestas fundamentales de cada componente y obtener a partir de ellas los estimadores para las amplitudes y fases instantáneas de las señales componentes.

Los resultados de estas simulaciones con SNR de 0 dB se presentan en la Fig. 2.5. Mostramos el mapa de intensidad 2D junto con los histogramas marginales para r_1^* y r_2^* , donde podemos ver que los criterios proporcionan buenos estimadores para los parámetros

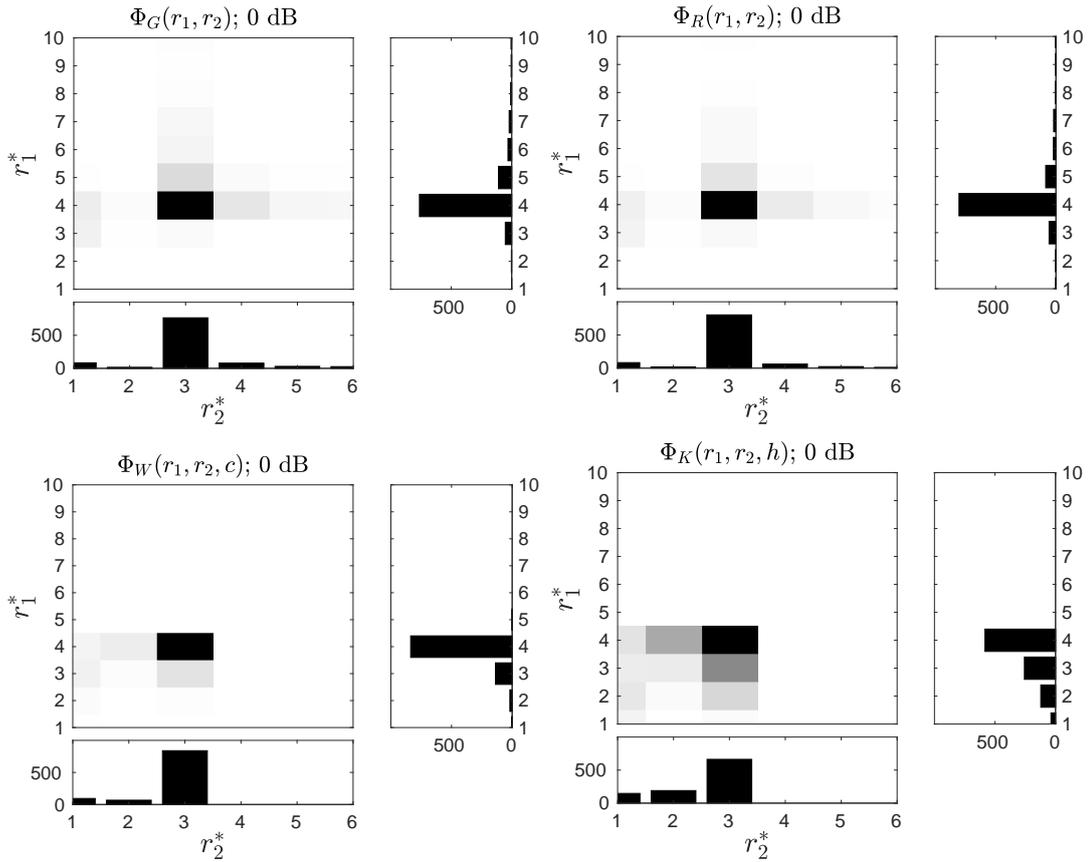


Figura 2.5: Resultados de la estimación del orden del ANHM en señales multicomponente $s(t) = s_1(t) + s_2(t)$ con SNR de 0 dB. Para cada criterio, se muestran el mapa de intensidad y los histogramas marginales. Todos los criterios estiman correctamente el orden de ambos componentes, como muestran las modas en los histogramas.

de orden del modelo r_1 y r_2 , ya que las modas de los histogramas son los órdenes correctos para el modelo utilizando todos los criterios. En particular, $\Phi_G(r_1, r_2)$ y $\Phi_R(r_1, r_2)$ tienen una dispersión más amplia y una mayor tendencia a sobreestimar el orden del modelo que $\Phi_W(r_1, r_2, c)$ y $\Phi_K(r_1, r_2, h)$.

En las Figs. 2.6 y 2.7 se muestran los resultados para SNR = 10 y 20 dB, respectivamente. Vemos que el desempeño de los criterios mejora a medida que la SNR aumenta, ya que los histogramas de orden óptimos se encuentran más concentrados en torno al par de órdenes correcto. Estos resultados indican que los criterios de selección del modelo de regresión trigonométrica son adecuados para la detección automática de las distintas WSFs que componen una señal con frecuencias instantáneas variables en el tiempo.

2.5. Aplicación 1: limpieza de ruido en señales de onda de pulso

Las técnicas de análisis de señales basadas en el ANHM se han aplicado a numerosas tareas de procesamiento de señales no estacionarias en el ámbito biomédico, incluyendo el modelado de señales de pulso [53], la separación de ECG fetal [55] y la clasificación

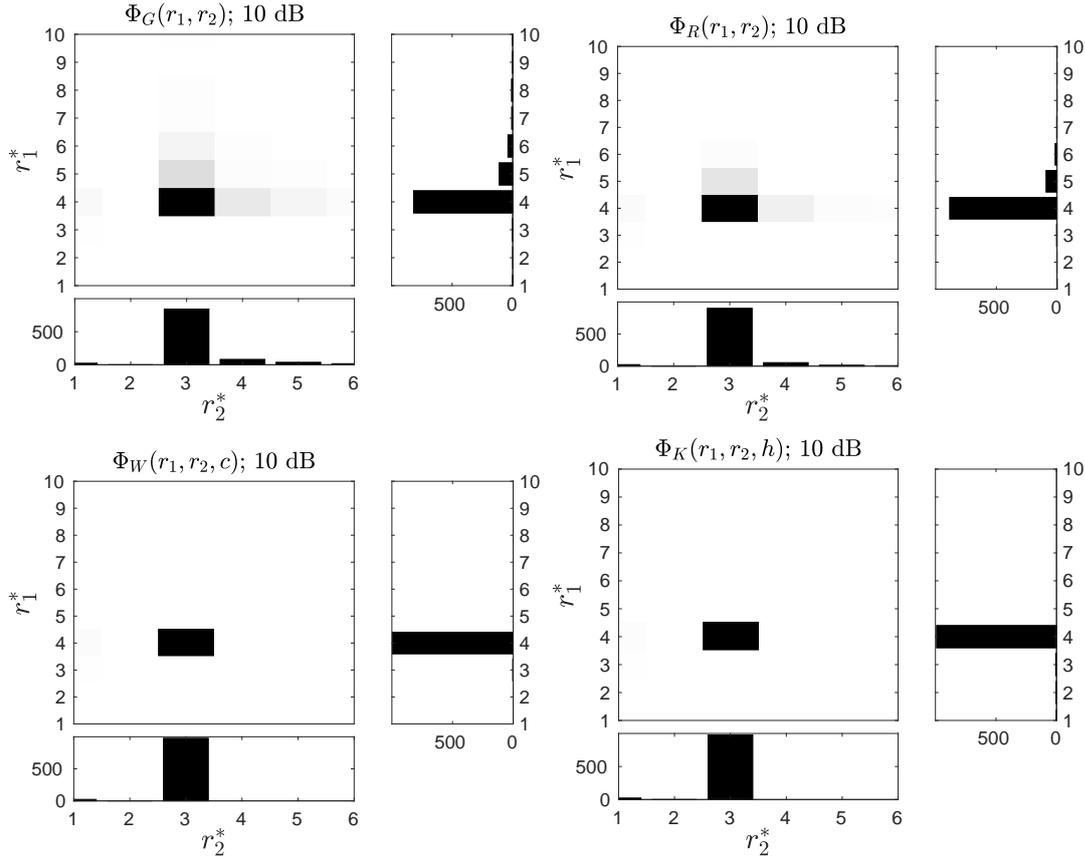


Figura 2.6: Resultados de la estimación del orden del ANHM en señales multicomponente $s(t) = s_1(t) + s_2(t)$ con SNR de 10 dB. Para cada criterio, se muestran el mapa de intensidad y los histogramas marginales. Todos los criterios estiman correctamente el orden de ambos componentes, como muestran las modas en los histogramas.

de latidos [80], entre otros. En esta sección, evaluaremos la viabilidad de los criterios de selección de modelos trigonométricos como forma de determinar el número óptimo de armónicos para una función de forma de onda s según el modelo (1.31), cuando se aplica a señales biomédicas contaminadas con ruido. En particular, analizaremos señales procedentes de la base de datos *Pulsewave Database* (PWDB) [81, 82]. Esta base de datos contiene señales de ondas de pulso generadas a partir de un modelo que simula la hemodinámica de individuos sanos. Con este modelo, es posible generar ondas de pulso correspondientes a diferentes tramos del sistema circulatorio. En estos experimentos, compararemos el desempeño de nuestro algoritmo para dos tramos diferentes del sistema vascular: la aorta abdominal y la arteria tibial anterior.

2.5.1. Diseño experimental

Con el fin de evaluar el rendimiento del ANHM cuando tratamos con señales biomédicas contaminadas con ruido, sintetizamos registros de ondas de pulso basadas en las formas de onda contenidas en la PWDB y añadimos una cantidad conocida de ruido a estas señales. Seguidamente, aplicamos el procedimiento de extracción de crestas y reconstruimos las señales utilizando el ANHM estimando automáticamente r_0 , dando como resultado las versiones reconstruidas $x_r(t)$. Por último, calculamos la SNR de salida de la

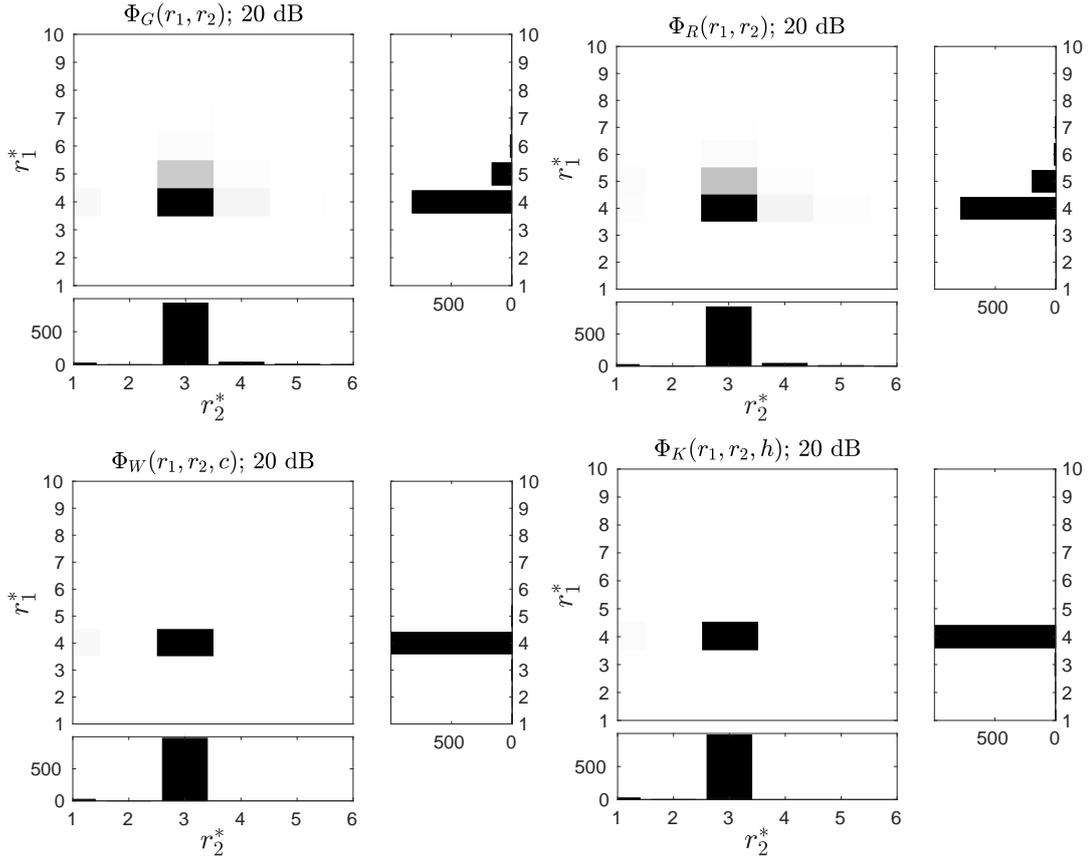


Figura 2.7: Resultados de la estimación del orden del ANHM en señales multicomponente $s(t) = s_1(t) + s_2(t)$ con SNR de 20 dB. Para cada criterio, se muestran el mapa de intensidad y los histogramas marginales. Todos los criterios estiman correctamente el orden de ambos componentes, como muestran las modas en los histogramas.

siguiente manera:

$$\text{SNR}_{out} = 20 \log \left(\frac{\|x(t)\|_2}{\|x(t) - x_r(t)\|_2} \right). \quad (2.18)$$

También comparamos los resultados de la aplicación de este procedimiento con el esquema de eliminación de ruido basado en el umbralado de la STFT, utilizando un umbral robusto [73]. Para nuestras simulaciones, aplicamos tanto umbrales suaves como duros, con el umbral fijado en $\eta = 3\hat{\sigma}\|g\|_2$, con $\hat{\sigma}$ dado por (2.7).

Las señales de onda de pulso sintetizadas tienen la forma

$$x(t) = (1 + 0,02\sqrt{t})s(2\pi\phi(t)), \quad (2.19)$$

siendo s una forma de onda de pulso elegida aleatoriamente de la base de datos, y $\phi(t)$ una función de fase sintética que emula la fase asociada a la variabilidad fisiológica de la frecuencia cardíaca (VFC):

$$\phi'(t) = 1 + 0,035a \sin(2\pi f_L t) + 0,035b \sin(2\pi f_H t), \quad (2.20)$$

con f_L correspondiente a la frecuencia asociada con el contenido energético de baja frecuencia (0,01 – 0,15 Hz) de la VFC, y f_H asociada con el contenido energético de alta

2.5. APLICACIÓN 1: LIMPIEZA DE RUIDO EN SEÑALES DE ONDA DE PULSO

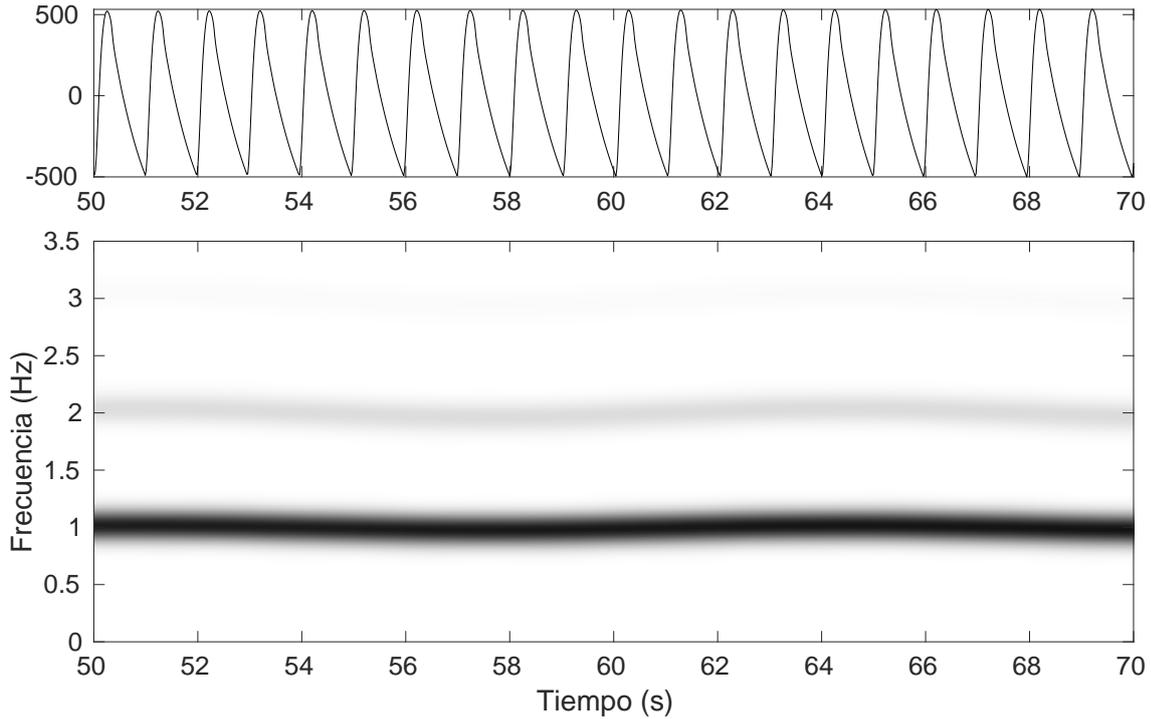


Figura 2.8: Señal de onda de pulso simulada. Primera fila: segmento de 20 segundos de la señal de onda de pulso $x(t)$ dada por (2.19). Esta señal corresponde a uno de los posibles patrones oscilatorios asociados al registro de la señal a nivel de la aorta abdominal. Segunda fila: Espectrograma de $x(t)$, donde se observan las crestas correspondientes al componente fundamental $\phi'(t)$ y los dos primeros armónicos.

frecuencia (0,15 – 0,4 Hz). Para este experimento, muestreamos f_L y f_H de distribuciones uniformes $\mathcal{U}_L(0,01, 0,15)$ y $\mathcal{U}_H(0,15, 0,4)$ respectivamente. Los coeficientes a y b se eligieron de modo que la relación de potencia de baja frecuencia a alta frecuencia (LHPR, del inglés *low-high power ratio*) fuera 3 y la desviación máxima de la frecuencia media ($\bar{f} = 1$ Hz) fuera $\pm 0,035 Hz$, para simular la VFC de un individuo sano tal como se presenta en [83]. Consideramos una duración de la señal de $T = 120$ segundos, muestreada a $f_s = 58,33$ Hz para una longitud total de la señal de 7000 muestras. La STFT se calcula como se ha descrito anteriormente, considerando un ancho para la función de ventaneo elegido de forma adaptativa utilizando un método basado en entropía de Rényi [84, 85]. Los parámetros elegidos para la extracción de crestas y la reconstrucción del modo fundamental fueron los mismos que en la Sec. 2.4.

La metodología paso a paso de este experimento se detalla a continuación:

1. Muestreamos aleatoriamente 70 formas de onda de pulso $s_{\{i:1 \leq i \leq 70\}}$ de cada sección considerada del sistema circulatorio.
2. Para cada forma de onda, generamos una señal AM-FM $x_i(t)$ utilizando (2.19) (mostramos un ejemplo en la Fig. 2.8) y luego añadimos ruido gaussiano de media cero a 3 niveles diferentes de SNR de entrada (SNR_{in}): 0, 10 y 20 dB.
3. Para cada señal ruidosa, estimamos el orden óptimo utilizando el criterio $\Phi_G(r)$ descrito en la Sec. 2.2.2 y aplicamos el algoritmo de reconstrucción de la Sec. 2.3,

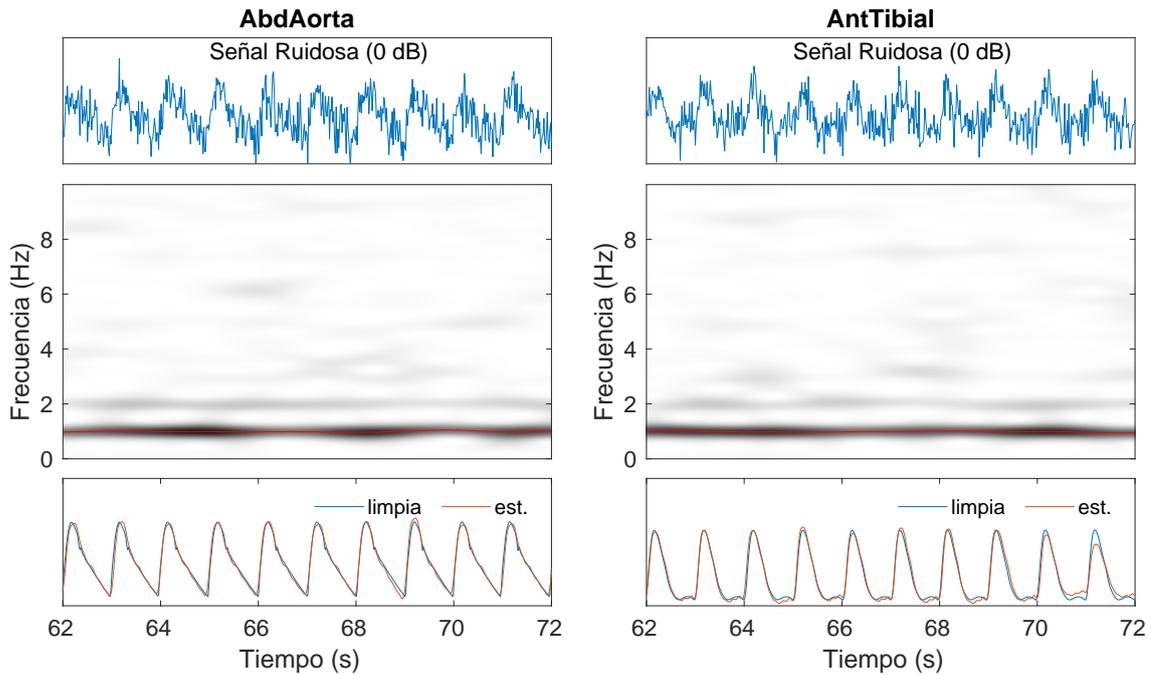


Figura 2.9: Resultados de limpieza de ruido para señales de onda de pulso. Primera fila: Porciones de 10 segundos de las señales de onda de pulso $x(t)$ de los tramos de aorta abdominal y tibial anterior contaminadas con ruido de 0 dB. Segunda fila: Espectrogramas de $x(t)$ para ambos tramos con la curva correspondiente a la cresta fundamental resaltada en rojo. Tercera fila: Señales reconstruidas utilizando el ANHM con estimación adaptativa del orden WSF (rojo) superpuestas a las señales de onda de pulso sin ruido (azul).

utilizando el mínimo del criterio como orden óptimo. Estas señales reconstruidas se denotan como $x_{r,i}(t)$.

4. Se obtuvieron versiones reconstruidas de $x_i(t)$ con un número fijo de armónicos, considerando los siguientes valores: $r = 3, 9, 18$ y r_{max} , para utilizar como resultados de referencia para el método adaptativo.
5. Comparamos nuestra técnica de reconstrucción con la de eliminación de ruido basada en umbralado de la STFT, fijamos un umbral η y reconstruimos la señal por integración numérica sobre la frecuencia de la STFT umbralada $\tilde{\mathbf{F}}$: $\mathbf{x}_{HT,i}(n) = \frac{2}{g(0)} \text{Re}(\sum_{k=1}^K \tilde{\mathbf{F}}(n, k))$. También reconstruimos la señal del mismo modo utilizando un umbral suave.
6. Calculamos SNR_{out} para todas las señales reconstruidas.

2.5.2. Resultados de limpieza de ruido

La Fig. 2.9 muestra, en la fila superior, señales de onda de pulso ruidosas, con ruido a 0 dB de SNR, para los dos tramos del sistema circulatorio que se consideraron. Debajo se muestran los espectrogramas correspondientes a estas señales, con las curvas correspondientes a los componentes fundamentales resaltadas en rojo. Claramente, la energía de la señal se concentra a lo largo de las crestas asociadas a la frecuencia fundamental $\phi'(t)$ y

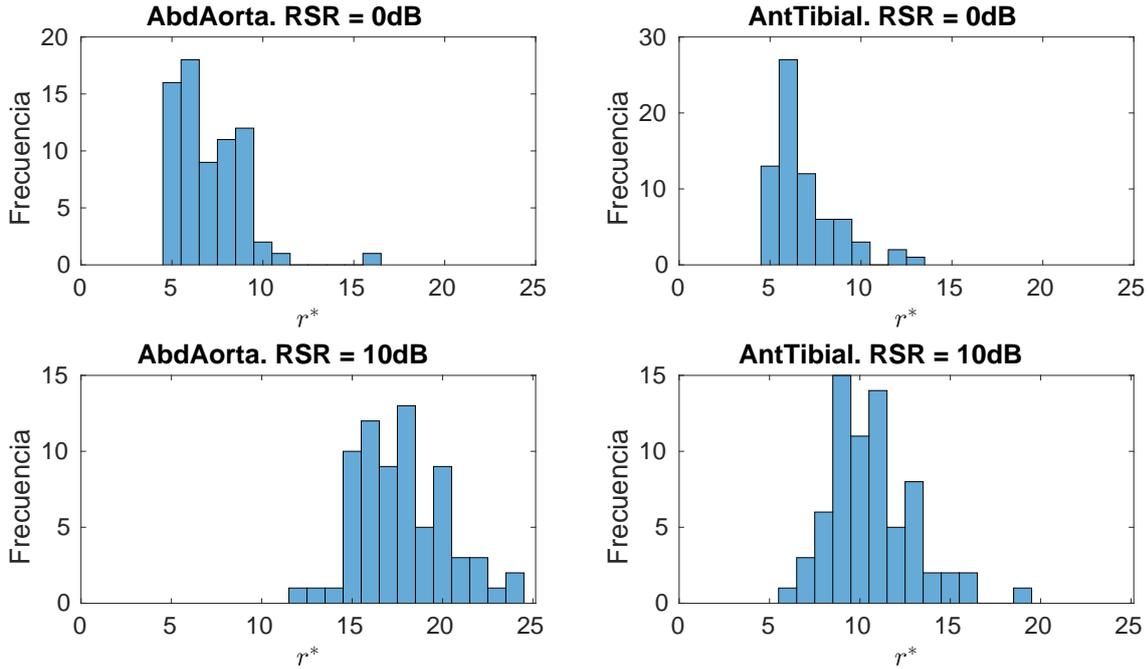


Figura 2.10: Resultados de estimación de orden óptimo en señales de onda de pulso. Columna izquierda: Histograma del orden óptimo del ANHM r^* para todas las formas de onda de la sección de la aorta abdominal para una SNR_{in} de 0 (arriba) y 10 (abajo) dB. Columna derecha: Lo mismo para las formas de onda de la sección tibial anterior.

sus múltiplos, mientras que el ruido tiene una distribución más homogénea a lo largo de todo el plano tiempo-frecuencia. En la parte inferior de la Fig. 2.9 se muestran las señales limpias originales con los resultados de la reconstrucción superpuestos en rojo.

Podemos ver que el procedimiento de reconstrucción basado en el ANHM, con estimación adaptativa del orden del modelo, puede recuperar con precisión la morfología original de la señal de onda de pulso, incluso en presencia de niveles considerables de ruido. Además, el ANHM se ajusta a la variabilidad de la forma de onda del pulso entre las diferentes porciones del sistema vascular.

Es esperable que el orden óptimo del modelo para las señales de una u otra sección difieran entre sí. Efectivamente, la Fig. 2.10 muestra histogramas del orden óptimo estimado con GCV para todas las formas de onda de pulso en cada sección del sistema circulatorio. En primer lugar, vemos que a medida que aumenta la SNR_{in} , la distribución se desplaza hacia valores mayores de r^* . Además, cuando comparamos ambos tramos al mismo nivel de SNR, vemos que las distribuciones también difieren. Esto valida nuestra suposición de que el procedimiento de estimación del orden óptimo del modelo contempla la variabilidad de la forma de onda entre señal de una misma sección, así como la variabilidad entre una sección y la otra.

Además, la Fig. 2.11 muestra los resultados del procedimiento de eliminación de ruido para cada método de reconstrucción, en función de la SNR de entrada. Los boxplots etiquetados como “A” corresponden a la estimación automática del número óptimo de armónicos para la WSF utilizando el criterio $\Phi_G(r)$. Los resultados de reconstrucción obtenidos considerando $r = r_{max}$ se denotan con “M”. Los resultados de la técnica de eliminación de ruido basada en umbralado usando un umbral duro y suave se etiquetan como

2.5. APLICACIÓN 1: LIMPIEZA DE RUIDO EN SEÑALES DE ONDA DE PULSO

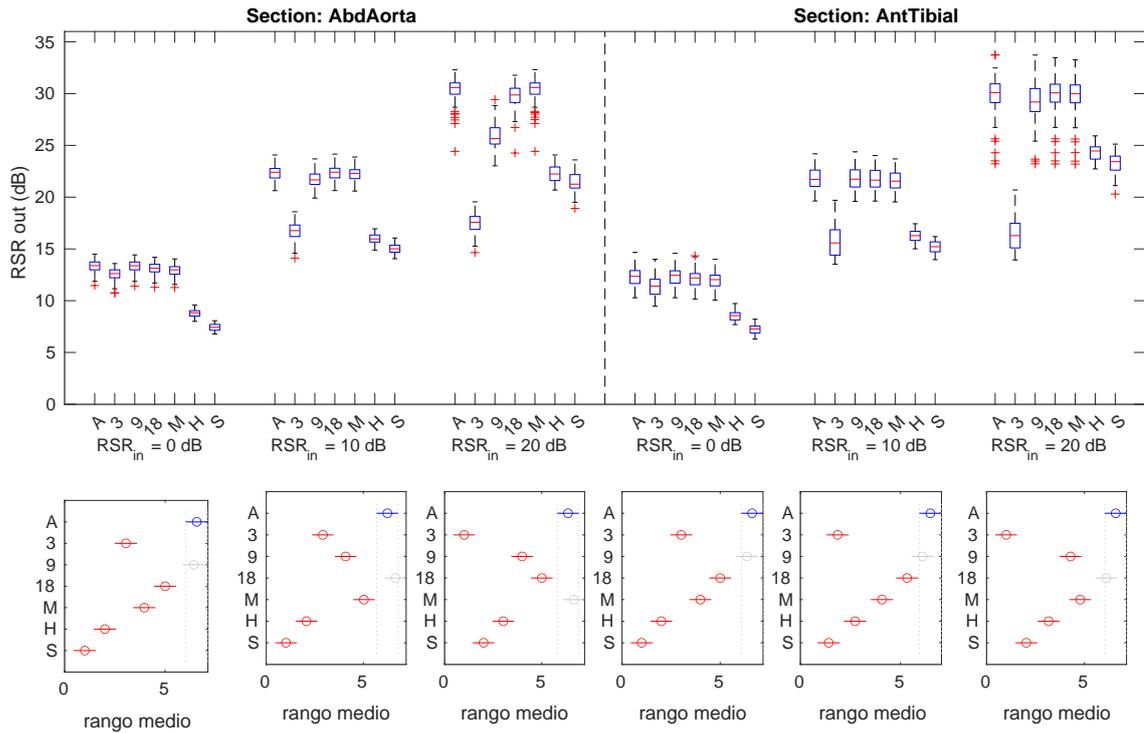


Figura 2.11: Estudio comparativo del algoritmo propuesto en señales de onda de pulso. Primera fila: Diagramas de caja del desempeño del algoritmo para la tarea de eliminación de ruido utilizando los distintos procedimientos de reconstrucción. Segunda fila: Rangos promedios para cada método de reconstrucción obtenidos a partir de la prueba no paramétrica de Friedman. Referencias: A - Orden óptimo del modelo WSF $r = r^*$; 3, 9, 18 - Órdenes fijos del modelo WSF; M - Orden máximo del modelo WSF $r = r_{max}$; H - Umbralado duro; S - Umbralado suave.

“H” y “S”, respectivamente. A partir de los gráficos de caja, podemos ver que el desempeño en la tarea de eliminación de ruido (medido a través de la SNR_{out}) de la mayoría de los métodos aumenta a medida que aumenta la SNR_{in} . Además, las reconstrucciones con un orden de modelo más bajo tienden a empeorar considerablemente para valores más altos de SNR_{in} . La reconstrucción con $r = 3$ tiene un rendimiento considerablemente inferior a partir de 10 dB de ruido de entrada y $r = 9$ tiene un rendimiento inferior para SNR de 20 dB para ambos tramos. En todos los casos, la reconstrucción basada en la estimación adaptativa del orden del modelo se mantiene sistemáticamente como uno de los métodos de mayor rendimiento. Por último, en la mayoría de los casos, el procedimiento de reconstrucción basado en el ANHM supera a las técnicas de umbralado para la mayoría de los valores de r .

2.5.3. Análisis estadístico de los resultados

Para obtener una mejor comparación entre los métodos, realizamos un análisis no paramétrico de la varianza (ANOVA) utilizando la prueba de Friedman [86]. Se eligió un nivel de significancia combinado de $\alpha_H = 0,05$ y se aplicó la corrección de Bonferroni para obtener un nivel de significancia para comparaciones múltiples de $\alpha_B = 0,05/21 \approx$

0,0024. Los rangos medios resultantes y sus intervalos de confianza se muestran en la fila inferior de la Fig. 2.11. A partir de estos gráficos, podemos ver que el método de estimación de orden adaptativo siempre tiene el rango más alto, o está empatado en el rango más alto, para todos los niveles de SNR_{in} . Además, el orden del modelo fijo que empata con el método adaptativo cambia a medida que aumenta el nivel de ruido. Para $SNR_{in} = 0$ dB, el orden fijo $r = 9$ se sitúa por encima de los demás órdenes fijos en ambos tramos. Para $SNR_{in} = 10$ dB, ahora tenemos que $r = 18$ ofrece el mejor rendimiento para la sección aorta abdominal. Para una SNR_{in} de 20 dB, las mejores opciones según los resultados de la prueba son $r = r_{max}$ y $r = 18$ para los tramos AbdAorta y AntTibial, respectivamente, y el método adaptativo es capaz de estimar con precisión estos valores. En resumen, los mejores resultados de limpieza de ruido se alcanzan para diferentes valores de r dependiendo tanto de la SNR_{in} como de la sección del sistema circulatorio analizada. En todos los casos, la estimación adaptativa del orden del modelo recupera el modelo óptimo con respecto al rendimiento de eliminación de ruido.

2.6. Aplicación 2: limpieza de ruido en señales cardíacas y respiratorias

En la sección anterior, mostramos que el ANHM puede utilizarse para eliminar el ruido de señales biomédicas simuladas, e ilustramos el valor de seleccionar el parámetro r de forma adaptativa. En esta sección, nos centraremos en la tarea de modelar y procesar señales fisiológicas reales. Demostraremos que es necesario elegir de forma adaptativa el orden r del ANHM para cada señal con el fin de tener en cuenta la variabilidad de la forma de onda entre pacientes. Analizaremos señales fisiológicas de la base de datos Fantasia Database, concretamente señales de ECG y respiratorias, disponibles a través del repositorio Physionet [87].

Esta base de datos se desarrolló originalmente para el estudio de la dinámica responsable de las fluctuaciones del intervalo entre latidos y su relación con la edad del paciente [88]. Está compuesta por 40 pacientes divididos en dos cohortes: 20 pacientes jóvenes (de 21 a 34 años) y 20 pacientes adultos mayores (de 68 a 85 años). Todos los pacientes se sometieron a un riguroso procedimiento de cribado para descartar posibles problemas de salud. Durante todo el estudio se registraron simultáneamente el ECG y las señales respiratorias con el paciente recostado en decúbito supino durante 120 minutos mientras veía la película *Fantasia* (Disney, 1940) para mantener la vigilia. Las grabaciones contienen una cantidad considerable de ruido y artefactos, entre los que se incluyen el desvío de la línea de base y las interferencias de la línea eléctrica.

En este experimento, se extrajeron segmentos de 200 segundos de longitud de estas señales, discretizadas a una frecuencia de muestreo de 250 Hz. Posteriormente, se aplicó el procedimiento de reconstrucción descrito en la Sec. 2.3 utilizando el criterio de validación cruzada generalizada ($\Phi_G(r)$) para elegir el orden óptimo r^* . Mostramos los espectrogramas de dos señales de la base de datos en la Fig. 2.12 donde las crestas fundamentales están resaltadas en rojo. Aquí aclaramos que, a diferencia de las señales estudiadas hasta el momento, la cresta fundamental no es la dominante para el caso de las señales de ECG. Debido a esto, para la extracción de la cresta de las señales de ECG, se utilizó la *de-shape STFT* [89] con el fin de obtener el componente fundamental de la señal. De las representaciones T-F se desprende claramente que la presencia de ruido

2.6. APLICACIÓN 2: LIMPIEZA DE RUIDO EN SEÑALES CARDÍACAS Y RESPIRATORIAS

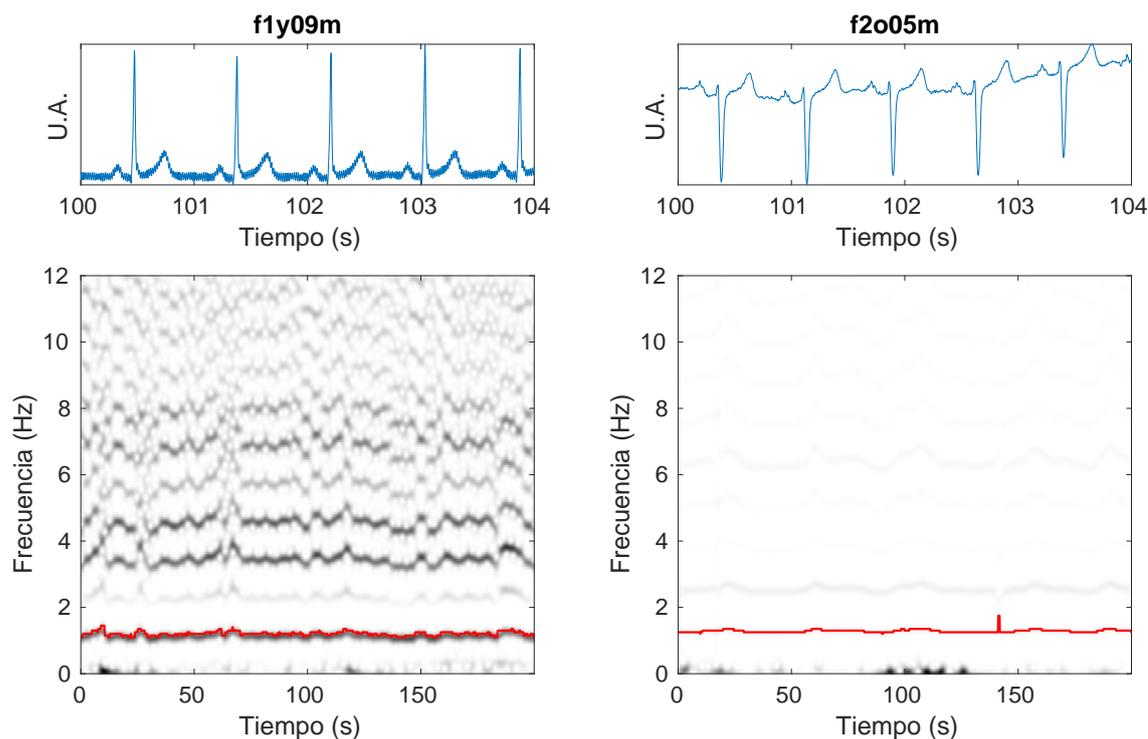


Figura 2.12: Espectrogramas de las señales de electrocardiografía de la base de datos Fantasia. Izquierda: espectrograma para la señal de ECG de un paciente de la cohorte de edad avanzada. Derecha: espectrograma para la señal de ECG de un paciente de la cohorte joven. En ambos casos las crestas de la frecuencia fundamental resaltadas en rojo.

y la desviación de la línea de base dificultan la estimación del número de componentes armónicos relevantes mediante la extracción individual de crestas. Además, estas representaciones también muestran que el número óptimo de armónicos que deben tenerse en cuenta para una estimación precisa de la WSF varía enormemente de un paciente a otro. Para cada paciente de cada cohorte, se obtuvieron las WSFs y se aplicó la fórmula de reconstrucción a las señales de ECG y respiratorias utilizando el Algoritmo 1 con estimación automática del orden r de las WSF.

En la Fig. 2.13 mostramos los resultados para un paciente de cada grupo de edad. Junto a las formas de onda estimadas para el ECG y las señales respiratorias, mostramos las señales originales en azul con la señal del ANHM superpuesta en negro. Vemos que la reconstrucción con el ANHM recupera la morfología general tanto del ECG como de las señales respiratorias y filtra el ruido procedente de la interferencia de la línea eléctrica (f1y09m) y de la desviación de la línea de base (f2o05m). Además, cabe señalar que el orden óptimo para cada señal es diferente y que, en general, se necesitan más armónicos para representar con precisión el ECG, ya que se trata de un fenómeno oscilatorio más complejo en comparación con la señal respiratoria. Para cada WSF, indicamos el orden óptimo estimado r^* , y el orden máximo posible r_{max} en función de la relación entre la frecuencia fundamental y la frecuencia de muestreo. Se realizó el mismo análisis para todos los pacientes de la base de datos y los resultados completos se presentan en las Figs. 2.14-2.17.

También incluimos la WSF estimada para cada señal en el intervalo $(-\pi, \pi)$. Vemos

2.6. APLICACIÓN 2: LIMPIEZA DE RUIDO EN SEÑALES CARDÍACAS Y RESPIRATORIAS

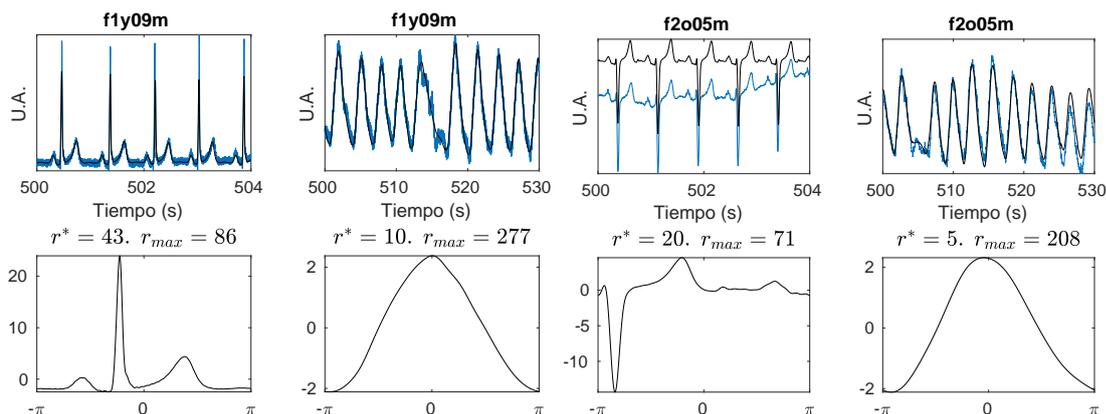


Figura 2.13: Resultados de reconstrucción de funciones de forma de onda para la base de datos Fantasia. Izquierda: Reconstrucciones mediante el ANHM (negro) superpuesta a las señales originales (azul) para un paciente de la cohorte joven. Se muestra también la función de forma de onda obtenida para las señales de ECG y respiratoria, respectivamente. Derecha: Lo mismo para las señales de un paciente de la cohorte de edad avanzada.

que el procedimiento de reconstrucción basado en el ANHM captura con precisión el patrón oscilatorio tanto para el ECG como para la respiración, al tiempo que elimina las interferencias de la línea eléctrica y el desvío de la línea de base presentes en las señales. Uno de los aspectos negativos de la reconstrucción es que el pico R en el ECG se subestima en la mayoría de los casos, aunque el grado de atenuación del pico R no es consistente para todas las señales. En algunas, la amplitud de la onda T supera a la del pico R.

Como podemos ver en las formas de onda de las Figs. 2.14-2.17, el número óptimo de términos trigonométricos obtenidos en base al criterio de selección son diferentes para la mayoría de los pacientes y notablemente bajos cuando se comparan con el orden máximo admisible r_{max} . Las señales respiratorias tienen una forma de onda más simple, por lo que se favorece un orden de modelo más bajo. En cambio, la morfología del ECG es más compleja y varía mucho de un paciente a otro, por lo que se requiere un modelo de orden superior. En particular, en el caso de las señales respiratorias, la reconstrucción del ANHM es capaz de preservar los cambios rápidos tanto en la frecuencia instantánea como en la amplitud, debido a la naturaleza del fenómeno respiratorio fisiológico y a las condiciones de adquisición de la señal. No obstante, el procedimiento de reconstrucción falla en algunos casos. Para el paciente $f2o04m$ (tercera y cuarta fila de la Fig. 2.16), el ANHM es incapaz de recuperar el complejo QRS del ECG y la amplitud de las ondas P y T se reduce considerablemente. En el caso del paciente $f2o09m$ (tercer y cuarta fila de la Fig. 2.16), el procedimiento de reconstrucción tampoco capta el complejo QRS y el orden óptimo de la WSF es llamativamente bajo (10 componentes armónicos).

La razón de este desempeño pobre para algunas señales de ECG se debe a que el ANHM es adecuado para representar señales *oscilatorias*, es decir, que se pueden ver como superposiciones de oscilaciones perturbadas en tiempo o frecuencia. El ECG, en cambio, presenta un comportamiento impulsivo debido a la presencia del complejo QRS. Esto provoca que el modelo no sea capaz de capturar adecuadamente la morfología del complejo QRS incluso al considerar un número elevado de armónicos. No obstante, los resultados obtenidos indican que la estimación automática del orden utilizando criterios

de selección de modelos trigonométricos es útil para el modelado de señales fisiológicas reales que tienen propiedades variables en el tiempo (fase y amplitud) en presencia de ruido.

2.7. Comentarios de fin de capítulo

En este capítulo, se discutió la tarea de determinar, de manera guiada por los datos, el número de armónicos necesarios para reconstruir la función de forma de onda de una señal no estacionaria a partir del modelo no armónico adaptativo. En este contexto, se decidió utilizar para esta tarea una serie de criterios de selección de modelos propuestos originalmente en el ámbito de la regresión trigonométrica. La incorporación de los criterios de selección de modelos nos permitió definir un algoritmo adaptativo capaz de reconstruir señales con formas de onda no sinusoidales y bajo condiciones de no estacionariedad.

Se evaluó el rendimiento de diferentes criterios de estimación del orden para el caso de señales no estacionarias, y todos los criterios proporcionaron una buena estimación del orden del modelo en la mayoría de las condiciones. Cabe destacar que estos no son los únicos criterios de selección posibles, pudiendo definirse criterios que tengan en cuenta otros componentes del modelo, como por ejemplo la tendencia variable en el tiempo o las características aleatorias del ruido que contamina la señal.

Respecto al desempeño del algoritmo propuesto, los resultados experimentales en señales fisiológicas simuladas y reales muestran que el ANHM es una herramienta adecuada para analizar señales contaminadas con niveles considerables de ruido. Además, se observó que el orden más adecuado para realizar la reconstrucción de la señal mediante el ANHM depende no sólo de la señal sino también de las condiciones de ruido presentes en los registros. Esto demuestra la importancia del uso de un criterio de selección automática del número de armónicos necesarios para estimar la WSF.

El método de selección de orden propuesto junto a los resultados presentados en esta sección dieron como resultado una publicación en congreso nacional [1] y una publicación en la revista *Signal Processing* de la editorial Elsevier [2].

2.7. COMENTARIOS DE FIN DE CAPÍTULO

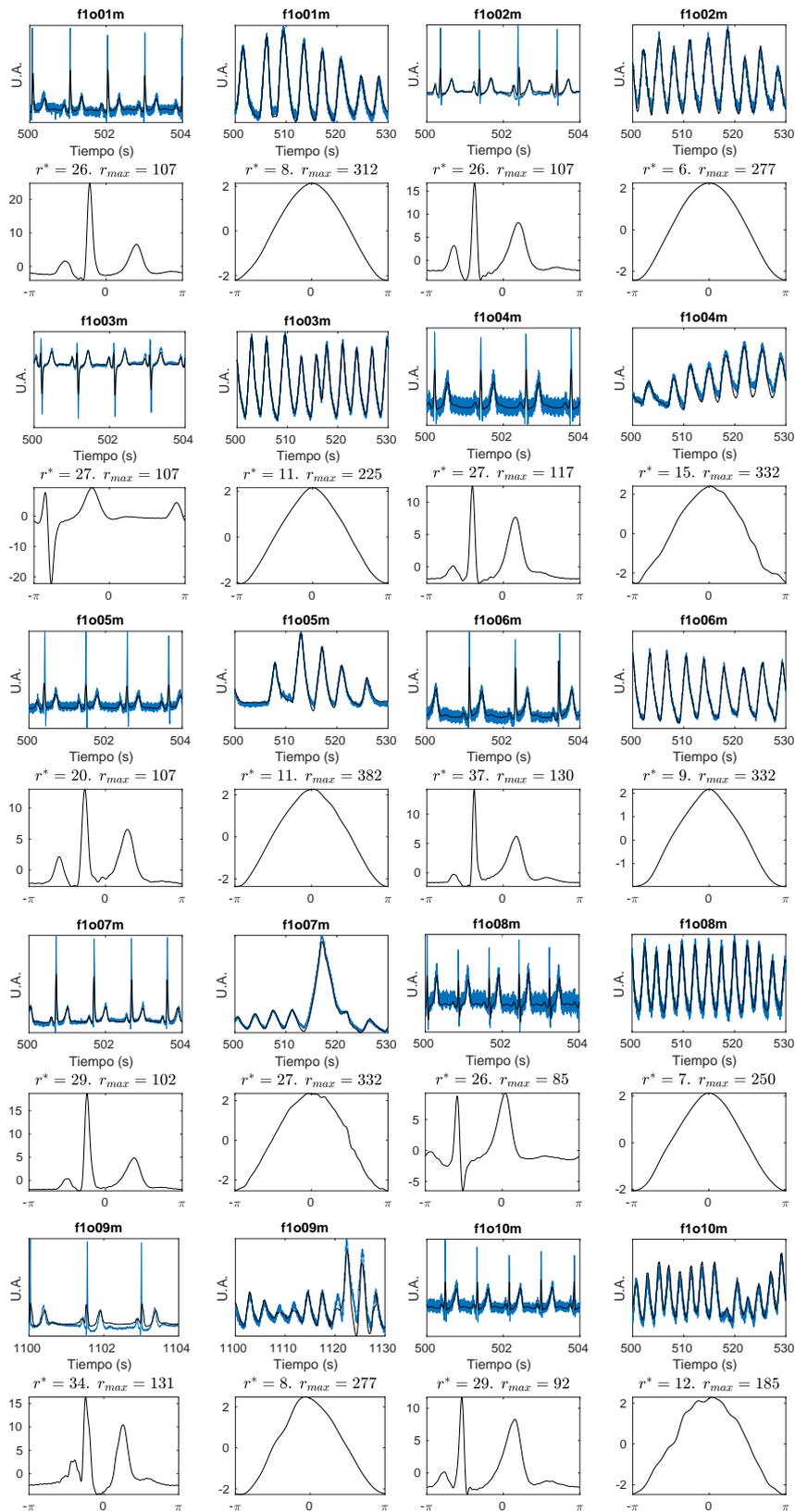


Figura 2.14: Resultados de reconstrucción de funciones de forma de onda para los 10 primeros sujetos de la cohorte de edad avanzada. Las señales respiratorias y de ECG originales se representan en azul con la señal reconstruida superpuesta. También se muestran las WSFs de cada señal.

2.7. COMENTARIOS DE FIN DE CAPÍTULO

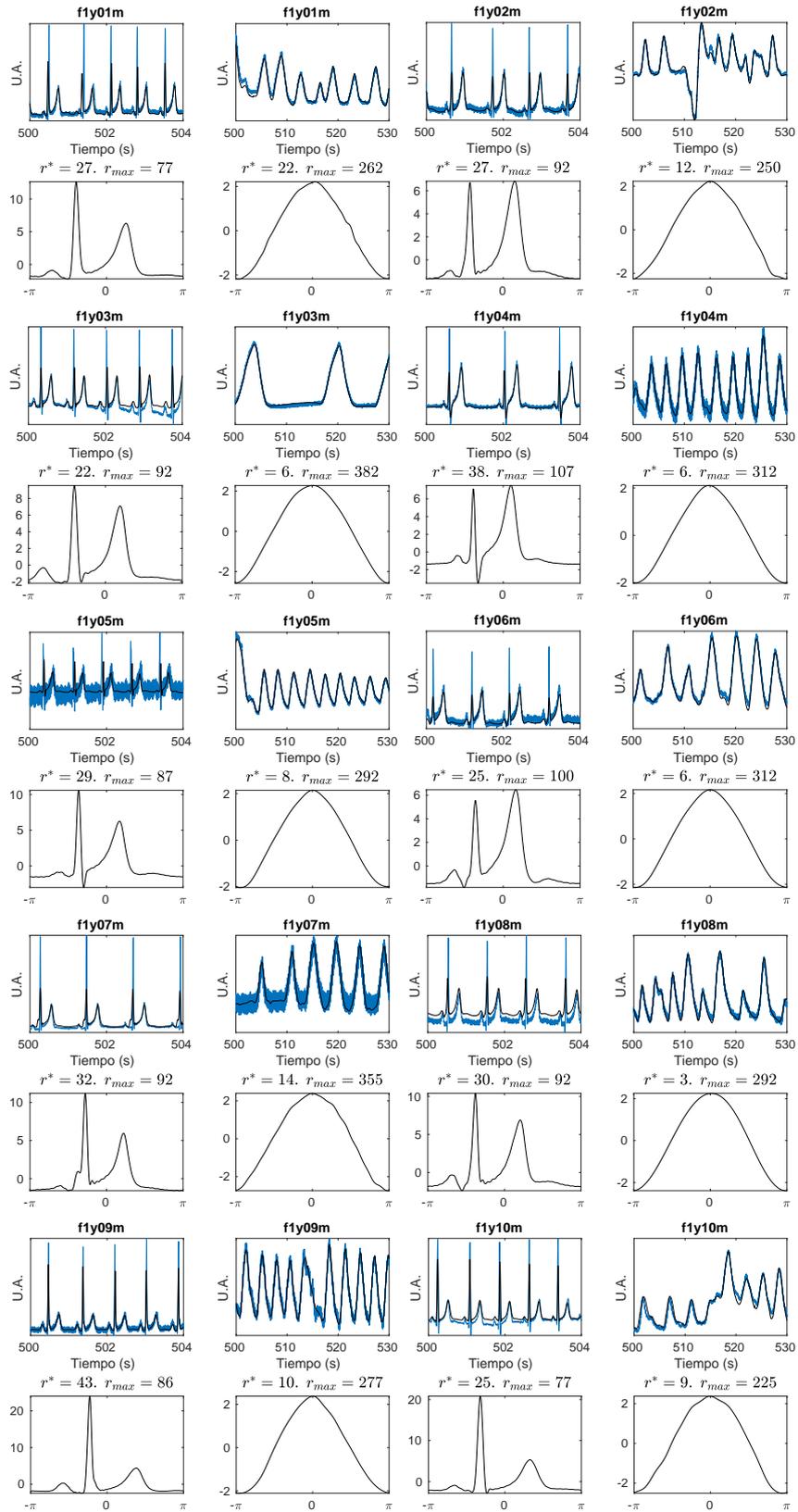


Figura 2.15: Resultados de reconstrucción de funciones de forma de onda para los 10 primeros sujetos de la cohorte de jóvenes. Las señales respiratorias y de ECG originales se representan en azul con la señal reconstruida superpuesta. También se muestran las WSFs de cada señal.

2.7. COMENTARIOS DE FIN DE CAPÍTULO

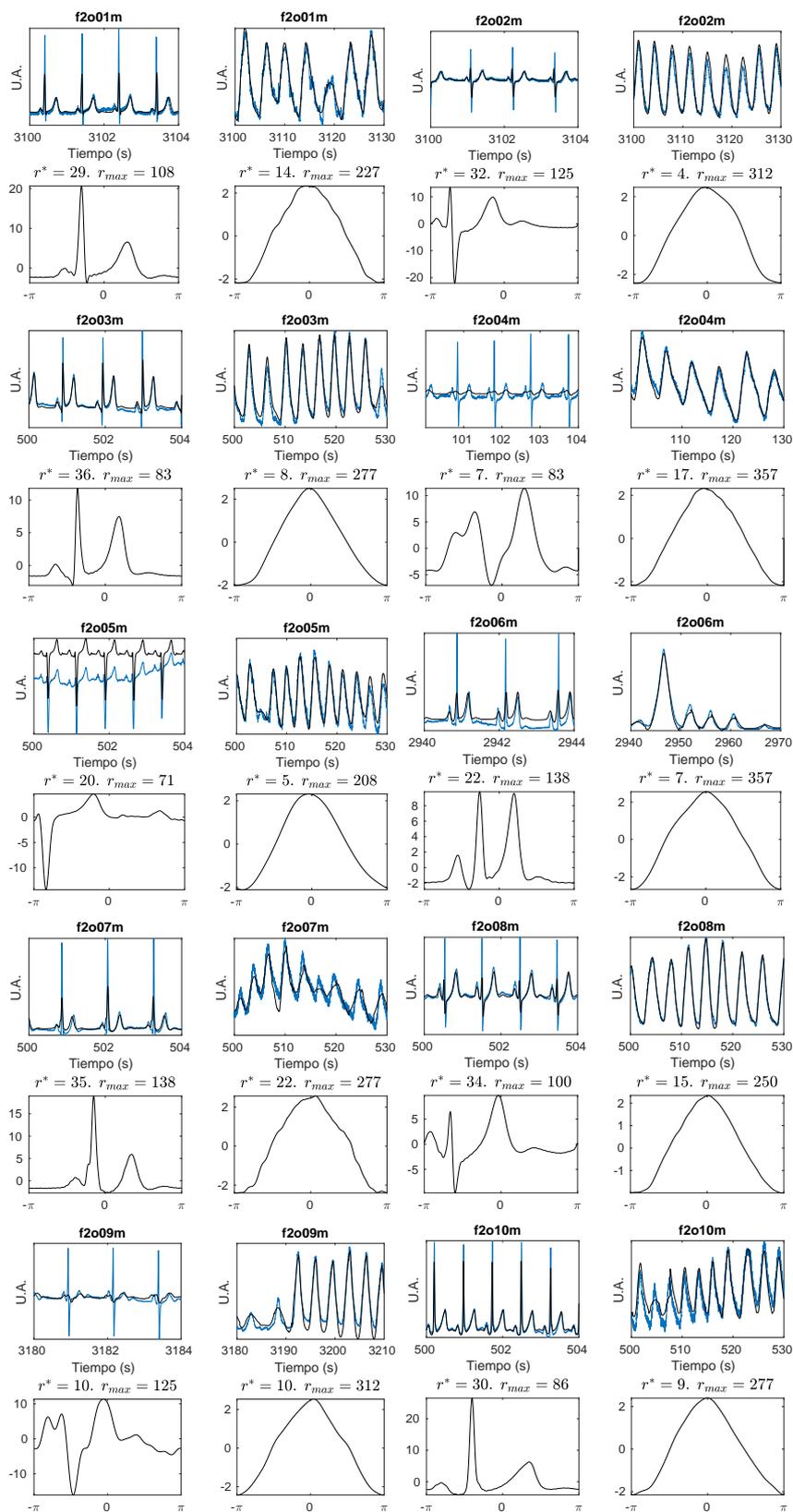


Figura 2.16: Resultados de reconstrucción de funciones de forma de onda para los 10 últimos sujetos de la cohorte de edad avanzada. Las señales respiratorias y de ECG originales se representan en azul con la señal reconstruida superpuesta. También se muestran las WSFs de cada señal.

2.7. COMENTARIOS DE FIN DE CAPÍTULO

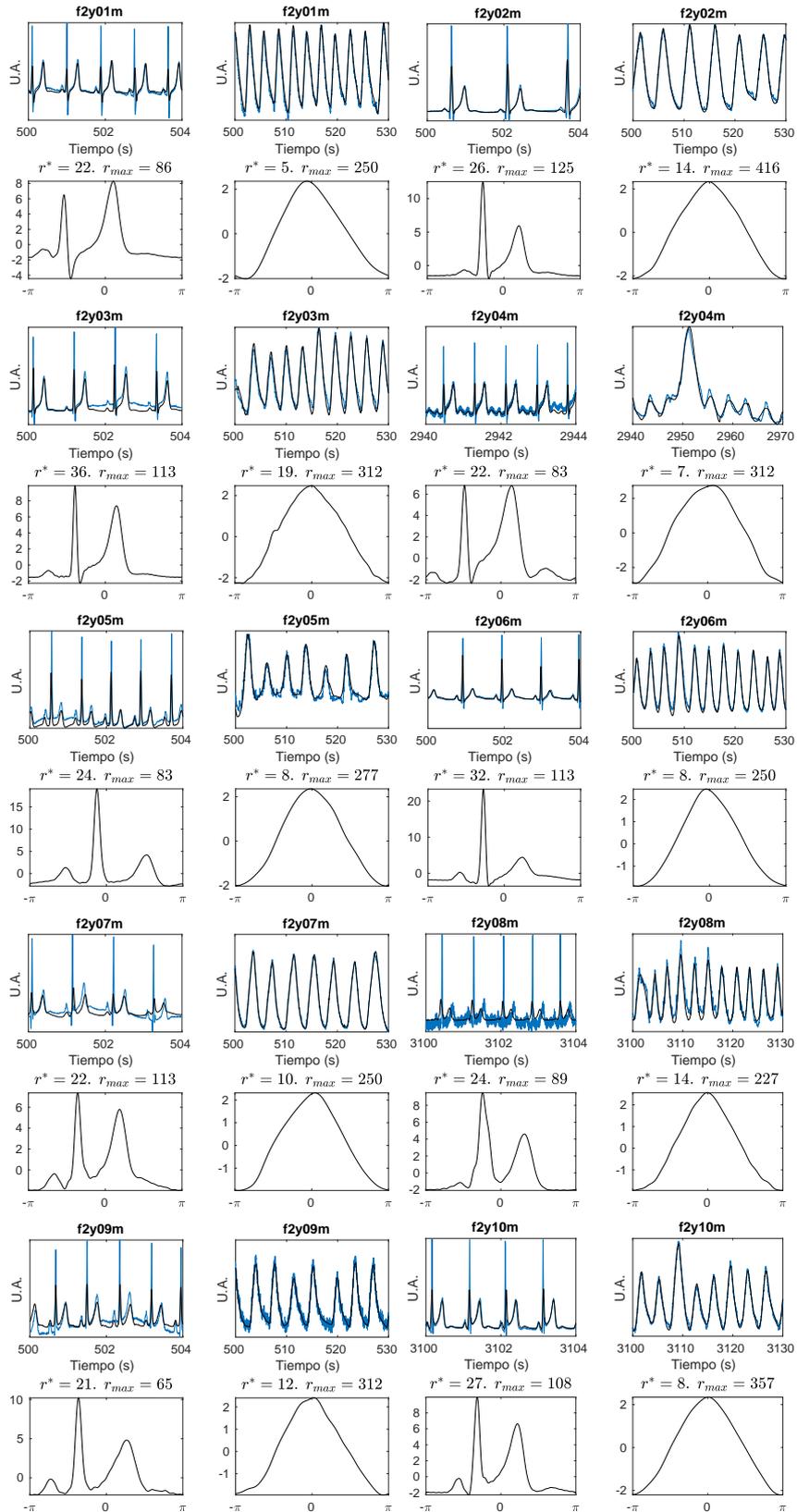


Figura 2.17: Resultados de reconstrucción de funciones de forma de onda para los 10 últimos sujetos de la cohorte de jóvenes. Las señales respiratorias y de ECG originales se representan en azul con la señal reconstruida superpuesta. También se muestran las WSFs de cada señal.

2.7. COMENTARIOS DE FIN DE CAPÍTULO

Capítulo 3

Modelo no armónico adaptativo con funciones de forma de onda variantes en el tiempo

El modelo no armónico adaptativo descrito hasta ahora considera una forma de onda *fija*, es decir, que no varía a lo largo del tiempo. De esta manera, la variabilidad ciclo a ciclo de la señal queda totalmente explicada a través de la modulación en amplitud y la modulación en fase. Como ya discutimos, la primera modifica la oscilación a partir de la amplificación o atenuación ciclo a ciclo de la onda, mientras que la segunda afecta la escala local de la oscilación mediante una contracción o dilatación ciclo a ciclo. Esta condición de forma de onda fija permite caracterizar la función forma de onda mediante los coeficientes de su descomposición en serie de Fourier. Además, en la Sec. 1.4.3 definimos una familia de funciones de forma de onda que cumplen con ciertas características espectrales. Sin embargo, comúnmente las señales reales no presentan un patrón oscilatorio fijo, sino que varía a lo largo del tiempo y dicha variación no puede ser explicada únicamente a través de la modulación en amplitud o en fase.

Para ilustrar esta condición, la Fig. 3.1 muestra algunos ejemplos de señales no estacionarias del mundo real, en particular señales biomédicas. La primera señal es un EEG de un recién nacido que se encuentra experimentando un ataque epiléptico. El registro está contaminado por ruido de alta frecuencia asociado a la contracción muscular. Además, vemos que el patrón oscilatorio tiende a cambiar con el tiempo al igual que la duración de cada ciclo. La segunda señal es un registro de neumografía por impedanciometría (IP), la cual está compuesta por un componente respiratorio de baja frecuencia y un componente cardíaco de alta frecuencia, teniendo ambas señales patrones variantes en el tiempo. Por último, se muestra un ECG de un paciente taquiarrítmico al inicio de un evento de fibrilación ventricular, donde se ve un cambio abrupto de forma de onda en los últimos segundos del registro. En general, esta variabilidad en la forma de las oscilaciones es debido a la dinámica subyacente de los sistemas que las generan, y recordamos aquí que esta variación temporal del patrón oscilatorio es un aspecto importante que debemos tener en cuenta al estudiar estos sistemas.

Los enfoques existentes para el análisis de formas de onda variables en el tiempo se basan en gran medida en la detección de puntos de referencia en el dominio temporal o en el análisis de la variabilidad espectral en el dominio de la frecuencia. Estas técnicas dependen de la señal analizada y pueden no ser adecuadas para casos patológicos. En

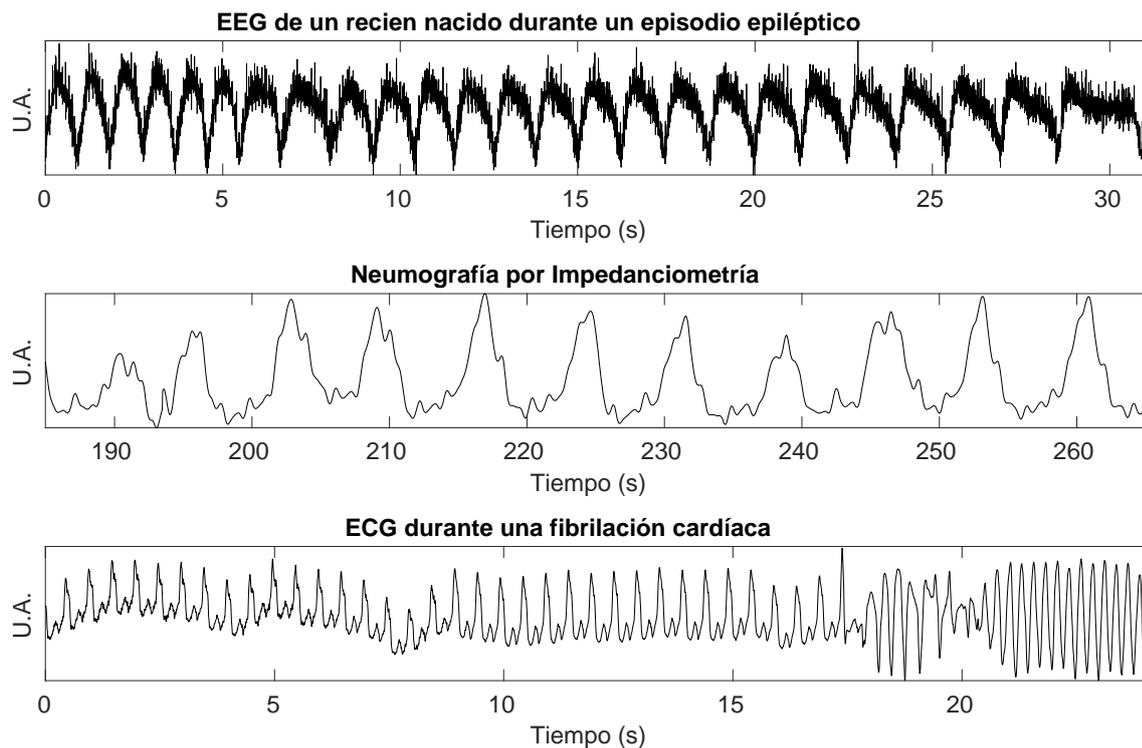


Figura 3.1: Ejemplos de señales biomédicas no estacionarias con forma de onda variable en el tiempo. Primera fila: señal de EEG de un recién nacido durante un episodio epiléptico. Segunda fila: señal de neumografía por impedanciometría compuesta por un componente respiratorio y otro cardíaco. Tercera fila: señal de ECG de un paciente al inicio de una fibrilación ventricular.

el estado de arte, se evidencia la ausencia de un marco general que permita el análisis de señales no estacionarias con forma de onda variable en el tiempo. Para abordar esta cuestión, Lin y cols. [89] propusieron una variante del ANHM que considera una forma de onda variable en el tiempo, definida a través de variaciones temporales en la amplitud y la frecuencia de los componentes armónicos de la forma de onda. Este modelo se postula como un buen candidato para analizar señales biomédicas con formas de onda variables. Sin embargo, su aplicabilidad sigue siendo limitada, principalmente porque no está claro cómo caracterizar la variabilidad de las amplitudes y frecuencias armónicas.

Aunque se han propuesto algunos métodos para analizar este tipo de señales, los mismos no llegan a capturar todas las fuentes de variabilidad de la forma de onda y, en algunos casos, no permiten llevar adelante el análisis de la forma de onda y su variabilidad en conjunto con tareas clásicas de procesamiento como son la limpieza de ruido o la descomposición de señales multicomponente. Otro aspecto a tener en cuenta es que al considerar modelos más complejos, la cantidad de parámetros a determinar aumenta considerablemente y muchas veces dichos parámetros deben ser establecidos por el operador a partir de información conocida sobre la señal bajo estudio. Surge entonces la necesidad de desarrollar una herramienta que permita, de forma guiada por los datos, recuperar la forma de onda variante en el tiempo a partir de señales no estacionarias, pudiendo estar contaminadas con diferentes fuentes de ruido, o bien mezcladas con otras señales de características similares.

En este capítulo, comenzaremos describiendo el modelo propuesto por Lin y cols. en la Sec. 3.1, el cual considera funciones de forma de onda variantes en el tiempo. En la Sec. 3.2, mencionaremos algunas técnicas existentes para evaluar dichas señales con WSFs variables en el tiempo. En la Sec. 3.3, describiremos nuestro algoritmo propuesto para la estimación adaptativa de la forma de onda variable en el tiempo de señales oscilatorias mediante una variante del ANHM y la aplicación de un método de ajuste de curvas para obtener una representación paramétrica de la WSF variante en el tiempo. A diferencia de los enfoques existentes, nuestra propuesta tiene en cuenta tanto la proporcionalidad no entera de las funciones de fase “armónicas” como también la variación temporal de las funciones de amplitud armónica. Este procedimiento de extracción de la forma de onda permite caracterizar la variabilidad de la forma de onda en señales AM-FM utilizando un vector de coeficientes de bajo orden. En la Sec. 3.4, definiremos una nueva familia de funciones de amplitud armónica, las cuales utilizaremos (junto con múltiplos no enteros de la fase) para caracterizar la forma de onda variante en el tiempo. En la Sec. 3.5, detallaremos el procedimiento adaptativo para obtener los parámetros de la función de forma de onda a partir de los datos, que incluye la selección automática de todos los hiperparámetros del algoritmo. Validaremos este algoritmo sobre señales sintéticas en la Sec. 3.6. Finalmente, en la Sec. 3.7 evaluaremos nuestro algoritmo en diferentes aplicaciones comunes en el campo del procesamiento de señales biomédicas.

3.1. Funciones de forma de onda variables en el tiempo

Para describir el comportamiento del modelo con forma de onda variante en el tiempo, consideremos primero el caso *monocomponente* (es decir, $K = 1$) del ANHM con forma de onda fija. Como ya vimos en el Cap. 1, si s pertenece a la clase $\mathcal{S}^{\delta, D, \theta}$, entonces podemos sustituirlo por su expansión en serie de Fourier

$$x_{\text{FIX}}(t) = A(t)s(2\pi\phi(t)) = A(t) \sum_{\ell=1}^{\infty} a_{\ell} \cos(2\pi\ell\phi(t) + \varphi_{\ell}). \quad (3.1)$$

Aquí, la morfología de la función de forma de onda s se caracteriza por los coeficientes de Fourier a_{ℓ} y los corrimientos en fase φ_{ℓ} . En [89], se presenta una generalización de (3.1), que admite una función de forma de onda variable en el tiempo:

$$x_{\text{VAR}}(t) = \sum_{\ell=1}^{\infty} B_{\ell}(t) \cos(2\pi\phi_{\ell}(t)). \quad (3.2)$$

Este modelo considera una señal no estacionaria con función de forma de onda variable en el tiempo como la superposición de componentes “armónicos” $B_{\ell}(t) \cos(2\pi\phi_{\ell}(t))$ que satisfacen las siguientes condiciones:

- (C3) $\phi_{\ell} \in C^2(\mathbb{R})$ con $|\phi'_{\ell}(t) - \ell\phi'_1(t)| \leq \epsilon_3\phi'_1(t)$, para todo t y para $\ell = 1, 2, \dots, \infty$.
- (C4) $B_{\ell} \in C^1(\mathbb{R}) \cap L^{\infty}(\mathbb{R})$, con $B_{\ell}(t) > 0$ y $B_{\ell}(t) \leq c(\ell)B_1(t)$, siendo $\{c(\ell)\}_{\ell=1}^{\infty}$ una secuencia ℓ^1 no negativa, para todo t y para $\ell = 1, 2, \dots, \infty$. Además, existe $N \in \mathbb{N}$ de modo que $\sum_{\ell=N+1}^{\infty} B_{\ell}(t) \leq \epsilon \sqrt{\sum_{\ell=1}^{\infty} B_{\ell}^2(t)}$ y $\sum_{\ell=N+1}^{\infty} \ell B_{\ell}(t) \leq C \sqrt{\sum_{\ell=1}^{\infty} B_{\ell}^2(t)}$.

(C5) $|B'_\ell(t)| \leq \epsilon c(\ell)\phi'_1(t)$ y $|\phi''_\ell(t)| \leq \epsilon \ell \phi'_1(t)$ para todo t y $\sup_{\ell; B_\ell \neq 0} \|\phi''_\ell\|_\infty = M$ para algún $M \geq 0$.

La condición (C3) indica que la modulación en frecuencia $\phi'_\ell(t)$ de cada componente es próxima, pero no exactamente igual, a un múltiplo entero de la frecuencia fundamental $\phi'_1(t)$. La condición (C4) da cuenta de la variabilidad de la forma de onda dada por la variación de la amplitud para cada armónico cuando la relación $B_\ell(t)/B_1(t)$ no es constante. La relación entre las amplitudes armónicas queda descrita por la secuencia de valores $c(\ell)$ para cada ℓ . Finalmente, la condición (C5) establece que la forma de onda varía *lentamente*, es decir, no hay cambios bruscos en el patrón oscilatorio de la señal. Obsérvese que podemos recuperar la expansión de Fourier de s mostrada en (3.1) cuando $\phi_\ell(t) = \ell\phi(t) + \varphi_\ell/(2\pi)$ y $B_\ell(t) = A(t)a_\ell$ para $\ell \geq 1$.

3.2. Métodos de análisis de señales con funciones de forma de onda variantes en el tiempo

Los modelos existentes para el estudio de señales no estacionarias, como son los modelos AH y ANH, resultan insuficientes para estudiar señales reales cuya forma de onda varía a lo largo del tiempo. Los modelos anteriores sólo pueden extraer un patrón oscilatorio “promedio”, no pudiendo representar la información asociada a la variabilidad de la forma de onda. En este contexto, diversos modelos alternativos han surgido con el fin de poder capturar la mayor cantidad de información posible asociada a la variación temporal de las oscilaciones. En las siguientes secciones describiremos algunas de las propuestas más relevantes.

3.2.1. *Multiresolution Mode Decomposition* (MMD)

Este modelo, propuesto en [60], considera a una señal multicomponente como una superposición de modos de la forma

$$f_k(t) = \sum_{n=-N/2}^{N/2-1} a_{k,n} \cos(2\pi n\phi_k(t)) s_{k,cn}(2\pi N\phi_k(t)) + \sum_{n=-N/2}^{N/2-1} b_{k,n} \sin(2\pi n\phi_k(t)) s_{k,sn}(2\pi N\phi_k(t)), \quad (3.3)$$

donde $f_k(t)$ se denomina como *multiresolution intrinsic mode function* (MIMF) y el conjunto de funciones $\{s_{cn}(\cdot), s_{sn}(\cdot)\}_{n=-N/2}^{N/2-1}$ se denomina *conjunto de funciones de forma*. Las sumatorias en (3.3) determinan el comportamiento variante en el tiempo de la forma de onda de la señal. Este modelo considera que los argumentos de las funciones seno, coseno, $s_{k,cn}$ y $s_{k,sn}$ pueden estimarse a partir de los datos y no establece ninguna condición adicional. La implementación de este modelo se enfoca exclusivamente en estimar las amplitudes y funciones de forma variantes en el tiempo.

Para estimar las funciones de forma $s_{k,cn}$ y $s_{k,sn}$, los autores proponen la utilización de un método de regresión recursivo basado en difeomorfismos (RDBR, por las siglas

en inglés de *recursive diffeomorphism-based regression*). Dicho algoritmo considera que, para cada modo, la función de fase $N_k\phi_k(t) : \mathbb{R} \rightarrow \mathbb{R}$ es un mapeo invertible de tal manera que tanto $N_k\phi_k(t)$ como $\frac{1}{N_k}\phi_k^{-1}(t)$ son funciones continuas y diferenciables. Gracias a esto, es posible definir una transformación que permite reescalar localmente la función $f_k(t)$, con fase no lineal $N_k\phi_k(t)$, en una nueva función $\tilde{f}_k(t)$ con fase lineal $N_k t$. Finalmente, la función de forma de onda s_{cn} se obtiene de resolver el problema RDBR sobre la función

$$s_{k,cn} \approx h_{k,cn} = f_k(\phi_k^{-1}(t)) \cos(2\pi n\phi_k(t)). \quad (3.4)$$

Análogamente, para obtener la función de forma de onda s_{sn} se aplica el algoritmo RDBR sobre la función

$$s_{k,sn} \approx h_{k,sn} = f_k(\phi_k^{-1}(t)) \sin(2\pi n\phi_k(t)). \quad (3.5)$$

Al igual que en el caso del ANHM, es necesario limitar la cantidad de términos de la sumatoria por lo cual la estimación de las funciones de forma de onda suele limitarse a una aproximación de ancho de banda limitada considerando solamente $2M + 1$ componentes de cada sumatoria, por lo cual las funciones $\{s_{sn}\}$ y $\{s_{cn}\}$ se estiman para $-M \leq n \leq M$.

3.2.2. *Shape Adaptive Mode Decomposition (SAMD)*

Con el objetivo de poder aplicar en la práctica el modelo (3.2), los autores de [59] propusieron utilizar la siguiente versión simplificada

$$x_{\text{SAMD}}(t) = A(t) \sum_{\ell=1}^{\infty} a_{\ell} \cos(2\pi\phi_{\ell}(t)), \quad (3.6)$$

satisfaciendo las condiciones (C1), (C2), (C3) y (C5). Podemos ver que el modelo (3.6) es equivalente a (3.2) cuando $B_{\ell}(t) = A(t)a_{\ell}$, es decir, cuando la amplitud modulada de el ℓ -ésimo componente armónico es múltiplo de la amplitud instantánea $A(t)$. Sin pérdida de generalidad, podemos suponer que $B_1(t) = A(t)$ y, en consecuencia, $a_1 = 1$. Por otro lado, la relación entre la fase instantánea de los armónicos y el componente fundamental se modela como un polinomio:

$$\phi_{\ell}(t) = \sum_i^I e_{i,\ell} \phi_1(t)^i, \text{ con } e_{i,\ell} \in \mathbb{R} \quad (3.7)$$

siendo $e_{1,1} = 1$ y $e_{i,1} = 0$ para $i > 1$, ya que la definición de las fases se basan en una estimación de la fase del primer armónico $\hat{\phi}_1(t)$. Esta representación para las fases $\phi_{\ell}(t)$ permite relajar levemente la condición de proporcionalidad entera entre la fase fundamental y la fase de los armónicos. Generalmente, los polinomios de primer orden son suficientes para dar una buena aproximación de la WSF variable en el tiempo, por lo que la ecuación se puede reducir a $\phi_{\ell}(t) = e_{\ell}\phi_1(t)$.

Para determinar los parámetros de este modelo, los autores proponen resolver un problema de regresión no lineal de la forma

$$\min_{\substack{a_\ell, b_\ell \\ e_\ell}} \left\| x(t) - A(t) \sum_{\ell=1}^D a_\ell \cos(2\pi e_\ell \phi_1(t)) + b_\ell \sin(2\pi e_\ell \phi_1(t)) \right\|_2^2, \quad (3.8)$$

Este modelo demostró ser superior a MMD tanto al aplicarse sobre señales sintéticas como señales biomédicas. Sin embargo, un aspecto que este modelo no logra capturar es la variabilidad de la forma de onda debido a los cambios en amplitud de los armónicos a lo largo del tiempo. Dicho comportamiento se contempla en la condición (C4) del ANHM con forma de onda variante en el tiempo, la cual no es tenida en cuenta por SAMD.

3.2.3. Modelo Oscilatorio de Funciones de Forma de Onda

Esta propuesta difiere de las anteriores en el sentido de que no busca encontrar una representación de los datos en el dominio temporal sino que realiza una segmentación de los ciclos de la señal y los representa como vectores de un espacio de alta dimensión. Para comenzar, se define la señal no estacionaria mediante un modelo aleatorio

$$Y(t) = \sum_{j \in \mathbb{Z}} \delta(t - t_j) \star s_j(t) + n(t), \quad (3.9)$$

donde t_j es el instante asociado a la j -ésima oscilación $s_j(t)$, $\delta(t - t_j)$ es un delta de Dirac centrado en t_j y el símbolo \star indicado la operación de convolución. Este modelo indica que la señal estará compuesta por una sucesión de ciclos u oscilaciones, donde el patrón oscilatorio del j -ésimo ciclo esta dado por s_j y se centra en el instante de tiempo t_j . Las oscilaciones s_j se distribuyen en una *manifold de formas de onda* de alta dimensión $\mathcal{M} \in H$. Se entiende que la dinámica del sistema a lo largo del tiempo puede modelarse mediante un proceso regresivo

$$(s_j, t_j) = T((s_{j-1}, t_{j-1}), (s_{j-2}, t_{j-2}), \dots), \quad (3.10)$$

donde T es un proceso dinámico discreto que explicita la dependencia de la forma y ubicación de los ciclos de la señal en función de los ciclos precedentes. Es claro que encontrar la dinámica subyacente T para el sistema es una tarea difícil y sólo sería factible encontrar esta ley en casos particulares. La ventaja de este modelo sobre modelos fenomenológicos en el tiempo como (3.3) o (3.6), es que captura la variabilidad ciclo a ciclo en la morfología de la oscilación y no se limita a cambios *lentos* en la forma de onda a lo largo del tiempo.

Debido a que los datos existen en un espacio de altas dimensiones, es necesario realizar una reducción de dimensionalidad para poder analizarlos adecuadamente. Para este fin, los autores proponen aplicar el algoritmo de *mapas de difusión* [90]. Este método, basado en la teoría de grafos, permite obtener una representación de baja dimensión de los datos manteniendo la estructura geométrica presente en el espacio original.

Para encontrar los mapas de difusión a partir del conjunto de ciclos s_j , definimos primeramente una nube de puntos $\{s_j\}_{j=1}^n \in \mathbb{R}^p$, donde asumimos que los puntos s_j se distribuyen en una *manifold* de menor dimensionalidad $d < p$. A partir de esta nube de puntos, definimos la *matriz de afinidad* W de $n \times n$ con entradas

$$W_{ij} = e^{-\frac{\|s_i - s_j\|_2^2}{h}}, \text{ para } i, j = 1, \dots, n. \quad (3.11)$$

3.3. PROPUESTA PARA EL MODELO NO ARMÓNICO ADAPTATIVO CON FORMAS DE ONDA VARIANTE EN EL TIEMPO

Posteriormente, se define la matriz de transición de un camino aleatorio a lo largo de la nube de puntos mediante

$$A(W) = D^{-1}W, \quad (3.12)$$

donde la matriz diagonal D esta dada por

$$D_{ii} = \sum_{j=1}^n W_{ij}, \text{ para } i = 1, \dots, n. \quad (3.13)$$

La matriz $A(W)$ puede descomponerse según $A(W) = U\Lambda V^T$, donde V contiene los autovectores derechos $\phi_1, \phi_2, \dots, \phi_n$ en sus columnas y la matriz diagonal Λ contiene los autovalores asociados $1 = \lambda_1 > \lambda_2 > \dots > \lambda_n$. Finalmente, podemos obtener una representación de baja dimensionalidad de los datos al calcular el mapa de difusión $\Phi_t^d : s_j \rightarrow (\lambda_2\phi_2(t), \lambda_3\phi_3(t), \dots, \lambda_{d+1}\phi_{d+1}(t))$. De esta forma, el agrupamiento de los puntos se respetará en el espacio de baja dimensión. Este algoritmo se aplicó para el análisis de señales de ECG de pacientes que son sometidos a anestesia total durante el proceso de intubación en terapia de cuidados intensivos [89]. Mediante este procedimiento, los autores muestran que la morfología de los ciclos se modifica considerablemente durante el proceso de intubación para luego recuperar su morfología natural. Este fenómeno queda reflejado en la representación de los ciclos en un espacio tridimensional, dónde los ciclos definen una trayectoria cerrada en el espacio 3D. Una de las principales limitaciones de esta propuesta es que no permite llevar adelante tareas de procesamiento de la señal como limpieza de ruido o descomposición de señales multicomponente. Además, los ciclos que se utilizan para formar la matriz de adyacencia pueden solaparse entre sí, por lo que habría información compartida entre los ciclos.

3.3. Propuesta para el modelo no armónico adaptativo con formas de onda variante en el tiempo

Para superar las limitaciones de los métodos existentes, proponemos un nuevo enfoque para la estimación de la WSF donde la variabilidad de la forma de onda queda descrita en conjunto por la proporcionalidad no entera de las fases armónicas y por la variación temporal de la amplitud relativa de los armónicos ($B_\ell(t)/B_1(t)$). Para describir nuestra propuesta, partimos de (3.2) y expandimos el primer término de la sumatoria

$$x(t) = B_1(t) \cos(2\pi\phi_1(t)) + \sum_{\ell=2}^{\infty} B_\ell(t) \cos(2\pi\phi_\ell(t)), \quad (3.14)$$

donde $B_\ell(t)$ es la modulación en amplitud del ℓ -ésimo “armónico” de la WSF. Vemos que el término $B_1(t) \cos(2\pi\phi_1(t))$ es el componente fundamental de $x(t)$ y, al compararlo con (3.6), identificamos $B_1(t) = A(t)$ suponiendo $a = 1$. Este componente fundamental codifica la variación global de la amplitud y la frecuencia de $x(t)$. En general, la modulación de amplitud global afecta a todos los armónicos de $x(t)$, por lo que dividimos la señal por $B_1(t)$ para obtener lo siguiente

$$\tilde{x}(t) = \frac{x(t)}{B_1(t)} = \cos(2\pi\phi_1(t)) + \sum_{\ell=2}^{\infty} \underbrace{\frac{B_\ell(t)}{B_1(t)}}_{Q_\ell(t)} \cos(2\pi\phi_\ell(t)). \quad (3.15)$$

3.3. PROPUESTA PARA EL MODELO NO ARMÓNICO ADAPTATIVO CON FORMAS DE ONDA VARIANTE EN EL TIEMPO

Esta expresión da lugar a una señal de forma de onda variable en el tiempo demodulada en amplitud. Según este nuevo modelo, la variación temporal de la forma de onda queda representada mediante la evolución temporal de las funciones armónicas de amplitud $Q_\ell(t) = B_\ell(t)/B_1(t)$ y las funciones armónicas de fase $\phi_\ell(t)$. Bajo esta consideración, la función de forma de onda no cambia si las funciones $Q_\ell(t)$ son constantes y si $\phi'_\ell(t) = \ell\phi'_1(t)$, con $\ell = 2, 3, \dots$. En general, la fase de cada componente puede contener un desfase constante φ_ℓ que difiere para cada armónico. Para tener esto en cuenta, reescribimos el término $\cos(2\pi\phi_\ell(t))$ como la suma de una onda seno y una onda coseno con una nueva función de fase $\Phi_\ell(t) = \phi_\ell(t) - \varphi_\ell$:

$$\tilde{x}(t) = \cos(2\pi\phi_1(t)) + \sum_{\ell=2}^{\infty} Q_\ell(t) (a_\ell \cos(2\pi\Phi_\ell(t)) + b_\ell \sin(2\pi\Phi_\ell(t))), \quad (3.16)$$

donde suponemos $\varphi_1 = 0$ y, en consecuencia, $b_1 = 0$. A partir de este nuevo planteo, y basándonos en el modelo SAMD, proponemos $\Phi_\ell(t) = e_\ell \hat{\phi}_1(t)$, con $e_\ell \in \mathbb{R}$ y siendo $\hat{\phi}_1(t)$ una estimación de la fase del componente fundamental obtenida a partir de la TFR de $\tilde{x}(t)$. Esta parametrización de las funciones de fase $\Phi_\ell(t)$ nos permite arribar al siguiente modelo

$$\tilde{x}(t) = \cos(2\pi\phi_1(t)) + \sum_{\ell=2}^{\infty} Q_\ell(t) (a_\ell \cos(2\pi e_\ell \phi_1(t)) + b_\ell \sin(2\pi e_\ell \phi_1(t))). \quad (3.17)$$

Cabe destacar que esta formulación del modelo introduce dos nuevos coeficientes, a_ℓ y b_ℓ , que deben estimarse para cada armónico. Se puede simplificar el modelo haciendo lo siguiente

$$\tilde{x}(t) = \cos(2\pi\phi_1(t)) + \sum_{\ell=2}^{\infty} Q_\ell(t) a_\ell \left(\cos(2\pi e_\ell \phi_1(t)) + \frac{b_\ell}{a_\ell} \sin(2\pi e_\ell \phi_1(t)) \right), \quad (3.18)$$

fijando $\alpha_\ell(t) = Q_\ell(t)a_\ell$ y $c_\ell = a_\ell/b_\ell$ obtenemos

$$\tilde{x}(t) = \cos(2\pi\phi_1(t)) + \sum_{\ell=2}^{\infty} \alpha_\ell(t) (\cos(2\pi e_\ell \phi_1(t)) + c_\ell \sin(2\pi e_\ell \phi_1(t))). \quad (3.19)$$

De esta forma, el desfase constante φ_ℓ queda codificado por los coeficientes c_ℓ . Nos referiremos indistintamente tanto a $Q_\ell(t)$ como a $\alpha_\ell(t)$ como *funciones de amplitud armónica* (HAFs, por las siglas en inglés de *harmonic amplitude functions*) dado que sólo se diferencian por una constante de escala.

3.3.1. Algoritmo para extracción de función de forma de onda variable en el tiempo

Para poder determinar los parámetros que definen la forma de onda variante en el tiempo, consideramos el siguiente problema de optimización por mínimos cuadrados:

$$\min_{\substack{c_\ell, e_\ell, \\ \alpha_\ell(t)}} \left\| \tilde{x}(t) - \sum_{\ell=1}^r \alpha_\ell(t) \Theta_\ell(t) \right\|_2^2, \quad (3.20)$$

donde $\alpha_\ell(t)$ es la HAF para el ℓ -ésimo componente armónico de la forma de onda variable en el tiempo y $\Theta_\ell(t) = \cos(2\pi e_\ell \phi_1(t)) + c_\ell \sin(2\pi e_\ell \phi_1(t))$. Fijamos $\alpha_1(t) = 1$, $c_1 = 0$, y $e_1 = 1$. Este problema puede pensarse como una regresión no lineal (NLR, del inglés *non-linear regression*) sobre el conjunto de parámetros $\{\alpha_\ell(t), c_\ell, e_\ell\}_{\ell=2}^r$. Los detalles sobre la resolución de este problema NLR se discutirán en la Sec. 3.5. Nótese aquí que sólo sumamos los primeros r componentes de la función de forma de onda variable en el tiempo (modulada en fase). Esto se basa en la condición (C2) para la clase de función de forma analítica (véase Def. 1.4.1), donde es razonable suponer que la WSF puede aproximarse adecuadamente mediante una función con un espectro de banda limitada. El parámetro r debe elegirse antes de resolver el problema de regresión no lineal. Para ello, utilizamos criterios de selección de modelos basados en la regresión trigonométrica. Como vimos en el Cap. 2, es posible aplicar estos criterios en el caso de señales no estacionarias para obtener una estimación precisa del número de componentes armónicos de la WSF.

3.4. Funciones de amplitud armónica

Un aspecto clave de nuestro modelo es cómo se caracteriza la variación de las funciones de amplitud armónica $\alpha_\ell(t) = Q_\ell(t)a_\ell$. Una descripción completa de cada amplitud armónica para cada instante de tiempo t conducirá a una representación excesivamente redundante de la información contenida en $x(t)$ y aumentará enormemente la carga computacional necesaria para resolver el problema de regresión. Por lo tanto, se requiere una representación más eficiente que permita capturar la variación temporal de las amplitudes armónicas. Esta consideración motiva la siguiente definición para una función de amplitud armónica:

Definición 3.4.1 (Función de amplitud armónica) *Dado $\ell \geq 2$, una función de amplitud armónica $\alpha_\ell(t) = Q_\ell(t)a_\ell$ como la dada en (3.19) es una función C^1 definida en un intervalo cerrado: $\mathbb{I} \rightarrow \mathbb{R}$, con norma-2 finita y*

$$|P_\ell(t) - \alpha_\ell(t)| < \epsilon_4, \text{ para } t \in \mathbb{I},$$

para un $\epsilon_4 > 0$ pequeño, donde $P_\ell(t)$ es una función interpoladora que utiliza I_ℓ nodos denotados como t_i , para $i = 1, \dots, I_\ell$.

3.4.1. Codificación de funciones de forma de onda variantes en el tiempo

Las funciones de amplitud armónica codifican información relevante sobre la variación temporal de la forma de onda. En base a la definición anterior, las HAFs son funciones suaves dependientes del tiempo que pueden aproximarse adecuadamente mediante una función interpoladora definida sobre un conjunto finito de nodos. Esto permite realizar una codificación comprimida de las HAFs con una pérdida mínima de información mediante interpolación utilizando nodos libres e interpoladores cúbicos. Este conjunto de I_ℓ nodos, localizados en los instantes de tiempo $t_{\ell,i}$ y con amplitudes $\alpha_{\ell,i} = \alpha_\ell(t_{\ell,i})$, se utilizarán junto a una función interpoladora $P_\ell(t)$, la cual se establece como una función definida a trozos que depende de la amplitud y localización de los nodos

$$P_\ell(t) = \sum_{i=1}^{I_\ell-1} P_{\ell,i}(t), \quad (3.21)$$

donde cada $P_{\ell,i}(t)$ es una función de soporte compacto definida sobre el subintervalo $[t_{\ell,i}; t_{\ell,i+1}] \subset [0, T]$ y $P_{\ell,i}(t_i) = \alpha_{\ell,i}$, para todos los nodos. Suele ser necesario establecer condiciones adicionales sobre la continuidad (primera derivada) y concavidad (segunda derivada) de la curva de interpolación, lo cual afectará la elección de las funciones de interpolación a utilizar.

En base a estas definiciones, las HAFs quedarán descritas mediante el conjunto de amplitudes nodales $\{\alpha_{\ell,i}\}$ y sus correspondientes ubicaciones temporales $\{t_{\ell,i}\}$, de forma que $\alpha_{\ell,i} = \alpha_\ell(t_{\ell,i})$ para $i = 1, \dots, I_\ell$. Generalmente, el número de nodos I_ℓ puede diferir para cada componente armónico. Consideramos nodos fijos al principio y al final de la señal: $t_{\ell,1} = 0$ y $t_{\ell,I_\ell} = T$ para todo $\ell \geq 2$. Los nodos interiores son libres de variar en su ubicación dadas las restricciones $0 < t_{\ell,i} < T$ y $t_{\ell,i-1} < t_{\ell,i} < t_{\ell,i+1}$ para todo $i = 2, \dots, I_\ell - 1$.

Dado el conjunto de amplitudes nodales $\{\alpha_{\ell,i}\}$ y localizaciones $\{t_{\ell,i}\}$, las HAFs se aproximan mediante una interpolación a trozos $\alpha_\ell(t) = \sum_{i=1}^{I_\ell-1} P_{\ell,i}(t)$, donde $P_{\ell,i}(t)$ es alguna función de interpolación adecuada definida sobre el subintervalo $I_{\ell,i} = [t_{\ell,i}, t_{\ell,i+1}]$ y $P_{\ell,i}(t_{\ell,i}) = \alpha_{\ell,i}$. Para nuestra propuesta, utilizamos interpoladores cúbicos hermitianos que conservan la forma [91] para la interpolación de las HAFs. Este tipo de interpolador presenta una serie de propiedades útiles. En primer lugar, es monótono en el intervalo $[t_{\ell,1}, t_{\ell,I_\ell}]$ si la secuencia de amplitudes de nodos $\{\alpha_{\ell,i}\}$ también es monótona. Además, si $\alpha_\ell(t)$ tiene un extremo local en $t_{\ell,i}$ entonces $P_{\ell,i}(t)$ también lo tiene. Los interpoladores cúbicos que conservan la forma tienen primeras derivadas continuas en los nodos, pero la continuidad de la segunda derivada no está garantizada.

3.4.2. Sobre la monotonidad y concavidad de las funciones de amplitud armónicas

Dada la Def. 3.4.1, necesitamos un esquema de interpolación que recupere la evolución temporal de la HAF lo más fielmente posible, dadas las condiciones de variación lenta y suavidad que establece el modelo (3.2). Además, es esperable que para una función monótona y convexa, la curva de interpolación resultante sea también monótona y convexa. Para conseguirlo, son útiles los interpoladores cúbicos que conservan la forma, ya que preservan la monotonidad y la concavidad de la curva interpolada. Otros interpoladores de uso común, como los splines cúbicos, dan lugar a curvas con ondulaciones indeseadas entre los nodos que afectan a la precisión de la curva estimada. Este fenómeno puede observarse en la Fig. 3.2 donde podemos ver la diferencia en la curva interpolada para nodos definidos sobre una función tangente hiperbólica.

3.5. Implementación numérica

Para implementar el algoritmo de estimación de forma de onda variante en el tiempo se requiere, en primera instancia, de la discretización de la señal. Para ello, fijamos una frecuencia de muestreo f_s en Hz (muestras/segundos) y muestreamos $x(t)$ a un intervalo regular dado por la tasa de muestreo $\Delta t = 1/f_s$ para obtener la señal discre-

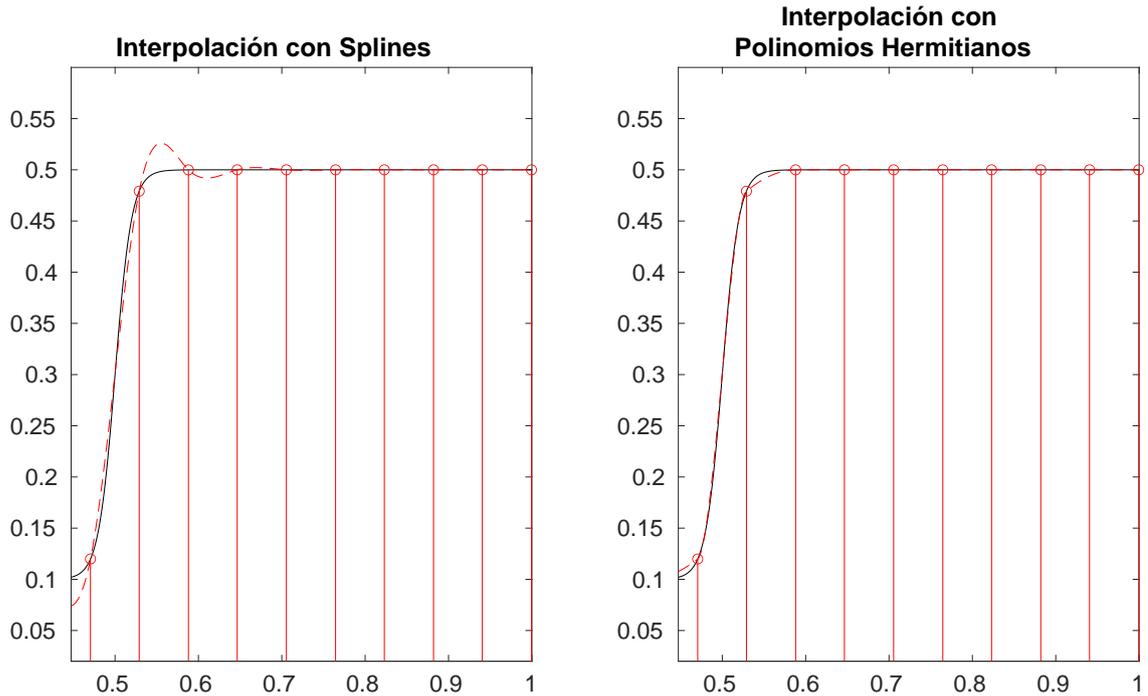


Figura 3.2: Demostración del efecto del interpolador sobre la estimación de las funciones de amplitud armónica. Izquierda: curva correspondiente a la tangente hiperbólica $0,3 + \tanh(t - 0,5)$ (en negro y línea continua) junto a la curva interpolada y los nodos de interpolación con splines cúbicos (en rojo y línea discontinua). Derecha: misma curva original (en negro y línea continua) junto a la curva interpolada y los nodos de interpolación con polinomios cúbicos hermitianos que conservan la forma (en rojo y línea discontinua).

ta $\mathbf{x}(n) = x(n\Delta t)$, con $1 \leq n \leq N$. Esta señal discreta tiene una duración finita de $T = N/f_s$ segundos. Dada la ecuación del modelo (3.19), las funciones de modulación de amplitud, modulación de fase y amplitud armónica también serán discretas: $\phi_1(n) = \phi_1(n\Delta t)$, $\mathbf{B}_1(n) = B_1(n\Delta t)$ y $\alpha_\ell(n) = \alpha_\ell(n\Delta t)$. Posteriormente, se realizarán una serie de pasos para obtener todos los elementos necesarios para plantear y resolver el problema de regresión no lineal descrito en (3.20). Los pasos a realizar para aplicar nuestro algoritmo se resumen en el diagrama de flujo de la Fig. 3.3. En las siguientes secciones, describiremos más en detalle diferentes aspectos de este algoritmo.

3.5.1. Efectos de borde y extensión de la señal

Un aspecto crucial a la hora de aplicar este algoritmo de extracción de forma de onda variable en el tiempo es evaluar el comportamiento de la solución cerca de los bordes de la señal. Una resolución directa del problema de NLR (3.20) demuestra que los efectos de borde se manifiestan de dos formas. En primer lugar, la ventana de análisis utilizada para calcular la STFT afecta a la resolución de la TFR en los bordes del espectrograma. Estos efectos deterioran la calidad de las estimaciones de $\mathbf{B}_1(n)$ y $\phi_1(n)$ en torno a las muestras inicial y final de la señal. El segundo problema se evidencia en el resultado mostrado en la Fig. 3.4, en la que se considera una señal de ejemplo

3.5. IMPLEMENTACIÓN NUMÉRICA

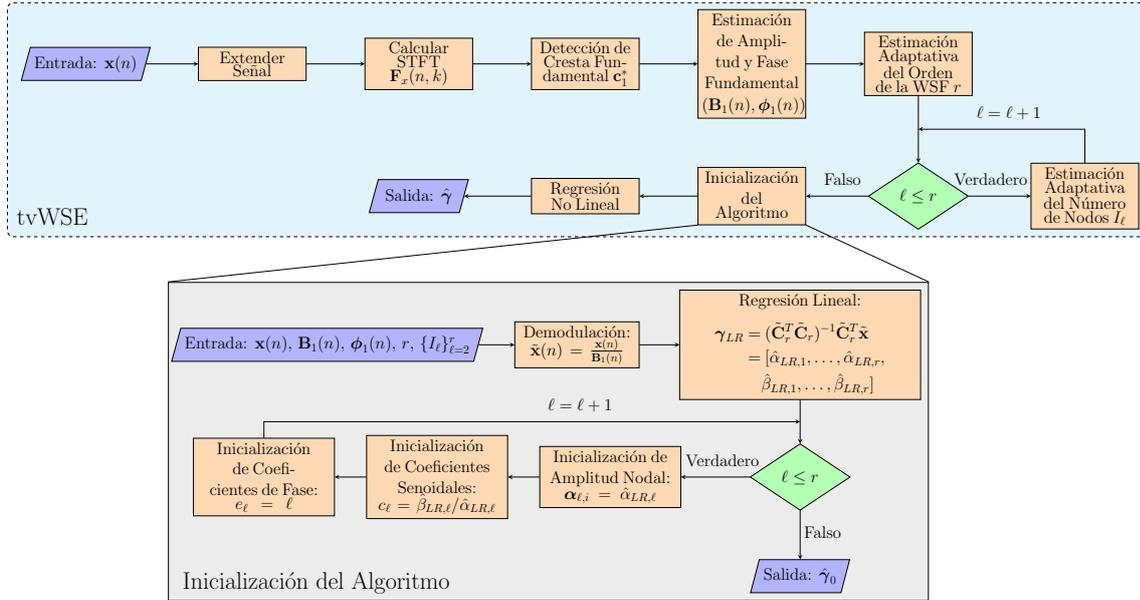


Figura 3.3: Diagrama de flujo del algoritmo adaptativo de extracción de funciones de forma de onda variables en el tiempo. En el recuadro gris se detallan los pasos para la obtención del vector inicial de coeficientes.

$$x(t) = \cos(2\pi\phi(t)) + \alpha_2(t) (\cos(2\pi e_2\phi(t)) + c_2 \sin(2\pi e_2\phi(t))), \quad (3.22)$$

donde la HAF del segundo armónico, denotada como $\alpha_2(t)$, evoluciona temporalmente de manera sinusoidal. Al comparar la HAF teórica con la estimada, vemos como los errores de estimación en el ajuste de curva basado en los nodos de interpolación provocan que la concavidad de la HAF no sea correcta. Esto se debe a que la interpolación cúbica con preservación de forma no puede especificar el valor de la primera derivada de la curva interpolada en los nodos extremos $t_{2,1}$ y t_{2,I_2} . Esto conduce a una estimación incorrecta de las amplitudes de los nodos en estos instantes y a un mal ajuste de la curva entre los nodos cercanos a los extremos.

Para sobrellevar estos efectos, se realiza como primer etapa del algoritmo una extensión de la señal en base a un modelo ARIMA estacional [92]. En primer lugar, se ajusta el modelo ARIMA estacional a partir de la información contenida en los 3 últimos ciclos y, a continuación, se utiliza el modelo ajustado para predecir las siguientes $N_p = 0,1N$ muestras. Los experimentos preliminares muestran que este valor de N_p ofrece un buen equilibrio entre el tiempo de cómputo y la atenuación de los efectos de borde. A continuación, extendemos la señal hacia atrás invirtiéndola: $x^*(t) = x(-t)$, y aplicando el mismo procedimiento. Nótese que esta extensión también añade un nodo fijo adicional al principio y al final para cada armónico a partir de $\ell = 2$, provocando entonces que el número total de coeficientes a estimar para cada señal sea $H = 2 \sum_{\ell=2}^r I_\ell + 2(r-1)$. Se elige este procedimiento de predicción mediante un modelo ARIMA por tratarse de un método de fácil implementación para la predicción de series temporales, siendo capaz de preservar el patrón oscilatorio al principio y al final de la señal.

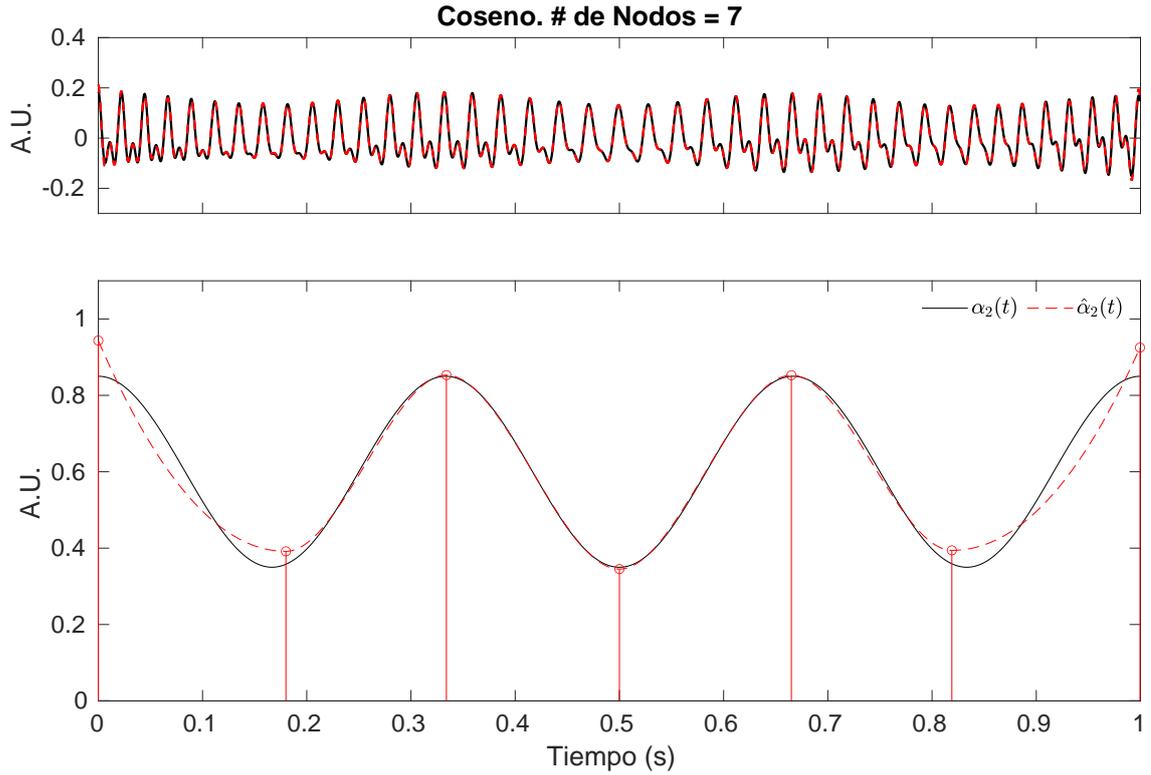


Figura 3.4: Demostración de los efectos de borde en la estimación de las funciones de amplitud armónica. Primera fila: Señal no estacionaria con forma de onda variante en el tiempo (negro, línea continua) junto con la curva resultante del algoritmo tvWSE superpuesta (rojo, línea discontinua). Segunda fila: HAF teórica para el segundo armónico (negro, línea continua), HAF interpolada junto a los nodos de interpolación (rojo, línea discontinua).

3.5.2. Estimación del número óptimo de nodos

El número de nodos de interpolación I_ℓ es un parámetro importante que puede variar de un armónico a otro y afecta en gran medida a la precisión de la curva estimada. Por lo tanto, es necesario realizar una estimación automática de I_ℓ para cada armónico de la señal. Para ello, en primer lugar observamos que para una señal modelada como en (3.19), se puede obtener una estimación aproximada de la versión complejizada del ℓ -ésimo armónico $\hat{y}_\ell(t)$ mediante la reconstrucción vertical de la STFT:

$$\hat{y}_\ell(t) = \frac{2}{g(0)} \int_{|f - c_\ell(t)| < \Delta_\ell} F_x^g(t, f) df, \quad (3.23)$$

donde Δ_ℓ es el ancho de banda de reconstrucción para el ℓ -ésimo armónico, y $c_\ell(t) \approx \phi'_\ell(t)$ es la estimación de la curva de frecuencia instantánea correspondiente a la ℓ -ésima cresta de $x(t)$. Aquí suponemos, sin pérdida de generalidad, que la señal $x(t)$ y, por lo tanto, los componentes armónicos $\hat{y}_\ell(t)$, están definidas en el intervalo $[0; 1]$. En esencia, $|\hat{y}_\ell(t)|$ es una estimación del módulo de la HAF del ℓ -ésimo armónico. Entonces, el número óptimo de nodos para interpolar la HAF puede estimarse considerando el módulo $|\hat{y}_\ell(t)|$ como una señal de ancho de banda limitado denotado como BW_ℓ . Para obtener

este ancho de banda, el espectro de energía (ES, del inglés *energy specturm*) de la amplitud armónica se calcula elevando al cuadrado el módulo de $\hat{Y}_\ell(f) = \mathcal{F}\{\hat{y}_\ell(t)\}$, donde \mathcal{F} es el operador de la transformada de Fourier.

El ancho de banda del espectro BW_ℓ puede estimarse como la coordenada f donde la energía espectral acumulativa normalizada

$$\text{CSE}(f) = \frac{\int_0^f |\hat{Y}_\ell(\xi)|^2 d\xi}{\int_0^\infty |\hat{Y}_\ell(\xi)|^2 d\xi}, \quad (3.24)$$

alcanza 90 % de la energía de la amplitud armónica estimada $|\hat{y}_\ell(t)|$. Utilizando el teorema de Nyquist, fijamos $I_\ell = 2 BW_\ell + 1$.

3.5.3. Inicialización del vector de coeficientes

La elección de los valores iniciales para los coeficientes del modelo es crucial para que el algoritmo de ajuste de curvas converja a una solución adecuada. Para nuestro algoritmo, proponemos una inicialización de “arranque en caliente” [93], donde en principio obtenemos los coeficientes del ajuste de la señal al ANHM de onda fija:

$$\hat{\gamma}_{LR} = (\tilde{\mathbf{C}}^T \tilde{\mathbf{C}})^{-1} \tilde{\mathbf{C}}^T \tilde{\mathbf{x}} = [\hat{\alpha}_{LR,1}, \dots, \hat{\alpha}_{LR,r}, \hat{\beta}_{LR,1}, \dots, \hat{\beta}_{LR,r}], \quad (3.25)$$

donde \mathbf{C} se define según (1.36) considerando $K = 1$. A continuación, las amplitudes iniciales de los nodos $\{\alpha_{\ell,i}\}$ se establecen iguales a los coeficientes de regresión lineal del coseno $\hat{\alpha}_{LR,\ell}$ para todos los nodos I_ℓ en cada armónico. Las ubicaciones iniciales de los nodos $t_{\ell,i}$ se establecen a partir de una cuadrícula uniforme de I_ℓ puntos en el intervalo $[0, T]$. Los coeficientes c_ℓ se inicializan como $\hat{\beta}_{LR,\ell}/\hat{\alpha}_{LR,\ell}$. Por último, los coeficientes de fase armónica se establecen en $e_\ell = \ell$ para $\ell = 2, \dots, r$. Estos coeficientes se concatenan para formar el vector inicial de parámetros γ_0 . Los pasos del procedimiento de inicialización del vector de coeficiente se muestran en la parte inferior del diagrama de flujo de la Fig. 3.3.

3.5.4. Estimación de los coeficientes del modelo

Dada la señal discreta $\tilde{\mathbf{x}}(n) = \mathbf{x}(n)/\mathbf{B}_1(n)$, queremos encontrar el vector óptimo de coeficientes de forma de onda variable en el tiempo

$$\gamma = \left[\{t_{\ell,i}\}_{i=2}^{I_\ell-1}, \{\alpha_{\ell,i}\}_{i=1}^{I_\ell}, c_\ell, e_\ell \right]_{\ell=2}^r \in \mathbb{R}^H. \quad (3.26)$$

Este vector de coeficientes óptimo puede estimarse resolviendo el problema de regresión no lineal por mínimos cuadrados (3.20) en tiempo discreto:

$$\hat{\gamma} = \min_{\gamma \in \mathbb{R}^H} \left\| \tilde{\mathbf{x}}(n) - \left[\sum_{\ell=1}^r \alpha_\ell(n) \odot \Theta_\ell(n) \right] \right\|_2^2, \quad (3.27)$$

donde $\alpha_\ell(n)$ es el resultado de interpolar las amplitudes de los nodos $\{\alpha_{\ell,i}\}$ en las posiciones de los nodos $\{t_{\ell,i}\}$ como se explicó en la Sec. 3.4.1 y $\Theta_\ell(n) = \cos(2\pi e_\ell \phi_1(n)) + c_\ell \sin(2\pi e_\ell \phi_1(n))$, para $n = 1, 2, \dots, N$.

Como se mencionó anteriormente, este problema de regresión es no lineal en los coeficientes del modelo dado (3.27), por lo que no existe una fórmula analítica para encontrar

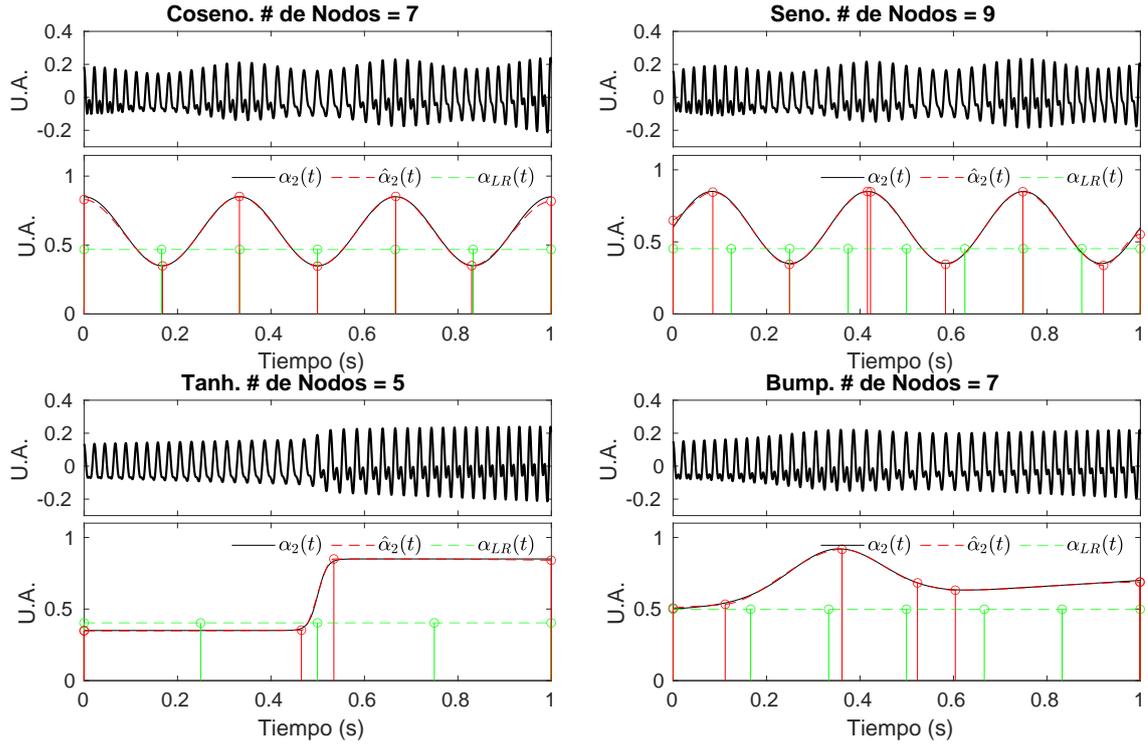


Figura 3.5: Demostración del método de estimación automática del número de nodos óptimos para la interpolación de las HAFs. Se muestran 4 ejemplos de señales con función de forma de onda variante en el tiempo según (3.28). Debajo de cada señal se muestran la HAF teórica (negro con línea continua) y la HAF estimada mediante el algoritmo propuesto (rojo con línea discontinua) junto con los nodos iniciales $\alpha_{i,LR}$ (puntos verdes) y los nodos óptimos $\hat{\alpha}_i$ (puntos rojos). El número de nodos que se reporta para cada señal fue obtenido de manera automática según el procedimiento detallado en la Sec. 3.5.2

la solución óptima del problema. Recurrimos, entonces, a un procedimiento iterativo de ajuste de curvas no lineales basado en el algoritmo de Levenberg-Marquardt [94].

Una vez obtenido el vector $\hat{\gamma}$, podemos estimar las funciones de amplitud armónica $\hat{\alpha}_\ell(n)$, para $\ell = 2, \dots, r$, mediante la interpolación con polinomios cúbicos hermitianos utilizando los nodos $\{\hat{\alpha}_{\ell,i}\}_{i=1}^{I_\ell}$ localizados en $\{\hat{t}_{\ell,i}\}_{i=1}^{I_\ell}$, notando que $\hat{t}_{\ell,1} = 1$ y $\hat{t}_{\ell,I_\ell} = N$ para $\ell = 2, \dots, r$.

Además, dada la relación $\hat{\alpha}(n) = \hat{\mathbf{Q}}_\ell(n)\hat{a}_\ell$, deducimos que $\hat{\mathbf{Q}}_\ell(n)$ codifica la variabilidad de la amplitud del ℓ -ésimo armónico *demodulado*. En cambio, las HAFs $\hat{\alpha}_\ell(n)$ codifican información sobre la variabilidad de la forma de onda junto con la amplitud armónica relativa dada por \hat{a}_ℓ . La relevancia de esta condición se discutirá en mayor detalle en la Sec. 5.3.

3.5.5. Ejemplo de aplicación

La combinación del procedimiento de extensión de la señal, la inicialización basada en el modelo WSF con forma de onda fija incorporando la selección adaptativa del número de armónicos del WSF, la estimación automática del número de nodos para las HAFs y la regresión no lineal conduce a un algoritmo de extracción de forma de onda variable en

el tiempo guiado por los datos. Esto implica que no es necesario establecer condiciones a priori sobre los parámetros del algoritmo para poder aplicarlo en diferentes contextos. Para ilustrar la potencialidad del algoritmo, consideraremos señales con funciones de forma de onda variante en el tiempo que se describen mediante la siguiente ecuación

$$x(t) = A(t) (\cos(2\pi\phi(t)) + \alpha_2(t) \cos(2\pi 2,005\phi(t))), \quad (3.28)$$

con $A(t) = \frac{1}{10}\sqrt{t+1}$, $\phi(t) = 40t + \frac{5}{2\pi} \sin(2\pi t)$. Respecto a la función de amplitud del segundo armónico, consideramos 4 formas diferentes para la evolución temporal de $\alpha_2(t)$, las cuales mostramos en la Fig. 3.5 junto con los nodos iniciales α_{LR} y las amplitudes nodales $\hat{\alpha}_{i,\ell}$, con ubicación en $\hat{t}_{i,\ell}$, obtenidos al resolver el problema de optimización (3.27). Vemos que al determinar de forma adaptativa el número de nodos I_ℓ en cada caso obtenemos una cantidad de nodos adecuada para describir cada una de las HAFs. Además, la condición de nodos libres le concede al algoritmo la capacidad ubicar los nodos de interpolación en las posiciones más adecuadas para obtener una buena estimación de las curvas asociadas a cada HAF. Finalmente, vemos que la extensión de la señal hacia atrás y hacia adelante permite mitigar los efectos de borde en la estimación manifestados en el caso estudiado en la Fig. 3.4.

Nos referiremos a este algoritmo de extracción de forma de onda variable en el tiempo totalmente adaptativo con el acrónimo tvWSE (de las siglas en inglés de *time-varying wave-shape extraction*). El algoritmo tvWSE se implementó utilizando MATLAB[®] versión R2021b. El código del método propuesto y todos los experimentos detallados en esta sección están disponibles en: <https://github.com/joaquinr-uner/tvWSE>.

3.6. Validación en señales sintéticas

En esta sección, mostramos los resultados de varios experimentos con datos simulados para diversas tareas de procesamiento de señales. Para estos experimentos, sin pérdida de generalidad, las señales sintéticas se limitan al intervalo $[0, 1]$ y su media se fija en cero. Para todos los experimentos de esta sección, las señales sintéticas se discretizaron utilizando una frecuencia de muestreo $f_s = 2000$ Hz.

Las estimaciones de amplitud y modulación de fase, $\hat{\mathbf{B}}_1(n)$ y $\hat{\phi}_1(n)$ respectivamente, se obtuvieron a partir de la reconstrucción alrededor de la cresta fundamental de la STFT, según (1.27). Nótese que suponemos que el componente fundamental corresponde a la cresta más energética del espectrograma, lo que puede no ser siempre el caso. Como mencionamos en el capítulo anterior, herramientas como la de-shape STFT [89] pueden utilizarse para superar este problema. Sin embargo, se asume la dominancia de la cresta fundamental para simplificar el análisis.

Para calcular la STFT, utilizamos una ventana gaussiana $\mathbf{g}(n) = e^{-\sigma n^2}$ con una varianza de ventana $\sigma = 10^{-4}$. La cresta fundamental, denominada $\mathbf{c}(n)$, se obtuvo utilizando el algoritmo de extracción de crestas detallado en la Sec. 1.3.3. El número de componentes armónicos de la señal se estimó mediante el procedimiento descrito en la Sec. 2.3 utilizando el criterio de selección de Wang.

Comparamos nuestra propuesta con las reconstrucciones obtenidas mediante el algoritmo basado en regresión lineal (LR, del inglés *lineal regression*) para estimar la WSF de una señal $x(t)$ siguiendo el ANHM con forma de onda fija, cuyo procedimiento fue detallado en la Sec. 1.4.4. Además, comparamos con el método SAMD, el cual conside-

3.6. VALIDACIÓN EN SEÑALES SINTÉTICAS

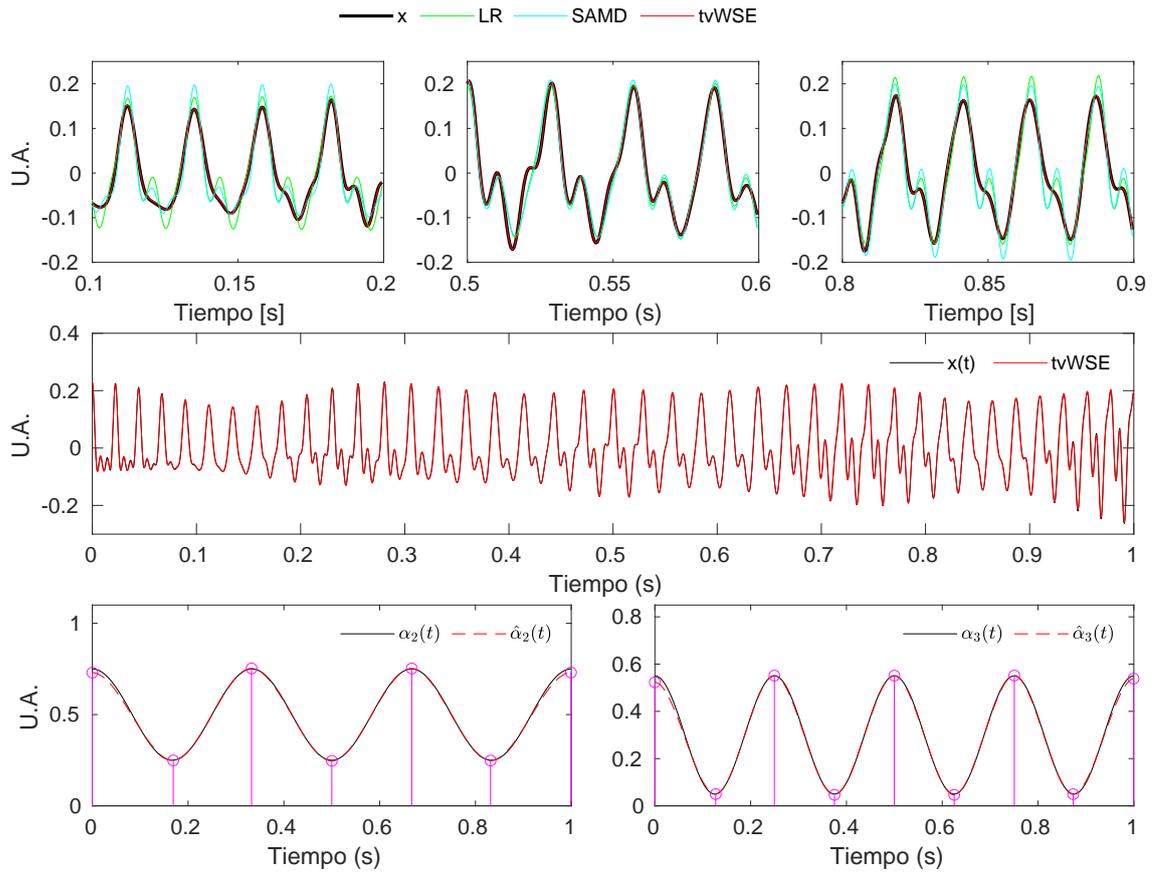


Figura 3.6: Resultados del método de estimación adaptativo de funciones de forma de onda variantes en el tiempo. Primera fila: segmentos de 100 milisegundos de duración al inicio, medio y final de las señales que resaltan la forma de onda variable en el tiempo con los resultados de LR, SAMD y tvWSE superpuestos. Segunda fila: señal original (negro) y reconstrucción con tvWSE (rojo). Tercera fila: funciones de amplitud armónica (línea completa) y estimación de las HAFs mediante nuestro algoritmo (línea discontinua) junto a los nodos óptimos (puntos).

ra una función de forma de onda variante en el tiempo. Los tres métodos se aplicaron a señales sintéticas para las tareas de estimación de forma de onda variable en el tiempo, descomposición y eliminación de ruido de señales monocomponentes y multicomponentes. En estos experimentos no se tuvo en cuenta el método MMD, debido a su lenta tasa de convergencia, que es dos órdenes de magnitud superior a la de los otros métodos. A su vez, el modelo oscilatorio de funciones de forma de onda no fue considerado ya que no permite realizar la limpieza de ruido o descomposición de la señal no estacionaria.

3.6.1. Caracterización de la función de forma de onda variante en el tiempo de una señal no estacionaria

Para demostrar la capacidad de nuestro algoritmo para recuperar la información asociada a las HAFs que componen la señal, consideramos la siguiente señal no estacionaria

con función de forma de onda variable en el tiempo

$$\begin{aligned}
 x(t) &= B_1(t) \left(\sum_{\ell=1}^3 \alpha_\ell(t) \cos(2\pi\phi_\ell(t)) \right), & (3.29) \\
 B_1(t) &= \frac{1}{10} \sqrt{t+1}, \quad \phi_1(t) = 40t + \frac{5}{2\pi} \sin(2\pi t), \\
 \alpha_1(t) &= 1, \quad \alpha_2(t) = 0,5 + 0,25 \cos(2\pi 3t), \quad \alpha_3(t) = 0,3 + 0,25 \cos(2\pi 4t), \\
 \phi_2(t) &= 2,005\phi_1(t), \quad \phi_3(t) = 2,995\phi_1(t).
 \end{aligned}$$

La señal se muestra en la segunda fila de la Fig. 3.6 con el resultado de reconstruir la señal con el vector óptimo de coeficientes $\hat{\gamma}$ superpuesto en rojo. Podemos ver que la forma de onda cambia con el tiempo y esta variación es capturada por el modelo mediante las HAFs $\alpha_\ell(t)$ y las funciones de fase $\phi_\ell(t)$. La estrecha correspondencia entre la señal original y nuestra estimación muestra cómo el algoritmo tvWSE es capaz de recuperar la información asociada a la forma de onda variable en el tiempo de $x(t)$. De hecho, observando la primera fila de la Fig. 3.6 vemos que los resultados utilizando LR y SAMD no puede capturar completamente la evolución temporal de la forma de onda. En la fila inferior de la Fig. 3.6, mostramos las HAFs teóricas $\alpha_2(t)$ y $\alpha_3(t)$, junto a los nodos óptimos obtenidos con nuestro algoritmo adaptativo y las curvas interpoladas trazadas con líneas discontinuas. Vemos que el patrón temporal de las amplitudes armónicas se recupera correctamente y que los efectos de borde se atenúan considerablemente gracias al proceso de extensión de la señal.

3.6.2. Limpieza de ruido en señales con función de forma de onda variable en el tiempo

En esta sección, evaluamos la calidad de nuestro procedimiento de extracción de formas de onda en presencia de ruido. Para ello, generamos 200 realizaciones de la señal ruidosa $x_n(t) = x(t) + n(t)$, donde $n(t)$ es ruido gaussiano de media cero y varianza σ^2 .

Para este experimento, se consideraron 5 niveles diferentes de SNR_{in} : 0, 5, 10, 15 y 20 dB. Además de la señal mostrada en la Fig. 3.6, estudiamos una familia más amplia de funciones de amplitud armónica. La fila superior de la Fig. 3.7 muestra una variedad de señales con formas de onda variables en el tiempo, dadas por diferentes combinaciones de las HAFs. Las dos primeras columnas muestran señales con una forma de onda suave y de variación lenta, dadas por las HAFs sinusoidales y polinómicas. Las señales de la tercera y cuarta columna presentan cambios más abruptos en la forma de onda, debido a la mayor velocidad de cambio de la tangente hiperbólica (\tanh) y a las modulaciones de amplitud de los bumps. Todas las señales consideradas en esta sección son de la forma $x_i(t) = B_1(t)s_i(2\pi\phi_1(t))$ para $i = 1, 2, 3, 4$, utilizando los mismos $B_1(t)$ y $\phi_1(t)$ de la sección anterior.

El desempeño de nuestro algoritmo se comparó de nuevo con las reconstrucciones LR y SAMD. Además, comparamos estos métodos basados en el ANHM con el procedimiento de limpieza de ruido por umbralado de la STFT, utilizando umbralado duro y suave. El desempeño de cada método se cuantificó utilizando la SNR de salida que se calcula utilizando la ecuación (2.18).

En la segunda y tercera fila de la Fig. 3.7 se representa la SNR_{out} de cada método, promediado sobre todas las realizaciones, en función de SNR_{in} junto con el intervalo de

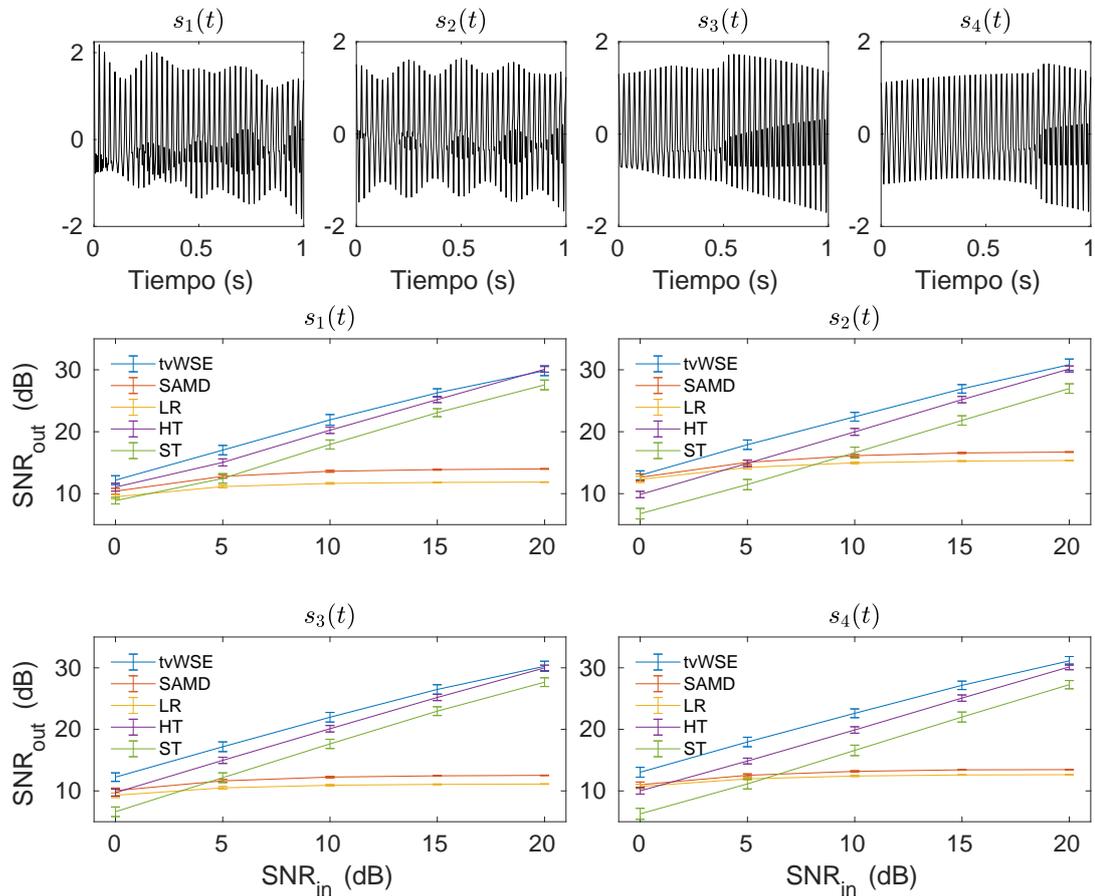


Figura 3.7: Limpieza de ruido para señales monocomponente con funciones de forma de onda variantes en el tiempo. Primera fila: Cuatro funciones de forma de onda diferentes que varían con el tiempo. $s_1(t)$ - la segunda y tercera HAF presentan una modulación cosenoidal. $s_2(t)$ - la segunda y tercera HAF presentan modulación lineal y cosenoidal, respectivamente. $s_3(t)$ - la segunda HAF tiene una modulación de tipo tangente hiperbólica (\tanh) y la tercera HAF tiene una modulación bump (gaussiana). $s_4(t)$ - la segunda y tercera HAF tienen modulación lineal y \tanh , respectivamente. Segunda y tercera fila: Resultados de la tarea de eliminación de ruido de cada método de eliminación de ruido para cada tipo de WSF variable en el tiempo. Referencias: LR - Regresión lineal - SAMD: Shape Adaptive Mode Decomposition - tvWSE: Extracción adaptativa de forma de onda variable en el tiempo - HT: umbralado duro de la STFT. ST: umbralado suave de la STFT.

3.6. VALIDACIÓN EN SEÑALES SINTÉTICAS

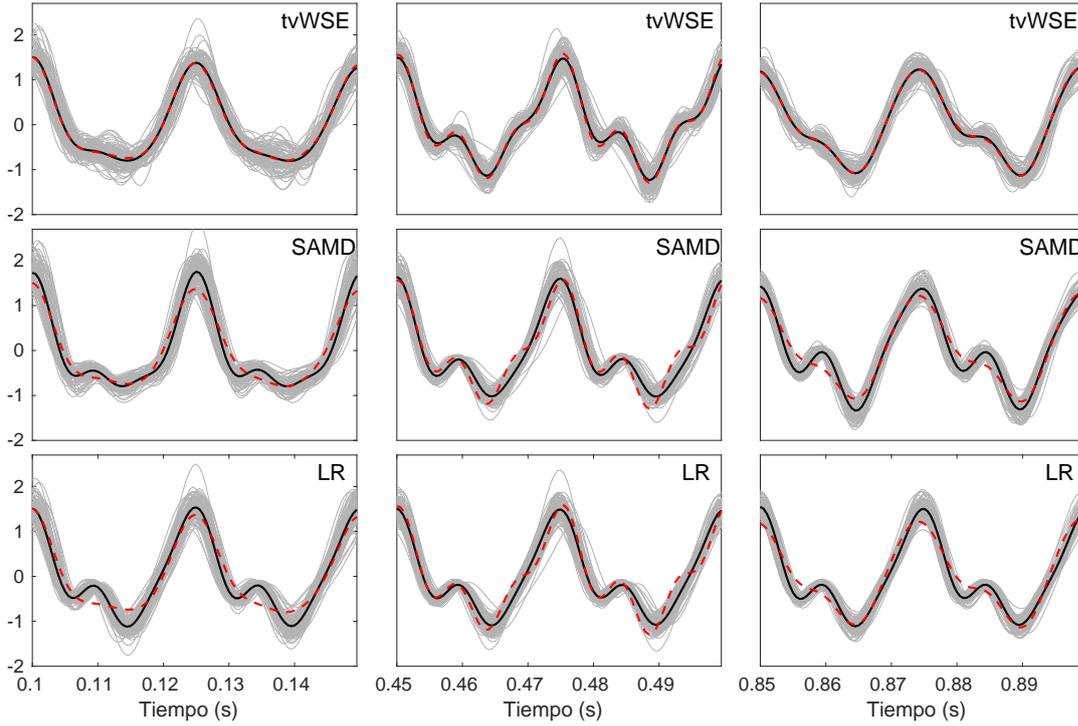


Figura 3.8: Comparación de las formas de onda extraídas con los diferentes métodos. Formas de onda extraídas con nuestro algoritmo (arriba), SAMD (centro) y LR (abajo). En cada fila, se representan en gris las ventanas de 50 milisegundos de longitud al inicio, centro y fin de las 200 realizaciones de la señal demodulada y sin ruido $\tilde{x}_{d,1}(t)$, para $\text{SNR}_{in} = 0$ dB. La forma de onda media de todas las realizaciones se representa en negro y la WSF original sin ruido $s_1(t)$ se representa con una línea discontinua roja.

desviación estándar. En primer lugar, observamos que a niveles bajos de SNR (0 dB), los tres métodos de reconstrucción basados en el ANHM y el método de umbralado duro de la STFT tienen un desempeño similar. El método STFT con umbral suave tiene un desempeño considerablemente inferior al resto en este nivel de SNR. Este menor desempeño para el método de umbral suave es esperable. Como se discute en [30] (p. 556), el umbral suave elimina los coeficientes TF con amplitud cercana al umbral, mientras que el umbral duro deja estos coeficientes sin cambios. A medida que aumenta la SNR_{in} , nuestro método y el de umbral duro tienden a obtener resultados similares, con el primero superando ligeramente al segundo. LR y SAMD muestran una rápida saturación en su desempeño, ya que no pueden seguir con precisión las variaciones temporales de la forma de onda. En el intervalo de 0 dB a 15 dB, nuestro método supera a todos los demás. A 20 dB, nuestro método alcanza una media de 30 dB de SNR_{out} , que es comparable a los resultados obtenidos mediante umbralado duro.

También comparamos la capacidad de los algoritmos basados en el ANHM para recuperar la forma de onda variante en el tiempo a partir de los datos ruidosos. Para ello, demodulamos las señales reconstruidas $\hat{x}_i(t)$ tanto en amplitud como en fase para obtener la señal demodulada:

$$\hat{x}_{d,i}(t) = \frac{\hat{x}_i\left(\frac{\hat{\phi}_1^{-1}(t)}{2\pi}\right)}{\hat{B}_1\left(\frac{\hat{\phi}_1^{-1}(t)}{2\pi}\right)}. \quad (3.30)$$

Dados estimadores suficientemente precisos de la amplitud y fase fundamentales, la señal demodulada aproxima a la forma de onda variable en el tiempo subyacente $\hat{x}_{d,i}(t) \approx s_i(t)$. La Fig. 3.8 muestra las estimaciones resultantes de la forma de onda variable en el tiempo para tvWSE, SAMD y LR para $s_1(t)$ contaminada con ruido de 0 dB. La línea negra gruesa y la línea roja discontinua representan la curva media de cada método y el valor real, respectivamente. En resumen, LR tiene un desempeño deficiente, ya que considera una forma de onda fija, lo que limita su capacidad para estimar con precisión la función de forma de onda variable en el tiempo. Tanto SAMD como nuestro método captan la variación temporal de la forma de onda y vemos cómo este último la sigue con mayor precisión al considerar la evolución temporal de las amplitudes armónicas.

3.6.3. Descomposición de señales multicomponente con funciones de forma de onda variables en el tiempo

Consideramos ahora la tarea de descomposición de señales multicomponentes cuando las señales componentes tienen funciones de forma de onda variantes en el tiempo. Para este análisis, denotaremos la señal multicomponente como $x(t) = x_1(t) + x_2(t)$, donde cada componente tiene una forma de onda variable en el tiempo. Así, $x_1(t)$ queda definida como

$$x_1(t) = B_{1,1}(t) \sum_{\ell=1}^3 \alpha_{1,\ell}(t) \cos(2\pi e_{1,\ell} \phi_{1,1}(t)), \quad (3.31)$$

$$B_{1,1}(t) = \sqrt{0,01t + 1,1}, \quad \phi_{1,1}(t) = 25t + \frac{5}{2\pi} \cos(2\pi t),$$

$$\alpha_{1,1}(t) = 1, \quad \alpha_{1,2}(t) = 0,5 + 0,2 \cos(2\pi 3t), \quad \alpha_{1,3}(t) = 0,3 + 0,2e^{-\left(\frac{t-0,25}{0,1}\right)^2},$$

$$\phi_{1,2}(t) = 2,005\phi_{1,1}(t), \quad \phi_{1,3} = 3,003\phi_{1,1}(t).$$

Análogamente, el segundo componente $x_2(t)$ es

$$x_2(t) = B_{2,1}(t) \sum_{\ell=1}^4 \alpha_{2,\ell}(t) \cos(2\pi e_{2,\ell} \phi_{2,1}(t)), \quad (3.32)$$

$$B_{2,1}(t) = 2 \log(t + 1,1) + 0,5, \quad \phi_{2,1}(t) = 100t + 7t^2,$$

$$\alpha_{2,1}(t) = 1, \quad \alpha_{2,2}(t) = 0,6 + 0,3t^2, \quad \alpha_{2,3}(t) = 0,4 + 0,5 \tanh(t - 0,5),$$

$$\alpha_{2,4}(t) = 0,3 + 0,3 \cos(2\pi 4t),$$

$$\phi_{2,2}(t) = 2,002\phi_{2,1}(t), \quad \phi_{2,3}(t) = 3,002\phi_{2,1}(t), \quad \phi_{2,4}(t) = 3,998\phi_{2,1}(t).$$

La descomposición se realizó de forma deflacionaria mediante un procedimiento de dos pasos en tiempo discreto. Para ello, primero resolvimos el siguiente problema de mínimos cuadrados

$$\hat{\gamma}_1 = \min_{\gamma_1 \in \mathbb{R}^H} \left\| \tilde{\mathbf{x}}(n) - \left[\sum_{\ell=1}^{r_1} \alpha_{1,\ell}(n) \odot \Theta_{1,\ell}(n) \right] \right\|_2^2, \quad (3.33)$$

con $\tilde{\mathbf{x}}(n) = \mathbf{x}(n)/\hat{\mathbf{B}}_{1,1}(n)$ y $\Theta_{1,\ell}(n) = \cos\left(2\pi e_{1,\ell} \hat{\phi}_{1,1}(n)\right) + c_{1,\ell} \sin\left(2\pi e_{1,\ell} \hat{\phi}_{1,1}(n)\right)$. Los estimadores $\hat{\mathbf{B}}_{1,1}(n)$ y $\hat{\phi}_{1,1}(n)$ del primer componente se obtienen a partir de la STFT

de \mathbf{x} . El vector de coeficientes óptimo $\hat{\gamma}_1$ se utilizó para reconstruir el primer componente $\hat{\mathbf{x}}_1(n)$. A continuación, calculamos la señal residual $\mathbf{x}_r(n) = \mathbf{x}(n) - \hat{\mathbf{x}}_1(n)$. Dada una estimación precisa del primer componente, podemos considerar $\mathbf{x}_r(n) \approx \mathbf{x}_2(n)$ y los coeficientes óptimos para el segundo componente se pueden obtener mediante la resolución de

$$\hat{\gamma}_2 = \min_{\gamma_2 \in \mathbb{R}^H} \left\| \tilde{\mathbf{x}}_r(n) - \left[\sum_{\ell=1}^{r_2} \alpha_{2,\ell}(n) \odot \Theta_{2,\ell}(n) \right] \right\|_2^2, \quad (3.34)$$

donde $\tilde{\mathbf{x}}_r(n) = \mathbf{x}_r(n)/\hat{\mathbf{B}}_{2,1}(n)$ y $\Theta_{2,\ell}(n) = \cos\left(2\pi e_\ell \hat{\phi}_{2,1}(n)\right) + c_{2,\ell} \sin\left(2\pi e_\ell \hat{\phi}_{2,1}(n)\right)$. Los estimadores $\hat{\mathbf{B}}_{2,1}(n)$ y $\hat{\phi}_{2,1}(n)$ del segundo componente se obtienen a partir de la STFT de \mathbf{x}_r . Este esquema de descomposición se aplicó a señales multicomponente ruidosas $x(t) + n(t)$, siendo $n(t)$ ruido gaussiano de media cero a 5 niveles SNR diferentes: 0, 5, 10, 15 y 20 dB. Para cada nivel de ruido, se generaron 100 realizaciones y se calculó la SNR de salida para cada componente. Como referencia, nuestro método se comparó con los resultados de LR y SAMD, que también son capaces de realizar simultáneamente la eliminación de ruido y la descomposición de señales oscilatorias multicomponente. Los métodos de umbralado no son capaces de realizar la descomposición de las señales, por lo que no se tuvieron en cuenta en este experimento. Las tres filas superiores de la Fig. 3.9 muestran los resultados de la descomposición para $x_1(t)$ y $x_2(t)$ por separado, y la superposición $\hat{x}_1(t) + \hat{x}_2(t)$, obtenida a 10 dB de ruido de entrada, comparada con la señal multicomponente $x(t)$. Nuestro algoritmo de descomposición recupera ambos componentes con precisión incluso cuando la forma de onda varía con el tiempo y en presencia de ruido.

Los resultados completos del experimento se muestran en la última fila de la Fig. 3.9, donde mostramos SNR_{out} frente a SNR_{in} para ambos componentes. De estas curvas, vemos que a partir de $\text{SNR}_{in} = 5$ dB, nuestro método supera a LR y SAMD en la tarea de descomposición y limpieza de ruido para ambos componentes. Conseguimos el mejor desempeño utilizando nuestro método aunque la varianza es algo mayor para todos los niveles de ruido de entrada. Muchos factores afectan al desempeño del algoritmo, desde las estimaciones de las modulaciones en amplitud y en fase de cada componente, el orden óptimo r y el número adaptativo de nodos. Es necesario seguir analizando esta tarea de descomposición para mejorar el desempeño general del algoritmo. No obstante, el método propuesto funciona mejor que los demás métodos basados en el ANHM para todos los niveles de ruido y compite con el umbralado de la STFT en el caso monocomponente.

3.6.4. Detección de instante de cambio en la función de forma de onda

En esta sección, consideramos el caso de señales no estacionarias con transiciones abruptas en su forma de onda. Para modelizar estas transiciones, consideramos la siguiente señal

$$x(t) = B_1(t) \cos(2\pi\phi_1(t)) + \sum_{\ell=1}^r B_\ell(t) \cos(2\pi e_\ell \phi_1(t)), \quad (3.35)$$

donde $B_1(t) = 0,1\sqrt{t+1}$, $\phi_1(t) = 40t + 5/(2\pi) \sin(2\pi t)$ y las HAFs

$$\alpha_\ell(t) = B_\ell(t)/B_1(t) = \mu + \lambda \tanh(\kappa(t - t_t)), \text{ para } \ell = 2, \dots, r, \quad (3.36)$$

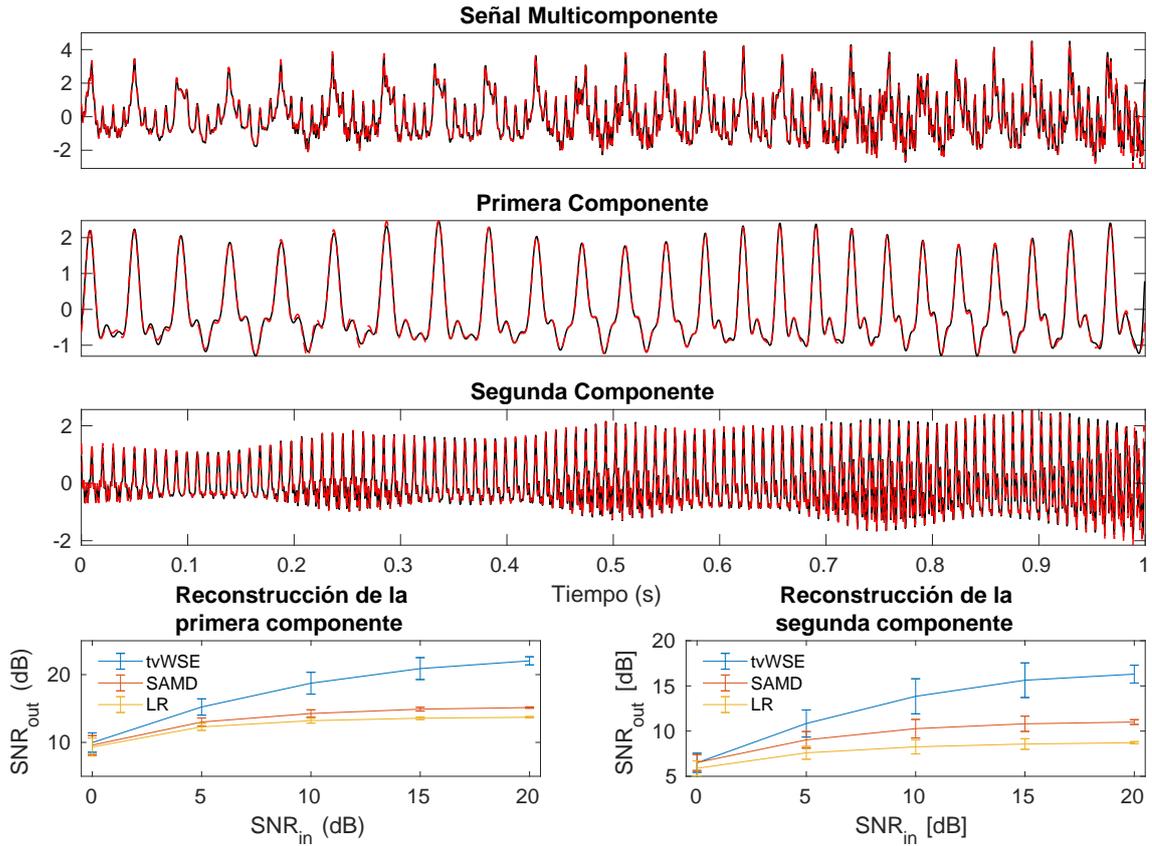


Figura 3.9: Limpieza de ruido y descomposición de señales multicomponente con funciones de forma de onda variantes en el tiempo. Primera fila: Señal multicomponente $x(t)$ representada en negro con el resultado de la reconstrucción para $\text{SNR}_{in} = 10$ dB superpuesto en línea roja discontinua. Segunda fila: Lo mismo para el primer componente $x_1(t)$. Tercera fila: Lo mismo el segundo componente $x_2(t)$. Cuarta fila: SNR_{out} en función de SNR_{in} para todos los métodos de descomposición de forma de onda para ambas señales componentes.

donde t_t es el tiempo de transición. Para $t < t_t$, la amplitud armónica tendería a $\mu - \lambda$, mientras que para $t > t_t$ la amplitud armónica tendería a $\mu + \lambda$. El parámetro κ controla la inclinación de la transición en el tiempo t_t . Este cambio en la amplitud armónica relativa se traduce en cambios bruscos en la forma de la onda. Para realizar la segmentación de la señal, hay que detectar los instantes de tiempo en los cuales la forma de onda cambia significativamente. Según el modelo propuesto, la única transición se produce en el tiempo t_t , por lo que el problema se reduce a localizar este punto de cambio. Observamos aquí que de los métodos discutidos hasta ahora, sólo nuestra propuesta permite la detección de estos cambios abruptos de forma de onda.

Utilizando nuestro procedimiento de extracción de forma de onda, las HAFs obtenidas por interpolación de los nodos $\alpha_{\ell,i}$ en las posiciones $t_{\ell,i}$ codifican la naturaleza de cambio temporal de la forma de onda. Teniendo esto en cuenta, se desarrolló un procedimiento adaptativo de estimación del tiempo de transición basado en las HAFs. Para cada HAF, el tiempo de transición se estima utilizando un algoritmo de detección de punto de cambio [95] y el tiempo promediado sobre todos las HAFs se reporta como el tiempo

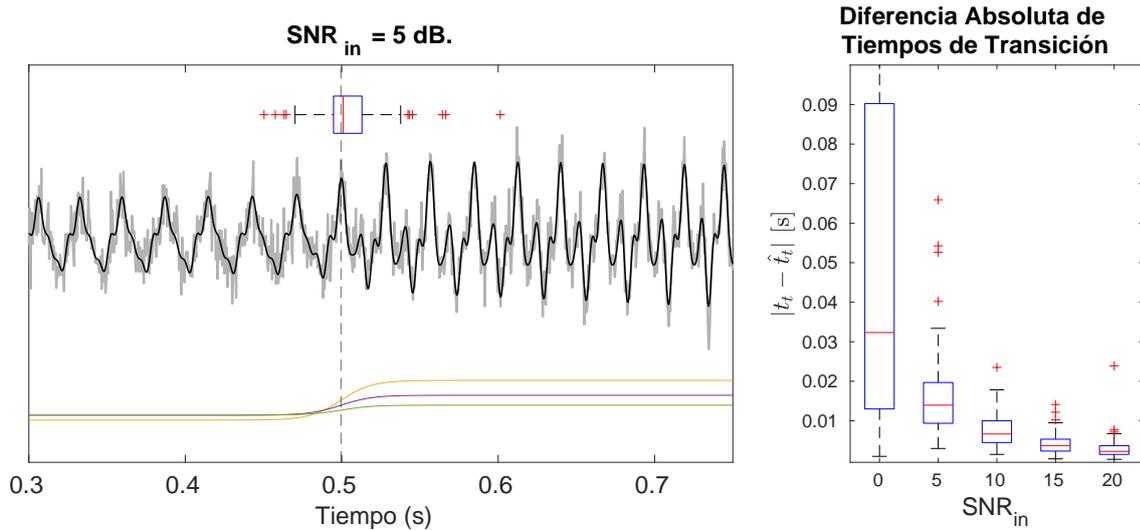


Figura 3.10: Detección de instante de cambio de la función de forma de onda. Izquierda: señal ruidosa no estacionaria con transición abrupta en su función de forma de onda representada en gris. La señal sin ruido se superpone en negro. Arriba se representa el diagrama de caja de los tiempos de transición estimados y abajo las HAFs teóricas. Derecha: diagramas de caja del error absoluto (EA) entre los instantes de transición reales y los estimados para distintos niveles de SNR_{in} .

de transición estimado \hat{t}_t . El panel izquierdo de la Fig. 3.10 muestra una señal de ejemplo con una transición brusca de forma de onda en $t = 0,5$. La señal se compone de 4 componentes armónicos ($r = 4$ en (3.35)).

Una realización ruidosa de la señal se muestra en gris con la señal sin ruido superpuesta en negro. Una línea vertical indica el instante de transición en $t_t = 0,5$. El primer experimento consistió en 100 realizaciones de la señal ruidosa que se analizaron utilizando nuestro algoritmo. Para cada una, se estimó el punto de transición y la distribución de \hat{t}_t se muestra como un boxplot sobre las señales representadas. Además, las HAFs teóricas se muestran debajo, evidenciando el efecto que tienen sobre la transición de forma de onda. Vemos en el diagrama de caja que el método puede estimar con precisión el tiempo de transición incluso en presencia de ruido. Posteriormente, se evaluó la robustez del algoritmo de segmentación propuesto. Para ello, se simularon 100 señales con diferentes transiciones abruptas de forma de onda. En este experimento, el tiempo de transición t_t se muestrea a partir de una distribución uniforme en el intervalo $[0,1; 0,9]$. El parámetro de transición κ se fija en 50. Los parámetros de amplitud μ y λ se muestrean a partir de distribuciones uniformes en $[0,1; 0,5]$ y $[0,1; 0,35]$, respectivamente. El número de componentes armónicos r se elige aleatoriamente para cada señal entre $r = 3$ y $r = 6$. Se añadió ruido gaussiano de media cero a las señales en cinco niveles de SNR_{in} : 0, 5, 10, 15 y 20 dB.

Las señales se discretizaron a una tasa de muestreo de $\Delta t = 0,5$ milisegundos. Nuestro método de extracción de la forma de onda se aplicó a cada señal simulada y el tiempo de transición se estimó a partir del punto de cambio de las HAFs, como se ha descrito anteriormente. Para cada una, se calculó el error absoluto (EA) entre el tiempo de transición real t_t y nuestra estimación. Los resultados de este experimento se muestran en el panel

derecho de la Fig. 3.10 mediante boxplots. Naturalmente, cuando la amplitud de ruido disminuye, el error de estimación para el instante de transición se reduce. Vemos que para niveles de SNR_{in} superiores a 0 dB la mayoría de los valores EA caen por debajo de 20 ms para una señal de 1 segundo de duración. Además, para $\text{SNR}_{in} > 10$ dB, la mediana de las distribuciones del EA son del orden de magnitud de la tasa de muestreo Δt . Hasta donde sabemos, no existe ningún otro método que aborde el problema de la detección de transiciones bruscas en la forma de onda de señales no estacionarias. Nuestra propuesta se vuelve entonces una alternativa interesante para abordar el análisis de señales no estacionarias cuyo patrón oscilatorio cambia de manera abrupta debido a cambios en los sistemas generadores.

3.7. Aplicaciones en señales biomédicas

Como ya hemos mencionado, los sistemas dinámicos del mundo real tienen un comportamiento altamente no lineal que da lugar a señales no estacionarias. En particular, los sistemas fisiológicos dan lugar a señales biomédicas oscilatorias con amplitud, frecuencia y forma de onda variables en el tiempo. Para mostrar la utilidad de nuestra propuesta, analizamos varias señales biomédicas incluyendo casos fisiológicos y patológicos. En primer lugar, mostramos un ejemplo de limpieza de ruido en una señal del EEG de un recién nacido durante un episodio de ataque epiléptico. A continuación, realizamos la descomposición de una señal de neumografía por impedanciometría (IP, del inglés *impedanciometry pneumography*) en sus componentes respiratorio y cardíaco. Por último, utilizamos nuestro algoritmo para estimar la forma de onda variable en el tiempo de una señal de ECG de un paciente antes y durante la fibrilación ventricular y utilizamos la información codificada en la función de amplitud del segundo armónico para segmentar la señal de manera adaptativa.

3.7.1. Aplicación 1: Limpieza de ruido en señales biomédicas

En esta sección, llevamos adelante el análisis de una señal de EEG de un recién nacido durante un episodio de ataque epiléptico. Esta señal forma parte de un conjunto de datos de acceso público [96]. El EEG se muestra en la primera fila de la Fig. 3.11, donde vemos claramente la presencia de ruido originado por la contracción muscular. Realizamos la limpieza de ruido de la señal EEG utilizando nuestro algoritmo y comparamos las formas de onda resultantes con las de SAMD y MMD. La señal se muestreó a 256 Hz y para este análisis se consideró una ventana de 20 segundos del evento convulsivo, lo que dio como resultado una señal discreta con 5120 muestras. Para la estimación de la amplitud y la fase fundamental, la STFT se calculó utilizando una ventana gaussiana con $\sigma = 2 \times 10^{-6}$, el salto de frecuencia máximo para el algoritmo de extracción de crestas se fijó en $I_f = 0,04$ Hz y el semiancho de reconstrucción del componente fundamental fue $\Delta = 0,4$ Hz. El orden del ANHM obtenido de manera automática para esta señal fue $r = 6$. Las formas de onda resultantes se muestran en la primera columna de la Fig. 3.11. Vemos que el resultado de nuestro método muestra más variabilidad entre los ciclos de la señal en comparación con los otros métodos. Particularmente al principio y al final de la señal, nuestra propuesta sigue la oscilación de la onda con mayor precisión.

Para comparar cualitativamente los resultados de la eliminación de ruido, se calculó la función de autocorrelación (ACF, del inglés *autocorrelation function*) de los residuos

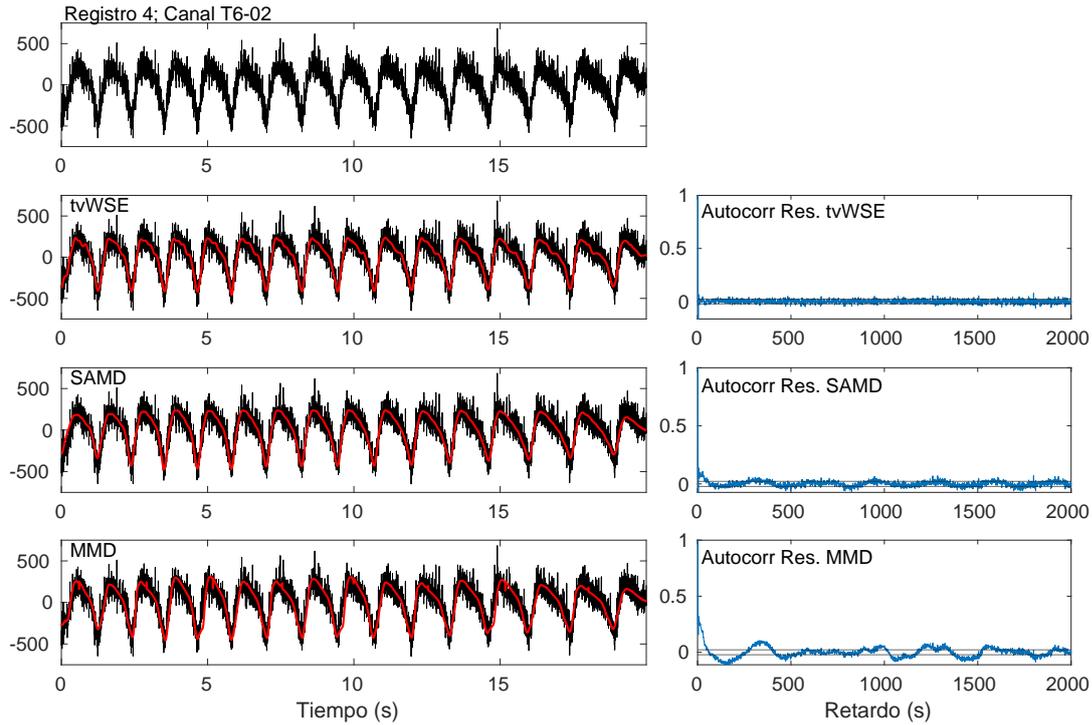


Figura 3.11: Limpieza de ruido en señales de electroencefalografía. Primera fila: Señal de EEG durante el evento epiléptico. Segunda fila: Señal de EEG con nuestro resultado superpuesto (izquierda) y la función de autocorrelación (ACF) del residuo entre las señales ruidosa y reconstruida (derecha). Tercera fila: Señal de EEG con el resultado de SAMD superpuesto (izquierda) y la ACF del residuo entre las señales (derecha). Cuarta fila: Señal de EEG con el resultado de MMD superpuesto (izquierda) y la ACF del residuo entre las señales (derecha).

para cada método

$$\text{ACF}(\tau) = \sum_{k=0}^N \mathbf{r}_i(k - \tau) \mathbf{r}_i(k), \text{ para } \tau = 0, \dots, \tau_{max}, \quad (3.37)$$

donde $\mathbf{r}_i(k) = \mathbf{x}(k) - \mathbf{x}_i(k)$ para la reconstrucción $\mathbf{x}_i(k)$ de cada uno de los algoritmos considerados. Las ACF normalizadas $\text{ACF}(\tau)/\text{ACF}(0)$ para tvWSE, SAMD y MMD, respectivamente, se representan en los paneles de la derecha de la Fig. 3.11, junto con el intervalo de 95 % de confianza: $[-1/\sqrt{N}, 1/\sqrt{N}]$. Vemos que nuestra propuesta conduce a un residuo con una estructura más cercana al ruido estacionario no correlacionado. Además, se calcularon la entropía espectral (SE, del inglés *spectral entropy*) del residuo y el coeficiente de correlación de Pearson (PCC, del inglés *Pearson's correlation coefficient*) entre el residuo y la señal reconstruida con cada uno de los métodos.

$$\text{SE}_i = \sum_{k=1}^N -\tilde{R}_i(k) \log_2(\tilde{R}_i(k)), \quad (3.38)$$

$$\text{PCC}_i = \frac{\text{Cov}(\mathbf{x}_i, \mathbf{r}_i)}{\sigma_{\mathbf{x}_i} \sigma_{\mathbf{r}_i}}, \quad (3.39)$$

donde \tilde{R}_i es el espectro de amplitud normalizado para el residuo r_i . La SE trata el espectro de potencia normalizado del residuo como una distribución de probabilidad y se calcula su entropía de Shannon. El ruido blanco no correlacionado se distribuye uniformemente en el dominio de la frecuencia, lo que da lugar a valores más altos de entropía espectral. Los valores de SE para MMD, SAMD y tvWSE fueron de 6,71, 7,10 y 7,30, respectivamente. Como referencia, el valor SE del ruido blanco uniforme con el mismo soporte es de 7,34. La entropía espectral del residuo para nuestra propuesta se aproxima más a este valor, lo que demuestra la mejora en la eliminación de ruido obtenida mediante nuestro procedimiento. PCC es una medida normalizada de la correlación lineal entre dos vectores de datos. Los valores más cercanos a 1 (-1) indican una fuerte correlación positiva (negativa). Cabe esperar que una mejor aproximación de la señal se traduzca en una menor correlación (PCC más cercana a 0) entre la estimación y el error residual. MMD, SAMD y tvWSE dieron como resultado valores de PCC de -0,25, -0,05 y -0,004, respectivamente. Estos valores muestran que nuestro algoritmo conduce a una mejor estimación de la forma de onda del episodio epiléptico sin ruido y un mejor desempeño de eliminación de ruido que los otros métodos.

3.7.2. Aplicación 2: Descomposición de señales multicomponente

En esta sección, aplicamos los algoritmos de estimación de forma de onda a un registro de IP. Este tipo de señal se compone normalmente de un componente respiratorio y un componente cardíaco, conocido como artefacto cardiogénico. Esta señal de IP se registró en un paciente que se sometió a un examen de broncoscopia flexible utilizando el monitor multiparámetro Philips Patient Monitor MP60 en el Chang Gung Memorial Hospital, en Linkou, Nuevo Taipei, Taiwán. El protocolo del estudio fue aprobado por la Junta de Revisión Institucional de la Fundación Médica Chang Gung (n° 104-0872C). La frecuencia de muestreo fue de 32 Hz. Se aplicaron SAMD, MMD y tvWSE a 60 segundos de la señal IP y se extrajeron los componentes respiratorio y cardíaco. La amplitud fundamental y la fase de ambos componentes se estimaron a partir del espectrograma utilizando $\sigma = 10^{-6}$, $I_f = 0,3$ Hz y $\Delta = 0,008$ Hz.

Mostramos las formas de onda resultantes para los tres métodos de descomposición en la Fig. 3.12. Con respecto al componente respiratorio, vemos que los tres métodos eliminan la tendencia de la señal y son capaces de recuperar la forma de onda del componente. El algoritmo tvWSE da como resultado una forma de onda más suave con una variabilidad de amplitud ligeramente superior. Esta forma de onda más suave es más atractiva para los clínicos y permite una mejor visualización de la información contenida en la señal respiratoria.

Con respecto al componente cardíaco, tanto nuestro método como SAMD recuperan una forma de onda suave y variable en el tiempo, en la que cada ciclo corresponde a un latido cardíaco, comparándola con la señal ECG registrada simultáneamente (representada en rojo). En particular, la forma de onda extraída utilizando nuestra propuesta muestra más variabilidad ciclo a ciclo que SAMD, como puede verse en los ciclos cardíacos segmentados en la última columna de la Fig 3.12. El componente cardíaco extraído con MMD muestra artefactos puntiformes que no están presentes en la señal original. Esta tarea de descomposición resulta interesante en el ámbito clínico ya que se puede obtener información clínica útil a partir del artefacto cardiogénico. En [57], la frecuencia cardíaca instantánea (IHR, del inglés *instantaneous heart rate*) se estima a partir del artefacto car-

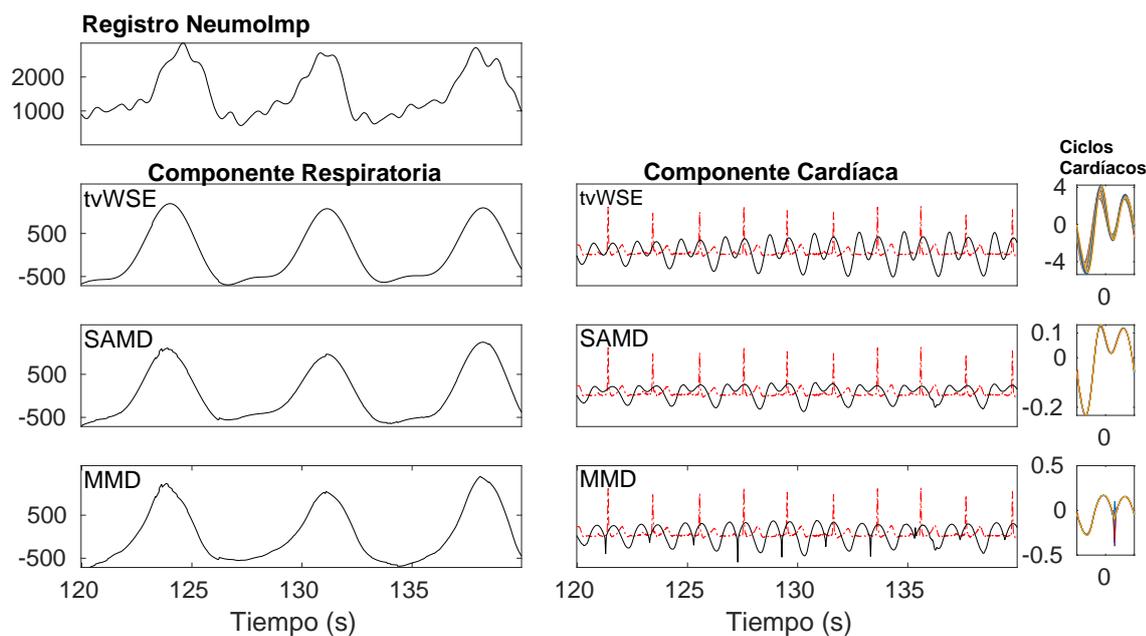


Figura 3.12: Descomposición de señal de neumografía por impedanciometría. Primera fila: 20 segundos de la grabación IP. Segunda fila: componentes respiratorios y cardíacos extraídos mediante nuestro algoritmo. Tercera fila: componentes respiratorios y cardíacos extraídos mediante SAMD. Cuarta fila: componentes respiratorios y cardíacos extraídos mediante MMD. En cada fila, la señal de ECG se representa junto al componente cardíaco con una línea discontinua roja. Los ciclos cardíacos segmentados con cada método se muestran en la columna de la derecha.

diogénico extraído de la señal IP. En [97], el volumen sistólico se estima a partir de las curvas de gasto cardíaco obtenidas de la señal de cardiografía por inductancia, tras la eliminación del componente respiratorio. El componente cardíaco extraído con forma de onda variable en el tiempo podría ser una alternativa útil para la estimación del gasto cardíaco.

3.7.3. Aplicación 3: Caracterización de transiciones en señales de electrocardiografía

En esta sección, analizamos una señal de ECG de un paciente taquiarrítmico antes y durante un episodio de fibrilación ventricular (FV) [98, 87]. La FV es una condición patológica en la que las cámaras inferiores del corazón se contraen rápidamente de forma descoordinada y provoca cambios considerables en la forma de onda del ECG. En este estudio, se analizaron 24 segundos de la señal de ECG, muestreada a 250 Hz, utilizando tvWSE, SAMD y MMD. El componente fundamental se estimó utilizando $\sigma = 5 \times 10^{-5}$, $I_f = 0,4$ Hz y $\Delta = 1,2$ Hz. El número de armónicos óptimo $r = 6$ se obtuvo utilizando el criterio de Wang.

Mostramos la señal de ECG en el panel superior de la Fig. 3.13 (correspondiente al registro cu03m de la *CU Ventricular Tachyarrhythmia Database*¹). Para ilustrar el

¹<https://physionet.org/content/cudb/1.0.0/>

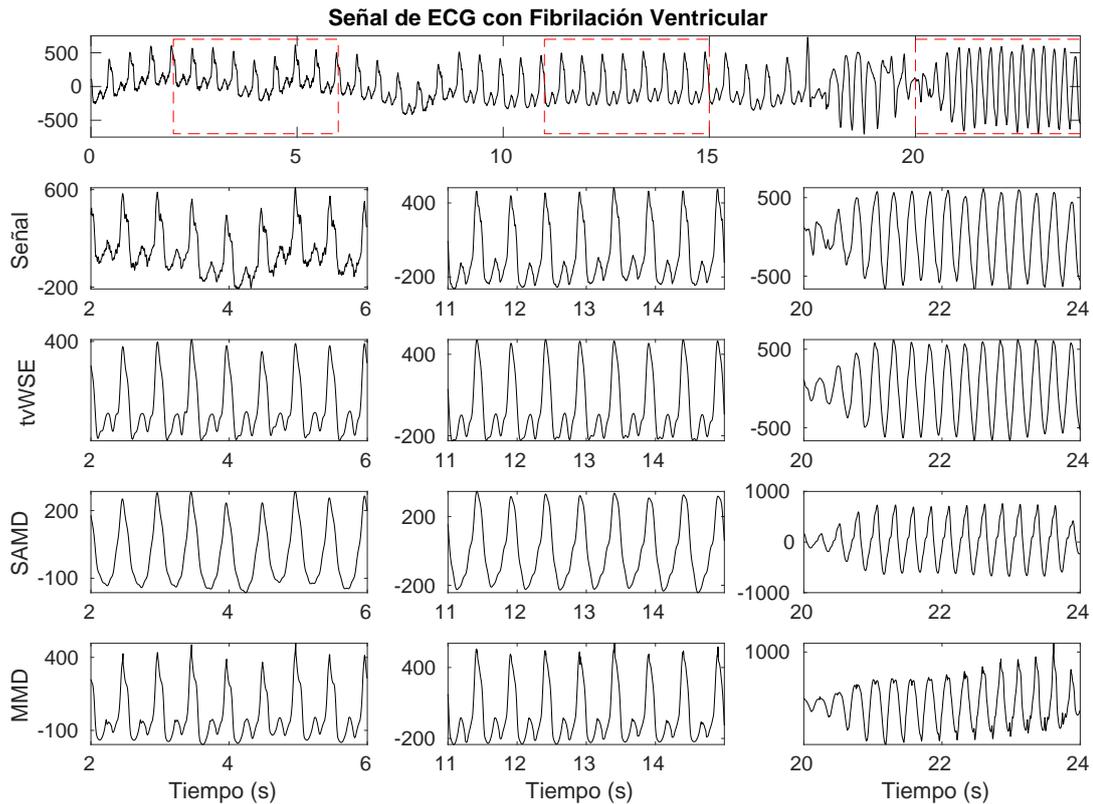


Figura 3.13: Reconstrucción de la señal de electrocardiografía durante fibrilación ventricular. Primera fila: señal de ECG con fibrilación ventricular que comienza en $t = 17$ s. Los segmentos ilustrativos del inicio, la mitad y el final de la señal se resaltan con recuadros rojos. Segunda fila: ventanas de 4 segundos de los segmentos resaltados de la señal original. Tercera fila: ventanas de 4 segundos de los segmentos resaltados de la forma de onda variable en el tiempo obtenida con nuestro método. Cuarta fila: el mismo resultado con SAMD. Quinta fila: el mismo resultado con MMD.

desempeño de la extracción de la forma de onda, se resaltan segmentos al inicio, mitad y final de la señal original con cuadrados rojos. Estos mismos segmentos se muestran en la segunda fila de la Fig. 3.13. Mostramos los segmentos correspondientes de la señal reconstruida obtenida con tvWSE, SAMD y MMD en la tercera, cuarta y quinta fila, respectivamente. Se observa un claro cambio en la forma de la onda entre la primera y segunda sección, asociadas al latido rítmico, y la última porción, asociada al evento de FV. Analizando las curvas, observamos en primer lugar que todos los métodos son capaces de realizar la *limpieza de ruido* y la *eliminación de la tendencia* al inicio de la señal, excepto MMD, que conserva parte de la tendencia en el segmento inicial. SAMD es incapaz de capturar completamente la forma de onda variable en el tiempo de la señal, en particular la forma de onda más compleja asociada con el latido rítmico del corazón. Al comparar nuestro método con el MMD, vemos que ambos recuperan la forma de onda del latido normal con gran precisión. En lo que respecta a la forma de onda de la fibrilación ventricular, el método MMD muestra artefactos puntiformes poco naturales en los picos y valles de los ciclos junto con una tendencia que aumenta lentamente. Además, en la primera porción de la señal, MMD conserva parte de la tendencia original de la señal, mientras que nuestro método es capaz de eliminar tanto la tendencia como el ruido de

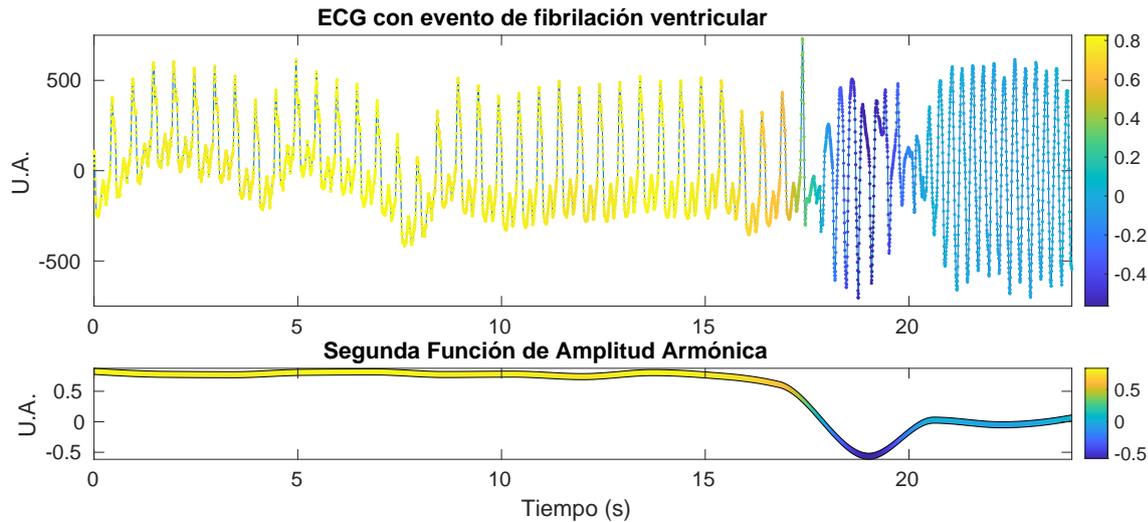


Figura 3.14: Caracterización de transiciones en la forma de onda de la señal de electrocardiografía. Primera fila: Señal de ECG original coloreada en función de $\alpha_2(t)$. Segunda fila: Función de amplitud del segundo armónico $\alpha_2(t)$.

baja amplitud.

Por último, mostramos de nuevo la señal ECG en el panel superior de la Fig. 3.14 con los ciclos coloreados en base a la función de amplitud del segundo armónico $\alpha_2(t)$. Esta última se muestra como una cinta coloreada en el panel inferior de la Fig. 3.14. Podemos ver cómo los diferentes valores de la amplitud armónica se asocian con las diferentes formas de onda de la señal. Comparando la curva HAF con la señal original mostrada en el panel superior, vemos una clara transición en el valor de $\alpha_2(t)$ en el instante en el que se inicia el evento FV. De este resultado se aprecia claramente que la evolución temporal de las HAFs codifica los cambios rápidos en la forma de la onda y que los eventos significativos en la señal pueden distinguirse a partir de los cambios en los parámetros del modelo. Esta información puede ser útil para la segmentación adaptativa y la detección de eventos en señales biomédicas en condiciones patológicas, así como la clasificación de señales según los cambios temporales de su patrón oscilatorio.

3.8. Comentarios de fin de capítulo

En este capítulo discutimos una variante del modelo no armónico adaptativo que permite considerar señales cuyas funciones de forma de onda cambian a lo largo del tiempo. Se presentaron métodos existentes en el estado del arte que buscan abordar el análisis y procesamiento de este tipo de señales, y se identificaron las limitantes de dichos métodos. Posteriormente, se propuso un nuevo algoritmo de extracción para estimar la WSF variable en el tiempo basado en el modelo no armónico adaptativo para el análisis de señales no estacionarias. El algoritmo considera tanto la fase armónica relativa mediante una condición de proporcionalidad no entera respecto a la fase fundamental, como así también las variaciones de amplitud relativa para caracterizar la variabilidad en la forma de onda. El método propuesto codifica la información asociada a las amplitudes armónicas variables en el tiempo a través de funciones de amplitud armónica que pueden representarse

como una serie de nodos libres que posteriormente se interpolan utilizando interpoladores cúbicos que conservan la forma de la curva interpolada.

Nuestro método se validó utilizando señales sintéticas y reales para las tareas de eliminación de ruido, descomposición y segmentación adaptativa. Comparamos nuestro método con los algoritmos existentes para la estimación de formas de onda. En el caso de señales sintéticas, nuestro método es capaz de caracterizar la naturaleza variable en el tiempo de las formas de onda con mucha más precisión que los métodos actuales, tanto para señales monocomponentes como multicomponentes, incluso en presencia de altos niveles de ruido.

Se obtuvieron resultados similares con señales reales. Para la eliminación de ruido del electroencefalograma epiléptico de recién nacidos, nuestra propuesta fue capaz de eliminar el ruido sin perder información relevante sobre la forma de onda variable en el tiempo de la señal, según las medidas consideradas. En cuanto a la descomposición de la señal de la neumografía por impedanciometría, nuestro método dio como resultado un componente respiratorio más suave que podría ser interpretado por los clínicos con mayor facilidad. Con respecto al componente cardíaco, nuestro método es capaz de recuperar una forma de onda con mayor variabilidad que los otros métodos, lo que cual resulta útil para obtener información sobre el comportamiento subyacente del sistema circulatorio. Al analizar un ECG durante un episodio de fibrilación ventricular, nuestro método fue capaz de seguir los cambios bruscos en la forma de onda de la señal con mucha más precisión que los métodos disponibles actualmente. Además, los coeficientes asociados a la segunda función de amplitud armónica variable en el tiempo se utilizaron para detectar transiciones bruscas en la forma de onda de la señal de ECG durante la fibrilación ventricular.

El algoritmo propuesto y los resultados presentados en este capítulo fueron incluidos en un artículo científico publicado en la revista *Signal Processing* de la editorial Elsevier [3].

3.8. COMENTARIOS DE FIN DE CAPÍTULO

Capítulo 4

Imputación de datos faltantes en señales biomédicas no estacionarias

La pérdida de información en las mediciones es un problema ubicuo en el sensado de magnitudes físicas. El tratamiento de esta condición depende de los conjuntos de datos y de los objetivos de su análisis, y por lo tanto ha surgido una gran diversidad de estrategias para abordar esta problemática. En este capítulo, nos centraremos en el problema de series temporales oscilatorias que presentan valores perdidos o faltantes. Exploraremos diferentes métodos existentes para la imputación de datos para su recuperación. Este problema surge frecuentemente en el ámbito de la informática médica, especialmente durante la monitorización prolongada de pacientes mediante sensores biomédicos. Factores como el movimiento del paciente, interrupciones en el sensado y los problemas de calibración pueden provocar la ausencia de valores en las series temporales biomédicas registradas. Aunque existen problemas similares en otros campos, nos centraremos principalmente en las aplicaciones biomédicas. En particular, las ideas discutidas en este capítulo pueden extenderse al análisis de series temporales relevantes en distintos ámbitos.

Comenzaremos describiendo la tarea de imputación de datos faltantes, y los diferentes fenómenos que pueden dar lugar a pérdidas de datos. Propondremos, además, una clasificación acorde a lo observado en los registros biomédicos. Luego, describiremos algunas estrategias existentes para el tratamiento de estos datos faltantes. El conjunto de métodos considerados contempla una gran diversidad en su fundamento teórico, así como las consideraciones prácticas para su implementación algorítmica. Respecto a esto último, haremos algunos comentarios sobre la elección de los parámetros de los algoritmos de imputación de datos faltantes. Posteriormente, describiremos nuestra propuesta para mejorar el desempeño de estos algoritmos, basándonos en la teoría del modelo no armónico adaptativo. Finalmente, validaremos el funcionamiento de nuestra propuesta en señales sintéticas y también en señales reales del ámbito biomédico.

4.1. Imputación de datos faltantes en series temporales

En [99, 100] se propuso una clasificación tradicional para la forma en la cual se manifiestan los datos faltantes en conjuntos de datos multidimensionales. Esta clasificación se basa en el problema general de datos faltantes, donde los registros se estudian como variables aleatorias. La descripción depende de los valores observados y no observados en un conjunto de datos. En primera instancia, los datos *faltantes completamente al azar*

(MCAR, por las siglas en inglés de *Missing Completely at Random*) se refiere a la situación en la que la falta de datos no depende de su valor, ya sea si los datos son observados o no observados. Aquí hay que tener en cuenta que esto no significa que el patrón de omisión de datos sea aleatorio. El caso de datos *faltantes al azar* (MAR, por las siglas en inglés de *Missing at Random*) se refiere a los casos en los que la falta de datos no depende de los datos no observados. MAR junto con el supuesto de que los parámetros del mecanismo de datos faltantes y el muestreo son diferentes se denomina *pérdidas ignorables*. Por último, la condición de datos *faltantes de manera no aleatoria* (MNAR, por las siglas en inglés de *Missing not at Random*) significa que la falta de datos depende de los datos no observados, o bien, que no es ignorable. Véase [99, 101] para una definición matemática rigurosa y [100] para ejemplos concretos en estudios clínicos.

La clasificación anterior es genérica y no tiene en cuenta las características de las series temporales. Aunque esta clasificación es factible cuando existe un modelo estadístico adecuado para una serie temporal [102, 103, 104, 105], es intrínsecamente difícil concebir un modelo de este tipo, especialmente en el ámbito biomédico. Con respecto a estas características, fenomenológicamente se pueden identificar tres tipos principales de “valores faltantes” para series temporales univariadas.

El primer tipo se puede denominar como datos *faltantes en el eje x*, como se ejemplifica en la señal del fotoplethismografía del panel superior de la Fig. 4.1. Matemáticamente, una señal $f(t)$ tiene datos faltantes en el eje x si se observa como

$$f(t) = \hat{f}(t)(1 - \chi_{I_{ms}}(t)), \quad (4.1)$$

donde $\hat{f}(t)$ representa a la señal sin falta de datos, $I_{ms} \subset \mathbb{R}$ define el/los *intervalo(s) de datos faltantes* y $\chi_{I_{ms}}$ es la función indicadora sobre I_{ms} . El problema de “predicción” (en inglés, *forecasting*) se alinea con este tipo de señales con datos faltantes en el eje x cuando $I_{ms} = (T, \infty)$, con T denotando el tiempo de muestreo actual.

El segundo tipo es el de datos *faltantes en el eje y*, como se ilustra en la señal de flujo de aire en el segundo panel de la Fig. 4.1. Una señal $f(t)$ con datos faltantes en el eje y se observa como, por ejemplo,

$$f(t) = \max\{\min\{\hat{f}(t), M\}, -N\}, \quad (4.2)$$

donde $M, N > 0$ normalmente provienen del desbordamiento aritmético. Este caso comúnmente recibe el nombre de *sobresaturación*.

El tercer tipo es el de datos faltantes *por corrupción*, demostrado con la señal de flujo de aire del tercer panel de la Fig. 4.1. Una señal $f(t)$ está corrompida si la observación para un instante t tiene la forma

$$f(t) = \hat{f}(t)(1 - \chi_{I_{ms}}(t)) + (c\hat{f}(t) + g(t))\chi_{I_{ms}}(t), \quad (4.3)$$

donde $c \geq 0$ cuantifica el escalado de la señal y g es una función que contiene información irrelevante. Dependiendo del escenario, g podría ser *ruido* estocástico o un *artefacto* determinístico. Aunque la señal corrompida incluye el caso comúnmente denominado como “señales ruidosas”, no nos centramos en este caso específico. En su lugar, nos centramos en la corrupción extrema, cuando $c = 0$ y g puede adoptar cualquier forma.

Aunque aquí nos centramos en la clasificación mencionada anteriormente en vez de la clasificación tradicional de [99], reconocemos su relevancia. Por ejemplo, en señales

4.1. IMPUTACIÓN DE DATOS FALTANTES EN SERIES TEMPORALES

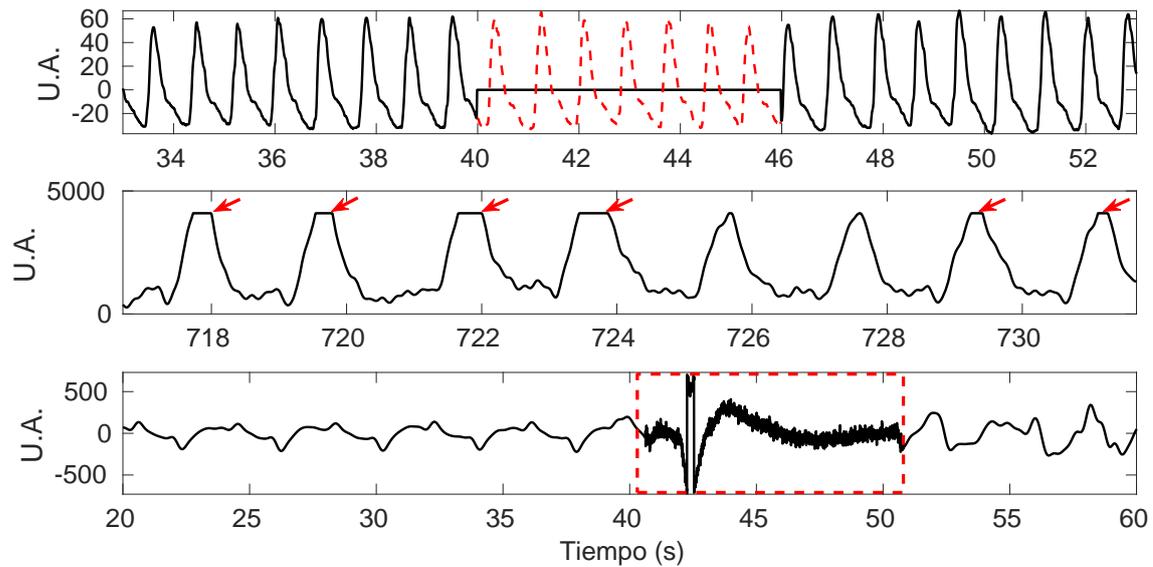


Figura 4.1: Ejemplos de señales con datos faltantes en el ámbito biomédico. Primera fila: señal de fotoplethysmografía en la que se aprecian los datos faltantes por la desconexión del sensor. Segunda fila: Señal de flujo de aire sobresaturada con picos saturados indicados con flechas rojas. Tercera fila: Señal de flujo de aire que muestra un segmento de baja calidad resaltado por el recuadro rojo.

biomédicas, el caso MCAR puede manifestarse cuando un sensor experimenta desconexiones aleatorias, dando lugar a una señal con datos faltantes en el eje x (ver Fig. 4.1(a)). Dependiendo de la situación, las señales corruptas causadas por artefactos de movimiento o interferencias (ver Fig. 4.1(c)) pueden asociarse al caso MAR o MCAR. La señal con datos faltantes en el eje y representada en la Fig. 4.1(b) puede considerarse como un ejemplo del caso MNAR.

Hasta donde sabemos, la investigación centrada en el tratamiento de las situaciones donde la falta de datos se da sobre el eje y es limitada. Normalmente, el tratamiento de los datos faltantes en y requiere información adicional. En los casos de señales corrompidas por artefactos, es posible convertirlas en configuraciones con datos faltantes en el eje x . Esta conversión implica identificar segmentos de señal de baja calidad y tratarlos como valores faltantes (asignando un valor de referencia como, por ejemplo, 0). Por lo tanto, a partir de este punto hablaremos principalmente de la configuración de datos faltantes en el eje x .

Existen dos estrategias habituales para tratar las señales con datos faltantes en el eje x . La primera es el método de “cortar y empalmar”, que consiste en eliminar los datos en los índices que faltan y analizar la señal con la información disponible. Aunque es sencillo, este método puede provocar la pérdida de detalles temporales importantes. El segundo método consiste en imputar valores en los instantes donde faltan datos. Diversos algoritmos de imputación pretenden generar nuevos datos que se ajusten a la dinámica subyacente, lo cual es crucial cuando los segmentos que faltan contienen información crítica acerca de, por ejemplo, el estado de salud de un paciente. Estos métodos van desde la sustitución por la media hasta soluciones más complejas basadas en modelos. En particular, la predicción, que es similar a un problema especializado de falta en el eje x , también puede utilizarse como método de imputación.

4.2. MÉTODOS PARA IMPUTACIÓN DE DATOS FALTANTES EN SERIES TEMPORALES

Las técnicas existentes para imputación de datos faltantes en señales no estacionarias abarcan algoritmos genéticos [106], interpolación [107], algoritmos difusos [108], aprendizaje múltiple [109], estimación bayesiana [110], enfoques basados en modelos autorregresivos [111], descomposición empírica en modos (EMD) [112, 113], y enfoques de aprendizaje profundo como BRITS [114], GP-VAE [115] e imputeGAN [116], entre otros. La principal diferencia entre los algoritmos tradicionales y los métodos de aprendizaje profundo es que los primeros pueden aplicarse directamente a las señales objetivo sin necesidad de largos periodos de entrenamiento y grandes volúmenes de datos.

En particular, los algoritmos basados en métodos de descomposición de señales aprovechan la estructura tiempo-frecuencia de las señales oscilatorias no estacionarias para llevar a cabo la imputación de datos faltantes. Por ejemplo, en [112], se propone una técnica de rellenado de huecos (*gap filling*) basada en la descomposición de la señal en IMFs mediante el algoritmo EMD. Este método aprovecha las oscilaciones lentas y la regularidad de las IMFs para mejorar la imputación de datos faltantes. De forma similar, en [113] los autores emplean una interpolación polinómica para rellenar los datos faltantes a nivel de las IMFs, reconstruyendo posteriormente la señal original mediante la superposición de modos. Sin embargo, la ventaja de este enfoque de descomposición basado en EMD sobre otras técnicas sólo se demostró para intervalos cortos de datos faltantes. También cabe señalar que las técnicas basadas en EMD pueden enfrentarse a problemas como la mezcla de modos inherente a EMD [117]. Además, sus fundamentos teóricos siguen siendo incompletos. Como alternativa, en este capítulo describiremos un algoritmo novedoso de imputación de datos perdidos basado en la teoría del ANHM, diseñado para evitar la mezcla de modos mediante la descomposición de las señales en términos de sus componentes armónicas.

El algoritmo propuesto, denominado *interpolación a nivel de los armónicos* (HaLI, por las siglas en inglés de *Harmonic Level Interpolation*), consta de dos etapas principales. La primera parte utiliza una técnica de imputación existente diseñada para señales oscilatorias para rellenar inicialmente los intervalos de datos faltantes. Para esta primera etapa, pueden utilizarse diferentes algoritmos de imputación inicial. Los algoritmos considerados en este trabajo para el estudio del desempeño de nuestra propuesta son los que se presentan en la Tabla 4.1. Describiremos estos algoritmos en mayor profundidad en la Sec. 4.2 y estableceremos criterios de selección de los parámetros para la implementación de dichos algoritmos en la Sec. 4.3. En la segunda etapa, HaLI realiza una descomposición armónica de la señal imputada y, posteriormente, interpola la amplitud y la fase de los armónicos obtenidos. Las diferentes etapas de HaLI se describirán en la Sec. 4.4. En la Sec. 4.5 mostraremos cómo los resultados obtenidos con HaLI evidencian una mejora del rendimiento, especialmente cuando la amplitud y la fase de los componentes armónicos de la señal siguen condiciones de regularidad específicas. Finalmente, en la Sec. 4.6 evaluaremos el desempeño de nuestro algoritmo en señales reales del ámbito biomédico.

4.2. Métodos para imputación de datos faltantes en series temporales

Como mencionamos en la sección anterior, el primer paso de nuestro algoritmo requiere un resultado de imputación preliminar obtenido mediante algún algoritmo existente. En esta sección, detallaremos los diferentes algoritmos de imputación de datos faltantes con-

4.2. MÉTODOS PARA IMPUTACIÓN DE DATOS FALTANTES EN SERIES TEMPORALES

Nombre	Referencia
Mapa de Retardos de Takens (TLM)	[118]
Estimación por Mínimos Cuadrados (LSE)	[119]
Descomposición Modal Dinámica (DMD)	[120]
Descomposición Modal Dinámica Extendida (EDMD)	[121]
Regresión por Procesos Gaussianos (GPR)	[122]
Regresión ARIMA con predicción hacia adelante	[92]
Regresión ARIMA con predicción hacia atrás	[92]
Modelo TBATS (Trigonometric Box-Cox, ARMA, Trend and Seasonal)	[123]
Búsqueda de Coincidencias No Lineales para TFRs Ralas (STF)	[124]
Procesos Ondita Localmente Estacionarios (LSW)	[125]
Suavizado No Lineal con Huecos Aleatorios y Valores Atípicos (DRAGO)	[126]

Tabla 4.1: Métodos de imputación inicial considerados en los experimentos.

siderados en nuestros experimentos, los cuáles se resumen en la Tabla 4.1. Esta lista no busca ser exhaustiva, aunque sí se han considerado diferentes métodos de imputación basados en enfoques diversos, yendo desde la búsqueda de patrones hasta predicción basada en modelos estadísticos.

4.2.1. Mapa de retardos de Takens

Antes de detallar el procedimiento de este método, recordemos que el teorema de *embedding* de Takens [118] establece que, en condiciones de suavidad, las propiedades del sistema dinámico subyacente que genera la serie temporal $y : \mathbb{Z} \rightarrow \mathbb{R}$ pueden representarse mediante una incrustación de los segmentos de la serie temporal con retardo. En concreto, se establece un vector con retardo de la forma

$$\mathbf{y}(k) = [y(k), y(k - \tau), \dots, y(k - (d_e - 1)\tau)]^\top \in \mathbb{R}^{d_e}, \quad (4.4)$$

donde $d_e \in \mathbb{N}$ es la dimensión de incrustación y $\tau > 0$ es el tiempo de retardo. Considerando condiciones de suavidad sobre la *manifold* subyacente que alberga la dinámica del sistema, el conjunto $\{\mathbf{y}(k)\}_k$ es difeomorfo a esta *manifold*. En este sentido, es posible recuperar la dinámica del sistema.

Esto nos permite definir un algoritmo de búsqueda de patrones, denominado mapa de retardos de Takens (TLM, por las siglas en inglés de *Takens' Lag Map*), a partir de la señal no estacionaria para poder imputar los datos faltantes. Para esto, primeramente fijamos un intervalo de datos faltantes con el índice inicial i_{ms} y la longitud de intervalo L . Seguidamente, se define el patrón de señal a la izquierda y a la derecha del intervalo de datos que faltan como

$$\mathbf{x}_l = [\mathbf{x}(i_{ms} - d), \dots, \mathbf{x}(i_{ms} - 1)], \quad (4.5)$$

$$\mathbf{x}_r = [\mathbf{x}(i_{ms} + L + 1), \dots, \mathbf{x}(i_{ms} + L + d - 1)]. \quad (4.6)$$

A continuación, se concatenan estos vectores para formar el patrón de referencia $\mathbf{x}_{tmp} = [\mathbf{x}_l \ \mathbf{x}_r]$.

Para estimar la porción faltante de la señal, se hace pasar una ventana deslizante de longitud $2d + L$ sobre la señal fuera del intervalo de datos faltantes, y se elige el patrón

4.2. MÉTODOS PARA IMPUTACIÓN DE DATOS FALTANTES EN SERIES TEMPORALES

de datos que mejor aproxima al patrón de referencia: $\mathbf{x}_p = [\mathbf{x}(p), \dots, \mathbf{x}(p + L - 1)]$. La calidad de esta aproximación se cuantifica a través de la distancia euclidiana entre \mathbf{x}_{tmp} y el vector $[\mathbf{x}_{p,l} \ \mathbf{x}_{p,r}]$, donde $\mathbf{x}_{p,l} = [\mathbf{x}(p-d), \dots, \mathbf{x}(p-1)]$ y $\mathbf{x}_{p,r} = [\mathbf{x}(p+L+1), \dots, \mathbf{x}(p+L+d-1)]$. A continuación, se repite el mismo procedimiento para todos los intervalos de valores faltantes.

Mediante este procedimiento, los datos imputados se eligen de forma que el comportamiento local de la señal alrededor de los valores candidatos sea similar al comportamiento alrededor del intervalo de datos faltantes. Desde el punto de vista técnico, establecemos $d_e = d$ y $\tau = 1$ para el *embedding* de Takens, y determinamos patrones “vecinos” en un subespacio proyectado. Aquí vale destacar que el algoritmo TLM es similar al método KNNimpute propuesto en [127] para microarreglos de ADN.

4.2.2. Estimación por mínimos cuadrados

La identificación de sistemas dinámicos no lineales es un enfoque habitual para la imputación de datos faltantes de series temporales. Generalmente, se ajusta un modelo paramétrico a los datos conocidos y, a continuación, se utiliza un procedimiento de predicción a partir del modelo para estimar los valores futuros de la serie temporal. En términos generales, el problema de identificar la dinámica de un sistema de la forma

$$\mathbf{Y} = f(\mathbf{X}), \quad (4.7)$$

implica estimar la función f que relaciona la entrada con la salida. Para la imputación de datos faltantes, consideramos el intervalo de datos faltantes como una serie temporal que debe predecirse a partir de los datos observados. Definimos la matriz $\mathbf{X} \in \mathbb{R}^{K \times M}$, que contiene los valores con retardo de la serie temporal, y la matriz $\mathbf{Y} \in \mathbb{R}^{K \times M}$ que contiene los valores correspondientes que deben predecirse, de la siguiente forma

$$\mathbf{X} = [\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{M-1}], \quad (4.8)$$

$$\mathbf{Y} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M], \quad (4.9)$$

donde $\mathbf{x}_l \in \mathbb{R}^K$ viene dado por $[x(i_{ms} - M - K + l), \dots, x(i_{ms} - M + l - 1)]$, para $1 \leq l \leq M$; es decir, cada columna de \mathbf{Y} se obtiene desplazando el elemento de la l -ésima columna de \mathbf{X} en una muestra. Una primera forma de estudiar este tipo de sistemas es considerando un modelo lineal de la forma:

$$\mathbf{Y} = \mathbf{A}\mathbf{X}, \quad (4.10)$$

donde la dinámica del sistema quedará descrita por la matriz $\mathbf{A} \in \mathbb{R}^{K \times K}$. A continuación, se ajusta un modelo predictivo a \mathbf{X} y \mathbf{Y} utilizando un algoritmo de estimación por mínimos cuadrados (LSE, del inglés *least-square estimation*)

$$\min_{\mathbf{A}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_2^2, \quad (4.11)$$

cuya solución se obtiene de aplicar las ecuaciones normales: $\tilde{\mathbf{A}} = \mathbf{Y}\mathbf{X}^T (\mathbf{X}\mathbf{X}^T)^{-1}$, si dicha matriz existe. Una vez ajustado el modelo de predicción, los valores faltantes se imputan estableciendo $\mathbf{x}_1(i_{ms} + m) = \tilde{\mathbf{A}}^m \mathbf{x}_K$, donde m es el índice de datos imputados.

4.2.3. Descomposición modal dinámica

El método denominado *dynamic mode decomposition* (DMD) es una técnica de identificación de sistemas que sigue la misma formulación que el método basado en LSE, pero en este caso la matriz \mathbf{A} se estima a partir de la descomposición en valores singulares de la matriz de datos \mathbf{X} :

$$\begin{aligned}\mathbf{Y} &= \mathbf{A}\mathbf{X} \\ \mathbf{X} &= \mathbf{U}\Sigma\mathbf{V}^T \\ \tilde{\mathbf{A}} &= \mathbf{U}^T\mathbf{Y}\mathbf{V}\Sigma^{-1}.\end{aligned}$$

Aquí, la matriz $\Sigma \in \mathbb{R}^{K \times M}$ es una matriz diagonal rectangular cuyas entradas diagonales $\Sigma_{ii} = \sigma_i$ son los valores singulares de la matriz de datos \mathbf{X} . Las matrices \mathbf{U} y \mathbf{V} son matrices cuadradas de tamaño $K \times K$ y $M \times M$, respectivamente, cuyas columnas forman una base ortonormal en el espacio de dimensión correspondiente.

Al igual que con el método de estimación por mínimos cuadrados, la imputación de los valores faltantes se realiza mediante la fórmula de predicción $\mathbf{x}_1(i_{ms} + m) = \tilde{\mathbf{A}}^m \mathbf{x}_K$.

4.2.4. Descomposición modal dinámica extendida

Formalmente, este método busca estimar el operador de evolución no lineal, denotado como \mathbf{F} , de un sistema dinámico en tiempo discreto: $\mathbf{x}(n+1) = \mathbf{F}(\mathbf{x}(n))$. Para lograrlo, el enfoque de descomposición modal dinámica extendida (EDMD, del inglés *Extended DMD*) aproxima el operador de Koopman \mathcal{K} , el cual es un operador *lineal* de dimensión infinita. Este operador gobierna la evolución de los *observables* del sistema. Estos observables son funciones definidas en el espacio de estados que definen la evolución temporal del sistema. Mientras que el operador de Koopman proporciona una linealización *global* del sistema dinámico, LSE y DMD realizan una linealización *local*.

En [121], los autores proponen un algoritmo guiado por los datos para aproximar los valores propios $\{\mu_j\}$ y los modos $\{\mathbf{v}_j\}$ de la estimación del operador de Koopman de dimensión finita denotado como $\hat{\mathbf{K}}$. En este trabajo, se demuestra que este método es una extensión de DMD, mejorando las estimaciones de los modos del operador de Koopman.

Además, en el estudio realizado por Hua y cols. [128], se introduce un algoritmo novedoso para estimar eficientemente los modos. Este algoritmo utiliza una función núcleo o *kernel* para proyectar los datos en un espacio de características de alta dimensión, lo que mejora la aproximación del operador de Koopman. Basándose en este procedimiento, los datos que faltan se imputan mediante la predicción en el intervalo de datos que faltan mediante

$$\mathbf{x}(n+1) = \sum_j \mu_k \mathbf{v}_j \phi_j(\mathbf{x}(n)), \quad (4.12)$$

donde $\{\phi_j(\mathbf{x})\}$ es un diccionario de funciones propias que mapean los observables \mathbf{x} en el espacio de estados. Al igual que con \mathcal{K} , las funciones propias de dimensión infinita se aproximan mediante la combinación lineal de un conjunto de funciones de dimensión finita $\psi_j(\mathbf{x})$ definidas en el espacio de características de dimensión J que, en general, es mucho mayor que K y M . Con el fin de evitar tener que resolver un problema de alta dimensión, las funciones $\psi(\mathbf{x})$ se definen implícitamente mediante el uso de un kernel

$\mathbf{k}(\mathbf{x}_i, \mathbf{x}_j) = \langle \boldsymbol{\psi}(\mathbf{x}_i), \boldsymbol{\psi}(\mathbf{x}_j) \rangle$, que es el producto interno entre dos funciones propias J -dimensionales. En nuestra implementación, se utiliza un kernel gaussiano $\mathbf{k}(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|_2^2/\gamma)$, donde el tamaño del kernel γ es elegido por el usuario.

4.2.5. Regresión por procesos gaussianos

La regresión mediante procesos gaussianos (GPR, del inglés *Gaussian process regression*) [129] es una técnica de aprendizaje automático utilizada para tareas de regresión. Esta técnica modela la relación entre las variables de entrada (características \mathbf{X}) y las variables de salida (respuestas \mathbf{Y}) de forma probabilística mediante $\mathbf{Y} = f(\mathbf{X})$. A diferencia de los métodos de identificación de sistema dinámicos, GPR estima la función f como una distribución de probabilidad sobre los posibles valores objetivo para una entrada dada:

$$f(\mathbf{X}) \sim \text{GP}(0, k(\mathbf{x}, \mathbf{x}')), \quad (4.13)$$

donde la función f se modeliza mediante un proceso gaussiano con media cero y función de covarianza $k(\mathbf{x}, \mathbf{x}')$. En la aplicación de imputación, se ajusta un GP utilizando los datos disponibles, lo que permite al modelo entrenado estimar los valores del intervalo faltante de una serie temporal basándose en las distribuciones de probabilidad a posteriori.

$$P(\mathbf{y}|f, \mathbf{X}) \sim \mathcal{N}(\mathbf{y}|f, \sigma^2 I), \quad (4.14)$$

donde σ^2 es la varianza del ruido contenido en la señal, el cual suponemos gaussiano de media cero. Para implementar el método de imputación GPR, utilizamos la función `fitrgp` de la *Statistical and Machine Learning Toolbox* de MATLAB®. Para entrenar el modelo, utilizamos la matriz \mathbf{X} definida en (4.8) y utilizamos la matriz \mathbf{Y} como objetivo. Este enfoque realiza la estimación dinámica del modelo utilizando un método no paramétrico, que puede ser beneficioso en los casos en que el sistema subyacente sea complejo y difícil de modelar paramétricamente.

4.2.6. Regresión ARIMA

ARIMA (del inglés, *Autoregressive Integrated Moving Average*) es una técnica de modelización de series temporales ampliamente adoptada y empleada para predecir valores futuros en series temporales no estacionarias [130]. Además, para capturar componentes periódicos dentro de la señal, se puede recurrir a los modelos ARIMA estacionales (SARIMA) [130]. El proceso de modelización implica la identificación de patrones de datos subyacentes mediante la evaluación de las funciones de autocorrelación y autocorrelación parcial. Una vez identificados estos patrones, el modelo se convierte en una valiosa herramienta para predecir valores futuros. En el contexto de la imputación de datos faltantes, el método consiste en ajustar el modelo ARIMA a los datos que preceden al intervalo de datos faltantes y, a continuación, imputar los datos que faltan utilizando el modelo ajustado mediante un proceso de predicción hacia adelante (*forward* ARIMA). Uno de los parámetros fundamentales de este método es el periodo de estacionalidad, denominado λ .

La predicción mediante el método ARIMA también puede realizarse utilizando un paso de predicción hacia atrás (*backward* ARIMA). Para esto, ajustamos el modelo ARIMA a la versión invertida de los datos que se encuentran a la derecha del intervalo faltante:

$$\mathbf{x}_{pos} = [x(N), x(N-1), \dots, x(i_{ms} + L + 1)]. \quad (4.15)$$

A continuación, el intervalo de datos que faltan se predice a partir del modelo ARIMA ajustado y se invierte para obtener los datos imputados. Al igual que en el método anterior, la periodicidad λ debe establecerse de antemano.

4.2.7. Modelo TBATS

El modelo TBATS (por las siglas en inglés de *trigonometric, Box-Cox transformation, ARMA errors, trend and seasonality*) es un modelo de espacio de estados diseñado para el análisis de series temporales [123]. Incorpora un suavizado exponencial para dar más peso a los datos recientes durante la estimación. La tendencia y los componentes estacionales también se suavizan mediante exponenciales para mejorar la estimación de los datos locales. Además, se aplica la transformación Box-Cox para normalizar los datos, mejorando sus características estadísticas. Se incluyen errores ARMA para captar la dinámica compleja de los datos. De forma similar a los modelos ARIMA, el modelo TBATS puede emplearse para la imputación de datos faltantes ajustando el modelo a los datos anteriores al intervalo faltante con un posterior paso de predicción de los valores faltantes. En [123], se propone un algoritmo eficiente para estimar los parámetros del modelo mediante la maximización de la función de verosimilitud de su representación en el espacio de estados.

4.2.8. Búsqueda de coincidencias no lineales para TFRs ralas

En [124], los autores proponen un método de descomposición de señales no estacionarias basado en un algoritmo de búsqueda de coincidencias (*matching pursuit*) a través de la resolución de un problema de optimización de la forma

$$\operatorname{argmín}_{a, \theta} \left\| f - \sum_{k=1}^K a_k(t) \cos \theta_k(t) \right\|_2^2, \quad (4.16)$$

donde cada componente se define como una función armónica modulada en fase y amplitud $a_k(t) \cos(\theta_k(t))$. Este método busca activamente funciones óptimas de amplitud y fase utilizando un diccionario de Fourier sobrecompleto, a la vez que promueve la rareza mediante la regularización de la norma-1 de los coeficientes de amplitud:

$$(a_k(t), \theta_k(t)) \in \operatorname{argmín}_{a, \theta} \gamma \|\hat{a}_k\|_1 + \left\| f - \sum_{k=1}^K a_k(t) \cos \theta_k(t) \right\|_2, \quad (4.17)$$

siendo \hat{a}_k los coeficientes de la descomposición en serie de Fourier de la modulación en amplitud $a_k(t)$. Para tratar señales con datos faltantes, se emplea un método de optimización de punto interior basado en gradientes conjugados preconditionados [131]. Este algoritmo requiere del preestablecimiento de parámetros, incluyendo un estimador inicial de la fase θ_0 y el número de componentes armónicos denotados como K .

4.2.9. Onditas localmente estacionarias

El modelo de onditas localmente estacionarias (LSW, por las siglas en inglés de *locally stationary wavelets*), propuesto por Nason y cols. [132], considera una serie temporal como una realización de un proceso estocástico $X_{t,T,t=0,1,\dots,T-1}$, donde $T = 2^J \geq 1$. Este modelo asume la siguiente representación para la señal:

$$X_{t,T} = \sum_{j=1}^J \sum_k w_{j,k;T} \psi_{j,k}(t) \zeta_{j,k}. \quad (4.18)$$

Aquí, $\psi_{j,k}(t)$ representa una familia de onditas discretas no decimadas donde la ondita madre $\psi(t)$ tiene soporte compacto, $\zeta_{j,k}$ es una secuencia aleatoria de incrementos ortonormales, y $w_{j,k;T}$ son los coeficientes de la transformada ondita.

Dentro del marco LSW, un conjunto de observaciones $X_{0,T}, \dots, X_{t-1,T}$ puede utilizarse para predecir la siguiente observación $X_{t,T}$ mediante una combinación lineal de las p observaciones más recientes, como se propone en [133]

$$\hat{X}_{t,T} = \sum_{s=t-p}^{t-1} b_{t-1-s,T} X_{s,T}. \quad (4.19)$$

El vector de coeficientes \mathbf{b} se obtiene al minimizar el error cuadrático medio de predicción (MSPE), lo que conduce a un conjunto de ecuaciones de Yule-Walker generalizadas. Los autores de [133] demuestran que la solución a las ecuaciones de Yule-Walker puede aproximarse minimizando

$$\mathbf{b}_{t,T}^T \mathbf{B}_{t,T} \mathbf{b}_{t,T}, \quad (4.20)$$

donde T indica la transposición y $\mathbf{B}_{t,T}$ es una matriz $(t+1) \times (t+1)$ que aproxima la matriz de covarianza de $X_{0,T}, \dots, X_{t-1,T}$. Según la teoría de procesos ondita, la matriz $\mathbf{B}_{t,T}$ puede calcularse a partir de la matriz de autocovarianza local (LACV) de $X_{t,T}$, que se obtiene a de aplicar un suavizado al periodograma ondita

$$\mathbf{I}_{j,k;T} = |d_{j,k;T}|^2, \quad (4.21)$$

siendo $d_{j,k;T}$ los coeficientes de la transformada ondita empírica de $X_{t,T}$.

4.2.10. Suavizado no lineal de datos con huecos aleatorios y valores atípicos

Este algoritmo, denominado DRAGO (por las siglas en inglés de *Data with Random Gaps and Outliers*), fue propuesto originalmente por Parekh y cols. [126] para estudiar las mediciones de la temperatura corporal central en presencia de segmentos de datos faltantes y valores atípicos. Para aplicar este método, se propuso el siguiente modelo de señal

$$\mathbf{S}\mathbf{y} = \mathbf{S}\mathbf{f} + \mathbf{x} + \mathbf{w}, \quad (4.22)$$

donde \mathbf{f} es una señal discretizada con cierta suavidad, \mathbf{w} es ruido gaussiano blanco aditivo, \mathbf{x} es una señal (rala) que representa a los valores atípicos y \mathbf{S} se denomina matriz de

4.3. SELECCIÓN DE LOS PARÁMETROS DE IMPUTACIÓN

muestreo. La matriz de muestreo indica las posiciones de las muestras faltantes. Para estimar la señal completa \mathbf{f} , se propone resolver el siguiente problema de optimización regularizado:

$$\{\hat{\mathbf{f}}, \hat{\mathbf{x}}\} = \arg \min_{\mathbf{f}, \mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{S}\mathbf{y} - \mathbf{S}\mathbf{f} - \mathbf{x}\|_2^2 + \frac{\lambda_1}{2} \|\mathbf{x}\|_1 + \frac{\lambda_2}{2} \|\mathbf{D}\mathbf{f}\|_2^2 \right\}, \quad (4.23)$$

donde \mathbf{D} es la matriz de diferencias de segundo orden, λ_1 y λ_2 son parámetros de regularización, y $\|\cdot\|_1$ indica la norma-1 del vector. Cuando se considera una señal sin valores atípicos ($\mathbf{x} = 0$), este problema se reduce a un operador de mínimos cuadrados ponderados suavizado. Para resolver (4.23), se propone un procedimiento de maximización-minimización (MM). Este enfoque MM resuelve una secuencia de problemas de optimización que son más fáciles de resolver que la minimización (4.23). Dado que la función objetivo es convexa, se puede alcanzar un mínimo global.

Este procedimiento MM se utiliza para implementar el algoritmo iterativo DRAGO, que puede resumirse en las siguientes operaciones para cada paso de iteración i :

$$\mathbf{A}^{(i)}(n, n) = \frac{\lambda_1}{|\mathbf{x}(n)^{(i)} + \lambda_1|}, \quad (4.24)$$

$$\mathbf{f}^{(i+1)} = \left(\mathbf{S}^T \mathbf{A}^{(i)} \mathbf{S} + \lambda_2 \mathbf{D}^T \mathbf{D} \right)^{-1} \mathbf{S}^T \mathbf{A}^{(i)} \mathbf{S} \mathbf{y}, \quad (4.25)$$

$$\mathbf{x}(n)^{(i+1)} = \frac{|\mathbf{x}(n)^{(i)}|}{|\mathbf{x}(n)^{(i)}| + \lambda_1} \mathbf{r}^{(i+1)}(n), \quad (4.26)$$

donde $\mathbf{A}^{(i)} \in \mathbb{R}^{N \times N}$ es una matriz diagonal y $\mathbf{r}^{(i+1)} = \mathbf{S}(\mathbf{y} - \mathbf{f}^{(i+1)}) \in \mathbb{R}^{N \times 1}$ para $n = 1, \dots, N$. Este procedimiento iterativo se realiza hasta la convergencia. En este algoritmo, los principales parámetros que afectan al rendimiento de la imputación de datos faltantes son los parámetros de regularización λ_1 y λ_2 .

4.3. Selección de los parámetros de imputación

La selección adecuada de los parámetros en los algoritmos descritos es un aspecto esencial y puede afectar en gran medida al rendimiento de los algoritmos. Normalmente, a menos que se disponga de un método automático, los parámetros se ajustan manualmente para lograr resultados satisfactorios. Para los algoritmos que no disponen de un procedimiento estándar de selección de parámetros, proporcionamos criterios empíricos basados en la práctica que pueden aplicarse a una amplia gama de series temporales oscilatorias, sin indagar en el fundamento teórico de dicha selección.

4.3.1. Mapa de retardos de Takens

Para el algoritmo TLM se debe establecer de antemano la longitud de la plantilla de comparación d . Esta longitud define el patrón que se va a comparar con el resto de la señal y puede influir significativamente en el resultado final. Dado que nos centramos en el estudio de señales oscilatorias, proponemos una forma intuitiva de definir d basada en la cuasi-periodicidad de la señal. Primeramente, definimos el periodo medio \tilde{T} como $\tilde{T} = 2\pi/f_p$, donde f_p es la frecuencia más energética de la señal que puede estimarse a partir

4.3. SELECCIÓN DE LOS PARÁMETROS DE IMPUTACIÓN

del espectro de la señal. A partir de \tilde{T} , definimos la longitud del patrón como $d = k\tilde{T}$ con $k \in \mathbb{N}$. Encontramos empíricamente que $k = 3$ conduce a resultados satisfactorios en la mayoría de las señales. Obsérvese que la naturaleza de variación lenta de la frecuencia instantánea significa que la señal puede aproximarse localmente mediante una función estacionaria con una pérdida mínima de precisión.

4.3.2. Estimación por mínimos cuadrados, DMD y EDMD

Debido a su implementación, los algoritmos LSE, DMD y EDMD incorporan dos parámetros ajustables: la dimensión del espacio de incrustación denotado como K , que está relacionado con el desplazamiento máximo entre las subseñales, y la longitud de las subseñales definida como M , la cual es utilizada en la construcción de la matriz del sistema \mathbf{A} . Cuando se trata de señales no estacionarias, es esencial establecer M para que sea mayor que el período medio de la señal para capturar la variabilidad interciclo de manera efectiva. Se recomienda seleccionar M de forma que abarque al menos tres ciclos completos de la señal dentro de cada subseñal.

La dimensión de incrustación K debe elegirse de forma que sea mayor que M . Los hallazgos teóricos en [119] ilustran que a medida que K se aproxima al infinito, la varianza del resultado de predicción disminuye asintóticamente, pero el sesgo aumenta. A partir de un análisis preliminar, se determinó que $K = 2,5M$ es un valor adecuado para la mayoría de los casos considerados en nuestros experimentos.

4.3.3. Regresión por procesos gaussianos

Al igual que los métodos anteriores, la regresión por procesos gaussianos predice una matriz de observaciones \mathbf{Y} a partir de una matriz de datos \mathbf{X} . Por lo tanto, también se deben establecer los parámetros constructivos de estas matrices, es decir, las dimensiones K y M . Estos parámetros se definen de manera análoga a los métodos LSE, DMD y EDMD.

Además, para el algoritmo GPR es necesario establecer la estructura de la matriz de covarianza $k(\mathbf{x}, \mathbf{x}')$. Existen diferentes tipos de matrices de covarianza o *kernels* que se usan frecuentemente en la práctica, como los kernels exponenciales o tipo Matern [134]. Este último tipo de funciones permite favorecer condiciones de suavidad para las curvas estimadas a partir del regresor. Sin embargo, al no poder conocer previamente las condiciones de suavidad de las señales a estimar y, por simplicidad, optamos por trabajar con matrices de covarianza de forma gaussiana (exponenciales al cuadrado).

4.3.4. Regresión ARIMA

Para ajustar un modelo SARIMA a los datos, ya sea antes o después del intervalo faltante para una buena imputación, es crucial estimar la periodicidad (o estacionalidad) λ de la señal debido a la naturaleza oscilatoria de las señales de interés. Sin embargo, hasta donde sabemos, no existe un enfoque adecuado para estimar la estacionalidad de las señales con frecuencia variable en el tiempo para SARIMA. Por lo tanto, consideramos el siguiente enfoque empírico. Primero estimamos la duración de cada ciclo de la señal como

$$T_k = \#\{\hat{\phi}_1(n) : k - 1 < \hat{\phi}_1(n) \leq k\}, \quad (4.27)$$

4.3. SELECCIÓN DE LOS PARÁMETROS DE IMPUTACIÓN

donde $\hat{\phi}_1$ es la fase estimada y $k \in \mathbb{N}$. Dado que la estacionalidad se supone fija en SARIMA, estimamos la estacionalidad para el método de regresión SARIMA hacia adelante por

$$\lambda_f = \frac{1}{N_c} \sum_{i=a-N_c}^{i=a} T_i, \quad (4.28)$$

donde a denota el ciclo inmediatamente anterior al intervalo de datos faltantes. Del mismo modo, para el método de regresión SARIMA hacia atrás, estimamos la estacionalidad por

$$\lambda_b = \frac{1}{N_c} \sum_{i=b}^{i=b+N_c} T_i, \quad (4.29)$$

donde b denota el primer ciclo inmediatamente posterior al intervalo de datos faltantes.

Tanto para la predicción hacia adelante o hacia atrás, N_c debe fijarse de antemano. Debido a la naturaleza variable en el tiempo del modelo considerado, son preferibles las estimaciones locales, ya que captan con mayor precisión la dinámica de los datos en las proximidades del intervalo de datos faltantes, pero N_c no debe ser demasiado pequeño para evitar una varianza elevada en la estimación. Como pauta general, una elección empírica $N_c = 3$ suele conducir a una buena estimación de los datos en el intervalo de datos faltantes. Otros parámetros, incluidos los órdenes del modelo y los parámetros, podrían estimarse siguiendo el procedimiento estándar para modelos estadísticos.

4.3.5. Modelo TBATS

El procedimiento general para obtener el modelo de espacio de estados adecuado para nuestros datos implica dos pasos principales: estimación de parámetros y selección de modelos. En el primer paso, se determinan los parámetros óptimos maximizando la verosimilitud de los datos dado un modelo específico. En el segundo paso, se selecciona el mejor modelo utilizando un criterio de selección. Para el modelo TBATS, se prueban varias configuraciones, como la aplicación de la transformación Box-Cox o la inclusión de errores ARMA en el modelo. En el contexto de las señales estudiadas en este trabajo, se presta especial atención a la periodicidad de los componentes estacionales y al parámetro de suavizado de la tendencia debido a la naturaleza variante en el tiempo de la frecuencia. La estimación de la periodicidad del componente fundamental, vinculado al componente estacional más largo del modelo, es crucial para garantizar una predicción precisa de los datos que faltan. Para tener en cuenta la variación temporal de la frecuencia, seguimos el método de estimación empírica utilizado para los modelos SARIMA analizados en la sección anterior.

4.3.6. Búsqueda de coincidencia no lineal para RTF ralas

En este método, es necesario proveer de una estimación inicial de la fase $\theta_0(t)$ y fijar el número de armónicos K para la descomposición. Respecto a la fase inicial $\theta_0(t)$, se define la función lineal $\theta_0(t) = 2\pi f_p t$, donde f_p se elige como la frecuencia de máxima energía del espectro de $x(t)$. El número de componentes armónicos K se estima aplicando los criterios de selección de modelos trigonométricos para el modelo no armónico adaptativo ajustado a los datos, como se describe en el Cap. 2.

4.3.7. Procesos ondita localmente estacionarios

El número p de índices pasados utilizados para estimar los coeficientes de predicción se estima automáticamente utilizando la función de autocorrelación parcial local (LPACF) propuesta en [135]. Los autores de [136] estudiaron el comportamiento de diferentes onditas madre y descubrieron que la ondita de Daubechies de fase extrema y orden 8 funciona adecuadamente en la mayoría de los casos. Además, para calcular $\mathbf{B}_{t,T}$, la matriz $(t+1) \times (t+1)$ que aproxima la matriz de covarianza de $X_{0,T}, \dots, X_{t-1,T}$ utilizada para encontrar el minimizador del MSPE, es necesario suavizar el periodograma ondita. El ancho de banda del periodograma suavizado se estima utilizando el procedimiento automático propuesto por Nason [137].

4.3.8. Suavizado no lineal de datos con huecos aleatorios y valores atípicos

Para este procedimiento, se deben fijar de antemano los parámetros de regularización λ_1 y λ_2 . Con respecto a λ_1 , los autores sugieren establecer el parámetro de regularización en un múltiplo de la desviación estándar del ruido σ : $\lambda_1 = c\sigma$, con $c \in 2, 3$. De este modo, este parámetro de regularización cumple la función de un umbral suave para el ruido blanco gaussiano que contamina la señal [138]. Con respecto a λ_2 , los autores en [138] muestran que cuando se considera el problema de optimización sin huecos ni valores atípicos, la solución puede interpretarse como un filtro lineal invariante en el tiempo (LTI) con respuesta en frecuencia

$$H(\omega) = \frac{1}{1 + 16\lambda_2 \sin^4(\omega/2)}. \quad (4.30)$$

La ecuación anterior puede resolverse para λ_2 y el valor del parámetro de regularización puede elegirse de modo que la banda de paso del filtro LTI contenga en su totalidad al espectro de la señal \mathbf{f} .

4.4. Algoritmo propuesto - HaLI

En esta sección, describiremos el algoritmo propuesto para la imputación de datos faltantes, el cuál denominamos como HaLI. En primer lugar, este algoritmo se basa en considerar a la señal no estacionaria según la siguiente variante del ANHM

$$x(t) = \sum_{k=1}^K \sum_{\ell=1}^{r_k} B_{k,\ell}(t) \cos(2\pi\phi_{k,\ell}(t)) + T(t) + n(t). \quad (4.31)$$

En este modelo, la señal se compone de K componentes o modos, cada uno de ellos con una función de forma de onda variable en el tiempo. Para el k -ésimo componente, la variabilidad de la forma de onda queda descrita mediante las funciones de amplitud $B_{k,\ell}(t)$ y frecuencia $\phi'_{k,\ell}(t)$, las cuales deben cumplir con las condiciones de variabilidad lenta de la forma de onda (C3-C5), mencionadas en la Sec. 3.1. Nótese que cada modo tendrá asociado un *orden armónico* r_k que deberá determinarse a partir de la señal. Además, la función $T(t)$ define la tendencia de la señal, la cual debe ser considerada ya que dicha tendencia acarrea información de interés en diferentes entornos biomédicos. Por ejemplo,

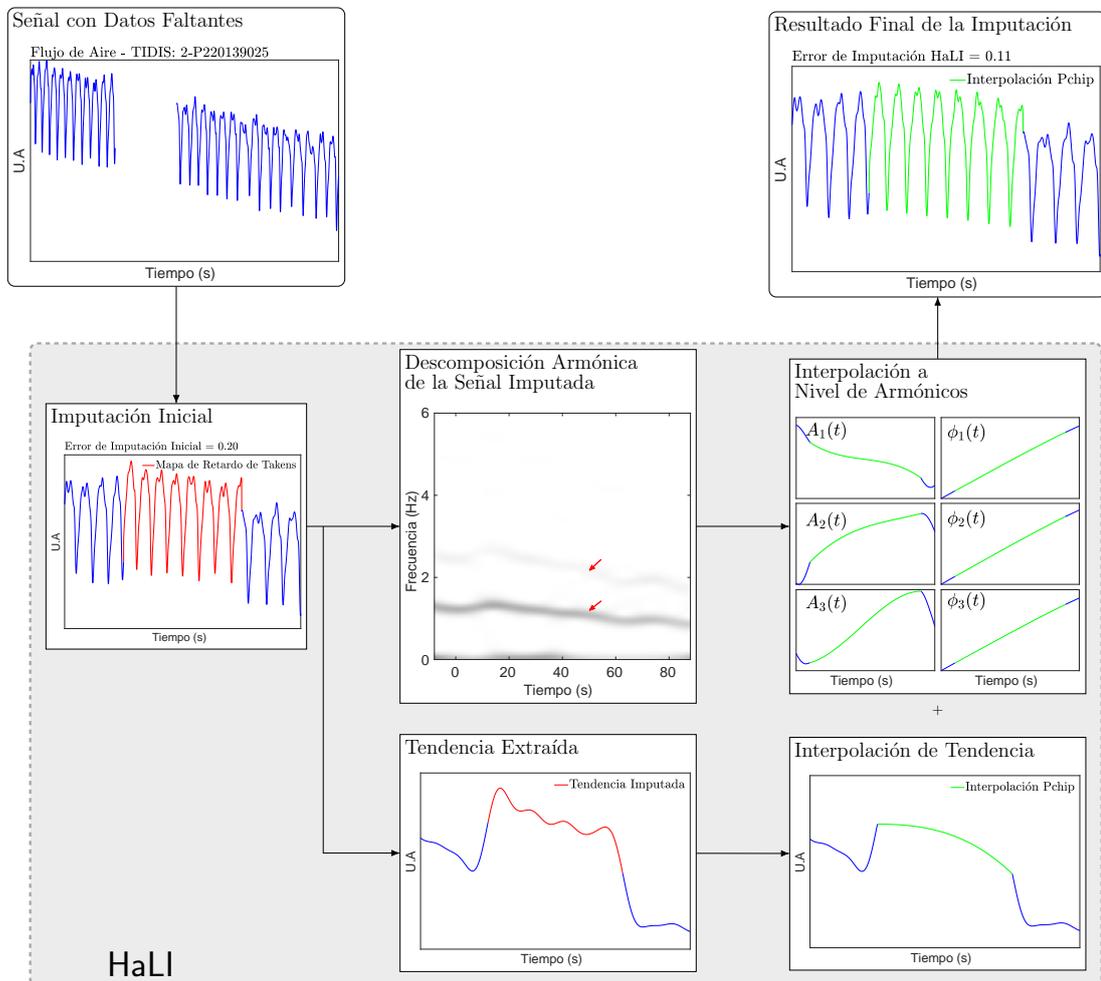


Figura 4.2: Método propuesto para imputación de datos faltantes basado en la interpolación de la descomposición armónica de señales no estacionarias. La primer etapa consiste en la aplicación de un método de imputación inicial para obtener un resultado preliminar. Luego, se realiza la descomposición armónica de la señal imputada y se extrae la tendencia. En la tercer etapa, se realiza la interpolación a nivel de las amplitudes y fases armónicas, y de la tendencia. La tendencia y los componentes armónicos interpolados se combinan para obtener como resultado final de algoritmo una versión mejorada de la señal con datos imputados en el intervalo faltante.

la presión arterial media (MAP, del inglés *mean arterial pressure*) obtenida como la media local (a nivel de cada ciclo) de la señal de presión sanguínea arterial (ABP, del inglés *arterial blood pressure*). En general, podemos concebir a la tendencia como una señal cuya variación es más lenta que la de todos los componentes armónicos, por lo cual el ancho de banda de $T(t)$ tendrá como cota superior: $\phi'_m(t) = \min_k \phi'_k(t)$ para todo t . Por último, el término $n(t)$ define el ruido estocástico presente en la señal, el cual suponemos gaussiano y de media cero.

El algoritmo HaLI consta de tres etapas principales. En primer lugar, se realiza una imputación inicial utilizando un método ya existente. En segundo lugar, se obtienen las amplitudes y fases armónicas descomponiendo la señal imputada a partir de una TFR. Por último, la imputación refinada se genera interpolando las amplitudes y fases armónicas en los intervalos de datos faltantes. Véase en la Fig. 4.2 un diagrama de flujo general de HaLI. Aquí notamos que HaLI sigue una idea similar a algoritmos como MissForest [139] o MISE [140] para datos genéricos, donde la imputación inicial implica el relleno de la media o la mediana, seguido de una imputación iterativa basada en aprendizaje automático. HaLI, en cambio, fue diseñado para tratar específicamente con series temporales oscilatorias no estacionarias.

Antes de profundizar en los detalles, es esencial abordar algunas consideraciones técnicas. La imputación inicial es de suma importancia para el HaLI. Este paso es fundamental porque la descomposición armónica es sensible a los efectos de borde inherentes a las herramientas de análisis TF. Concretamente, para la señal $f(t)(1 - \chi_{I_{ms}}(t))$ con datos faltantes en x , debido a la discontinuidad e irregularidad en los bordes de I_{ms} , la TFR se verá “borrosa” en esta región, lo cual degrada la descomposición armónica.

Aunque existe la posibilidad de no realizar la imputación inicial si se dispusiera de un algoritmo de descomposición armónica que sea inmune a los efectos de borde, hasta donde sabemos no existe una herramienta con esta capacidad. Al emplear una imputación inicial, se alivia el efecto en los bordes de los intervalos de datos faltantes y es posible lograr una descomposición armónica más adecuada, especialmente cerca de los límites de I_{ms} . No obstante, los armónicos sobre I_{ms} dependen en gran medida de esta imputación inicial, la cual puede no ser del todo acertada. La principal innovación del algoritmo propuesto consiste en mejorar la calidad de la imputación sobre I_{ms} interpolando adecuadamente los armónicos desde $\mathbb{R} \setminus I_{ms}$ hacia I_{ms} . Los pasos del algoritmo se discuten con mayor detalle en las siguientes secciones.

4.4.1. Etapa 1: Imputación inicial de datos

El primer paso de HaLI implica aplicar una técnica de imputación sobre la señal con datos faltantes. Este paso toma la señal incompleta x y los intervalos de valores faltantes como entrada. Como ya mencionamos antes, cada algoritmo contempla un conjunto de parámetros que deben ajustarse de antemano para garantizar un buen resultado en la tarea de imputación. Denotemos la señal resultante con valores imputados como $\mathbf{x}_1 \in \mathbb{R}^N$.

Entre los algoritmos considerados, destacamos el algoritmo TLM, el cual supera sistemáticamente a otros, como mostraremos en las Sec. 4.5 y 4.6. Este método tiene como objetivo localizar un segmento en la señal, sin datos faltantes, que se acerque a los “datos esperados” en el intervalo de datos faltantes. La elección de una medida de similitud adecuada es crucial para obtener resultados eficaces, ya que permite identificar un segmento con valores que pueden utilizarse para imputar los datos que faltan.

4.4.2. Etapa 2: Separación de la tendencia y descomposición armónica de la señal imputada

Una vez obtenido el resultado de imputación inicial \mathbf{x}_1 , aplicamos un algoritmo de descomposición armónica para descomponer \mathbf{x}_1 en su tendencia y armónicos. Existen diversos algoritmos de descomposición armónica. Sin embargo, para no desviarnos del objetivo de este algoritmo, aplicamos el algoritmo basado en la transformada de Fourier de tiempo corto (STFT). Para esto, primero se obtiene la STFT $\mathbf{F}_{\mathbf{x}_1}^{\mathbf{g}}$, como se describe en el Algo. 1 de la Sec. 2.3, utilizando una ventana gaussiana $\mathbf{g} = e^{-\sigma n^2}$. El ancho de la ventana se elige siguiendo la regla empírica de que el soporte temporal de \mathbf{g} debe contener entre 5 y 8 ciclos del componente oscilatorio de interés para obtener una resolución tiempo-frecuencia adecuada. La longitud media del ciclo se estima localmente a partir del período medio \tilde{T} de \mathbf{x}_1 .

A partir de \mathbf{F} , hallamos la cresta fundamental asociada a cada modo de la señal. Para ello, calculamos la de-shape STFT \mathbf{W} , la cual nos permite obtener una TFR donde se atenúa la amplitud de los armónicos y resulta más fácil detectar de la frecuencia instantánea del componente fundamental. Encontramos las crestas de $|\mathbf{W}|^2$ utilizando el algoritmo voraz de detección de crestas descrito en la Sec. 1.3.3. Como resultado se obtiene el conjunto de crestas fundamentales $\{\mathbf{c}_k^*(n)\}_{k=1}^K$, donde $\mathbf{c}_k^*(n) \approx \phi_{k,1}(n)$. Debido a la suposición de continuidad de las frecuencias instantáneas, y por tanto de las crestas asociadas, este algoritmo limita la región de búsqueda para cada paso temporal a una banda de frecuencias $\text{FB}_{n,k} = [\mathbf{c}_k(n-1) - \text{FB}, \mathbf{c}_k(n-1) + \text{FB}]$ alrededor de la estimación de la cresta en el instante de tiempo anterior, donde $\text{FB} > 0$ define el salto de frecuencia máximo para pasos consecutivos. Para este algoritmo, el salto de frecuencia máximo FB se establece en $10f_s/N$.

Basándonos en las crestas fundamentales estimadas $\{\mathbf{c}_k^*(n)\}_{k=1}^K$, realizamos la eliminación de la tendencia de la señal restando la tendencia estimada $\hat{\mathbf{T}}_1$ de \mathbf{x}_1 , donde

$$\hat{\mathbf{T}}_1(n) = \frac{2}{g(0)} \text{Re} \left\{ \sum_{1 \leq j < \underline{\mathbf{c}}^* - \Delta} \mathbf{F}(n, j) \right\}. \quad (4.32)$$

Aquí, $\underline{\mathbf{c}}^* = \min_k \mathbf{c}_k^*(n)$ y Δ se elige cerca de la medida del semi-soporte de $\hat{\mathbf{g}}(k) = \mathcal{F}(\mathbf{g}(n))$, como explicamos en la Sec. 1.3.3.

Para la descomposición armónica, recordamos que el componente fundamental complejo de la k -ésima función de forma de onda de \mathbf{x}_1 puede recuperarse integrando la STFT alrededor de la cresta \mathbf{c}_k^* :

$$\mathbf{y}_{k,1}(n) = \frac{1}{\mathbf{g}(0)} \sum_{|j - \mathbf{c}_k^*(n)| < \Delta} \mathbf{F}(n, j), \quad (4.33)$$

donde $n = 1, \dots, N$. A continuación, la amplitud fundamental se estima como $\tilde{\mathbf{B}}_{k,1}(n) = |\mathbf{y}_{k,1}(n)|$ y la fase asociada, denotada como $\tilde{\phi}_{k,1}$, se estima desarrollando la fase de $\mathbf{y}_{k,1}$.

Una vez obtenido el armónico fundamental de cada modo, extraemos los armónicos de orden superior. Para ello, obtenemos la estimación de r_k , denotada como r_k^* , para cada modo utilizando algún criterio de selección de modelos basado en regresión trigonométrica. Tras determinar r_k^* , ahora debemos hallar la cresta del ℓ -ésimo armónico del k -ésimo

componente $\mathbf{c}_{k,\ell}(n)$ para $2 \leq \ell \leq r_k^*$. Dada la condición (C3) descrita en la Sec. 3.1, sabemos que $\mathbf{c}_{k,\ell}(n) \approx \ell \mathbf{c}_k(n)$. La cresta armónica $\mathbf{c}_{k,\ell}^*(n)$ se encuentra entonces mediante una maximización de la energía de la cresta en las proximidades de $\ell \mathbf{c}_k^*(n)$.

Por último, la modulación de amplitud del ℓ -ésimo armónico se estima como $\tilde{\mathbf{B}}_{k,\ell}(n) = |\mathbf{y}_{k,\ell}(n)|$ para $n = 1, \dots, N$ y su fase, denotada como $\tilde{\phi}_{k,\ell}$, se estima desarrollando la fase de $\mathbf{y}_{k,\ell}$, donde $\mathbf{y}_{k,\ell}(n)$ es el ℓ -ésimo componente complejo de el k -ésimo modo de \mathbf{x}_1 obtenido usando (4.33) y sustituyendo $\mathbf{c}_k^*(n)$ por $\mathbf{c}_{k,\ell}^*(n)$.

Obsérvese que podríamos emplear algoritmos avanzados de análisis de TF como la transformada *synchrosqueezing* [42], algoritmos más recientes de detección de crestas [141, 142] o fórmulas más sofisticadas de reconstrucción basadas en fase. Sin embargo, en aras de la simplicidad del algoritmo, la eficiencia computacional y la precisión, optamos por seguir este enfoque simplificado.

4.4.3. Etapa 3: Interpolación a nivel de los armónicos

En esta etapa se utiliza la descomposición armónica de \mathbf{x}_1 para mejorar la imputación mediante un recorte de la amplitud y la fase de los armónicos, seguido de la interpolación dentro del intervalo o intervalos de datos faltantes. En esta etapa se pueden aplicar diferentes estrategias de interpolación para las funciones de amplitud y fase. Debido a la forma regular de las curvas asociadas a estas funciones, es esperable obtener curvas más suaves con pocas oscilaciones. Consideramos dos técnicas de interpolación eficaces pero sencillas: la interpolación mediante splines cúbicos y la interpolación cúbica hermitiana que conserva la forma (pchip) [91]. Como vimos en la Sec. 3.4.2, los splines tienden a generar oscilaciones en el intervalo de datos, mientras que el método pchip produce curvas más suaves con menos oscilaciones. El procedimiento específico para esta etapa se detalla a continuación:

1. Las amplitudes armónicas $\tilde{\mathbf{B}}_{k,\ell}$ y las fases $\tilde{\phi}_{k,\ell}$ se recortan en el intervalo de datos faltantes I_{ms} .
2. Las amplitudes y fases se interpolan en I_{ms} , resultando en una estimación de amplitud interpolada $\bar{\mathbf{B}}_{k,\ell}$ y una estimación de fase interpolada $\bar{\phi}_{k,\ell}$.
3. El resultado final de la imputación \mathbf{x}_2 se construye por superposición armónica en el intervalo de datos que faltan:

$$\mathbf{x}_2(n) := \sum_{k=1}^K \sum_{\ell=1}^{r_k^*} \bar{\mathbf{B}}_{k,\ell}(n) \cos(2\pi \bar{\phi}_{k,\ell}(n)),$$

para $n \in I_{\text{ms}}$. La parte restante de \mathbf{x}_2 satisface

$$\mathbf{x}_2(n) := \mathbf{x}_1(n), \quad (4.34)$$

para $n \in \{1, \dots, N\} \setminus I_{\text{ms}}$.

Los valores faltantes de la tendencia $\hat{\mathbf{T}}_1$ se interpolan directamente para producir la estimación final $\hat{\mathbf{T}}_2$. Este método de interpolación se elige porque las tendencias variables en el tiempo suelen mostrar variaciones suaves y carecen de oscilaciones locales a escala de los intervalos de datos faltantes.

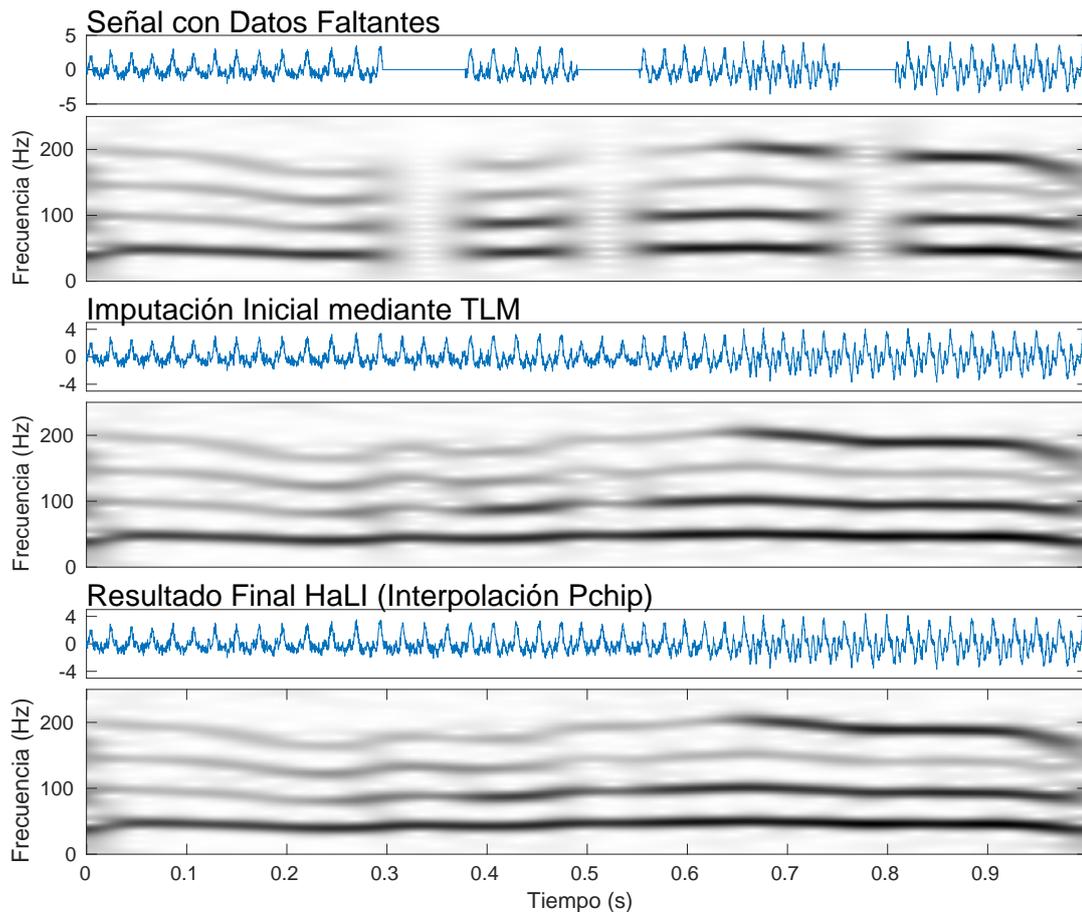


Figura 4.3: Ejemplo de señal sintética con datos faltantes. Primera y segunda fila: Ejemplo de señal sintética ruidosa con forma de onda variable en el tiempo y 20% de datos faltantes junto al módulo de su STFT. Tercera y cuarta fila: Versión imputada de la señal utilizando el mapa de retardos de Takens en el paso de imputación inicial junto al módulo de su STFT. Quinta y sexta fila: Resultado final del algoritmo HaLI junto al módulo de su STFT.

La salida resultante del algoritmo HaLI es la señal imputada $x_2 + \hat{T}_2$. Es importante destacar que este resultado final de la imputación representa una estimación de x carente de componentes aleatorias asociadas al ruido. La eliminación de ruido de la señal es una tarea común en varias aplicaciones de procesamiento de señales. En esencia, nuestro método propuesto no sólo imputa los datos que faltan, sino que también realiza una limpieza de ruido de la señal simultáneamente, un tema que exploraremos más a fondo en la siguiente sección.

4.5. Validación en señales sintéticas

En esta sección, presentamos una serie de experimentos sobre señales sintéticas para mostrar la eficacia del algoritmo HaLI. Las señales sintéticas se generaron a partir del ANHM, considerando formas de onda variables en el tiempo. El objetivo de estos experimentos es mostrar la eficacia de nuestro método en la restauración de datos faltantes en segmentos donde cambian los patrones de oscilación. Teniendo en cuenta las múltiples

opciones para la imputación inicial del Paso 1 del algoritmo, llevamos a cabo un estudio comparativo que abarca los algoritmos enumerados en la Tabla 4.1.

El código de MATLAB® para HaLI está disponible en un repositorio en línea¹, donde los lectores también pueden encontrar códigos que reproducen los resultados reportados en esta tesis. Asimismo, el algoritmo se ofrece como un paquete listo para usar, el cual también se encuentra disponible².

En esta sección, para especificar qué algoritmos se utilizan en la primera y tercera etapa del algoritmo HaLI, adoptamos la notación HaLI[IMP](int), donde [IMP] indica el método de imputación empleado en el paso inicial y (int) representa el tipo de interpolador utilizado en el paso final. El método de interpolación elegido se indica mediante int = s para splines cúbicos o int = p para interpolación pchip.

Consideramos señales simuladas que satisfacen el ANHM monocomponente (1.31). Generamos estas señales definiendo la amplitud y la fase instantáneas como

$$B_1(t) = \sqrt{t+1}, \quad \phi_1(t) = 50t + 5/(2\pi) \cos(2\pi t) + Y(t),$$

respectivamente. Aquí, $Y(t)$ es un proceso aleatorio definido como

$$Y(t) = \int_0^t \frac{R_B(u)}{\|R_B(u)\|_\infty} du, \quad (4.35)$$

donde $R_B(t)$ es un ruido blanco gaussiano estándar promediado en el tiempo. Las fases armónicas se fijan en $\phi_\ell(t) = e_\ell \phi_1(t)$ para $\ell \geq 2$, donde e_ℓ es una variable aleatoria uniforme con soporte en $[0,95\ell, 1,05\ell]$. Las amplitudes armónicas $B_\ell(t)$ están relacionadas con $B_1(t)$ a través de la condición (C4) de modo que, para cada armónico, $\alpha_\ell(t) = B_\ell(t)/B_1(t)$ es lentamente variable en el tiempo e independiente de la variación de los otros armónicos. Las señales se sintetizan utilizando una frecuencia de muestreo $f_s = 4000$ Hz y una duración de $T_s = 1$ s.

Para cada señal, simulamos valores omitidos con distintos porcentajes de omisión $P_{ms} \in [5\%, 20\%]$ y generamos tres intervalos de datos faltantes con longitudes aleatorias de tal forma que

$$L_1 + L_2 + L_3 = N \frac{P_{ms}}{100}, \quad (4.36)$$

donde L_i , con $i = 1, 2, 3$, es la longitud del i -ésimo intervalo de datos faltantes. En cuanto al componente de ruido $\Phi(t)$, consideramos una distribución gaussiana de media cero con dos valores de varianza correspondientes a SNR de 10 y 20 dB. También aplicamos el algoritmo HaLI a versiones sin ruido añadido para establecer una referencia. En total, se generaron 100 señales según el ANHM para cada nivel de ruido. Sobre cada una, se determinaron intervalos de datos faltantes ubicados aleatoriamente con distinto porcentaje de omisión de datos y nivel de ruido. Por lo tanto, procesamos un total de 400 señales aleatorias para cada nivel de SNR.

La Fig. 4.3 ilustra el efecto de la imputación inicial de datos. Las dos filas superiores muestran la señal ruidosa original (SNR = 10 dB) con 20% de datos faltantes en tres intervalos y el módulo de su STFT ($|\mathbf{F}|$), revelando discontinuidades en el plano TF

¹https://github.com/joaquir-uner/MSI_HarmDecomp

²https://github.com/joaquir-uner/harmonic_imputation

4.5. VALIDACIÓN EN SEÑALES SINTÉTICAS

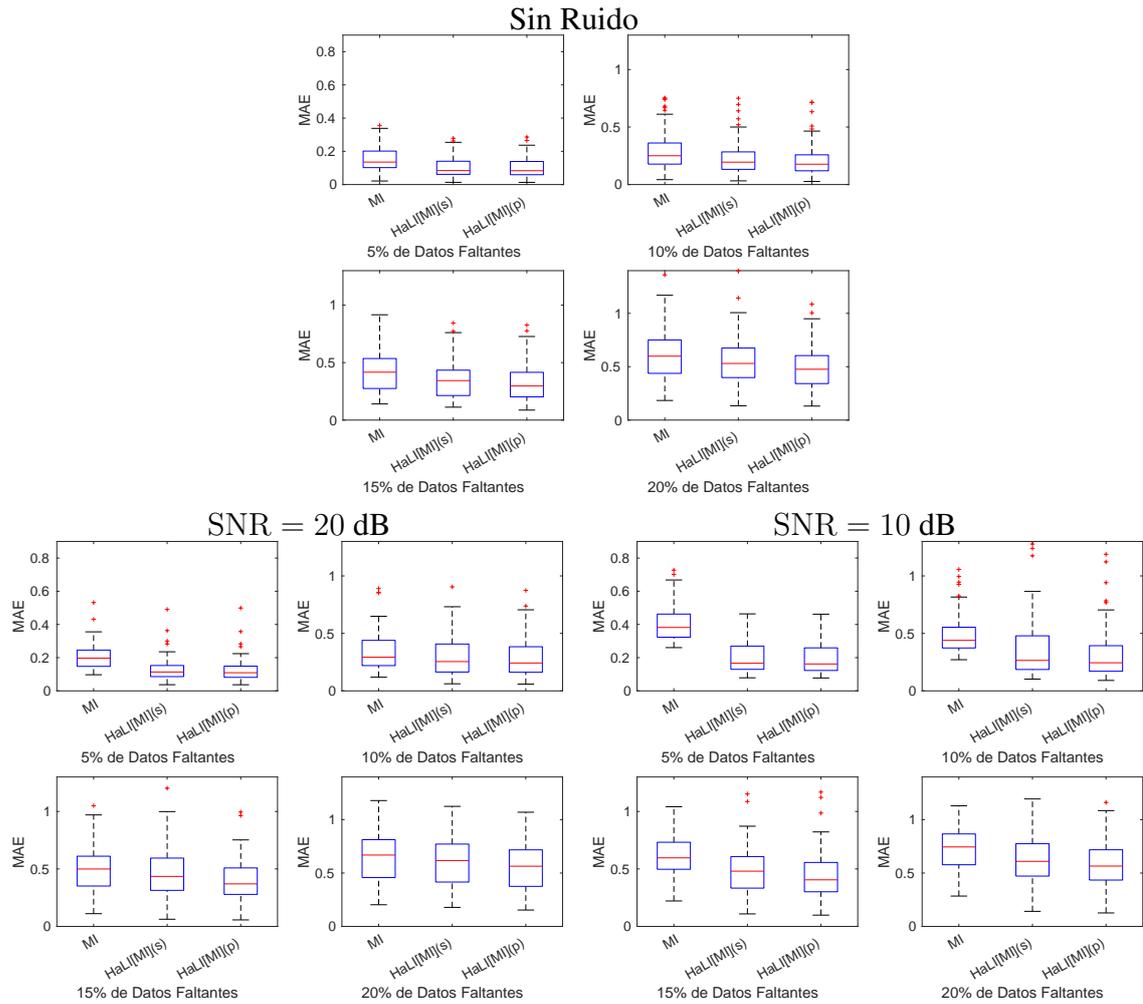


Figura 4.4: Resultados de imputación de datos faltantes en señales sintéticas. Gráfico de cajas del MAE para señales sintéticas con diferentes niveles de ruido. Para cada porcentaje de datos faltantes, se muestran las cajas para el mejor método de imputación (MI) y el algoritmo completo HaLI utilizando los esquemas de interpolación mediante splines (HaLI[MI](s)) y pchip (HaLI[MI](p)). En ambos casos, los pasos de descomposición armónica e interpolación se realizan sobre el resultado del método de imputación inicial de mejor rendimiento basado en el MAE. Se encontraron diferencias estadísticamente significativas ($p < 0,0167$) entre los valores medios de los tres métodos según la prueba de rango con signo de Wilcoxon (véase la Tabla 4.2 para más detalles).

y borrosidad de las crestas armónicas. Las filas tercera y cuarta muestran la señal después de la imputación inicial y el módulo de su STFT, donde se reducen los artefactos debidos a los datos que faltan. El algoritmo HaLI mejora aún más esta RTF mediante la descomposición armónica y la interpolación, como puede verse en las dos últimas filas de la Fig. 4.3.

Para evaluar cada método de imputación, utilizamos como medida de desempeño el error absoluto promedio (MAE, por las siglas en inglés de *mean absolute error*) entre la señal imputada y la señal original, sin datos faltantes y sin ruido. En primer lugar, evaluamos los diferentes métodos de imputación inicial detallados en la Tabla 4.1. El método óptimo de imputación inicial se denota como MI (Mejor Imputación). Se calculan el MAE

4.5. VALIDACIÓN EN SEÑALES SINTÉTICAS

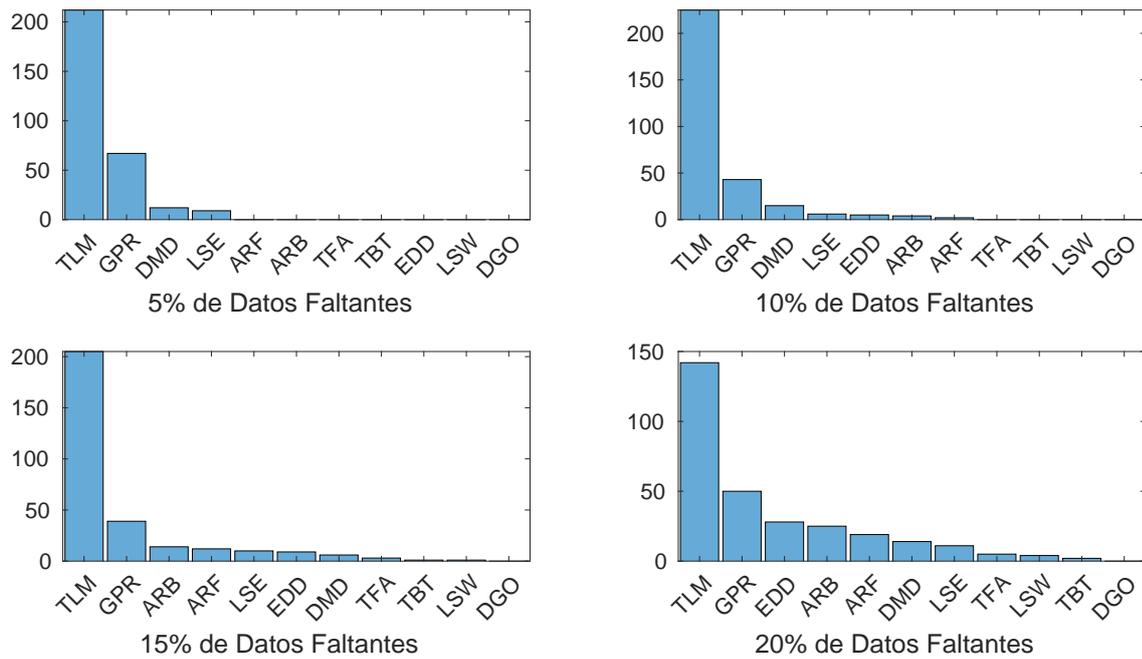


Figura 4.5: Desempeño de los métodos de imputación inicial para señales sintéticas. Se muestran los histogramas de frecuencias de mejor desempeño para cada método de imputación inicial en señales sintéticas para cada porcentaje de omisión de datos. Referencias: TLM: Mapa de retardos de Takens. LSE: Estimación por mínimos cuadrados. DMD: Descomposición modal dinámica. EDD: Descomposición modal dinámica extendida. GPR: Regresión por procesos gaussianos. ARF: Predicción por Regresión ARIMA con paso hacia adelante. ARB: Predicción por Regresión ARIMA con paso hacia atrás. TBT: Predicción mediante modelo TBATS. TFA: Búsqueda de coincidencias en TFR ralas. LSW: Procesos ondita localmente estacionarios. DGO: Suavizado no lineal de datos con huecos aleatorios y valores atípicos.

de los resultados finales obtenidos con HaLI y se comparan con los de la imputación inicial. Se utilizó la prueba de hipótesis con signo de Wilcoxon con corrección de Bonferroni para determinar si las diferencias en los valores medios de los MAE antes y después de la interpolación son estadísticamente significativas.

La Fig. 4.4 muestra los resultados del experimento mediante gráficos de caja, y la Tabla 4.2 presenta los valores MAE medios y los valores p correspondientes para la comparación entre la mejor imputación inicial y el algoritmo HaLI completo utilizando ambos métodos de interpolación. Los resultados demuestran que el algoritmo HaLI supera sistemáticamente a la imputación inicial para los diferentes porcentajes de omisión, métodos de interpolación y niveles de ruido. En particular, la interpolación cúbica que preserva la forma supera sistemáticamente a la interpolación con splines cúbicos en todos los escenarios, lo que se atribuye a la reducción de las oscilaciones de pchip en comparación con los splines. Además, la prueba de rango con signo de Wilcoxon valida la mejora estadísticamente significativa en todos los casos al emplear la interpolación spline o pchip (valor $p < 0,0001$), respecto a la imputación inicial.

Para comparar las estrategias de imputación en el primer paso, presentamos histogramas de frecuencia para el método de imputación inicial que ofrece el mejor rendimiento en

4.5. VALIDACIÓN EN SEÑALES SINTÉTICAS

		$P_{ms} [\%]$			
		5	10	15	20
Sin Ruido	MAE_I	0.1352	0.2505	0.4167	0.5990
	MAE_S	0.0842	0.1948	0.3436	0.5304
	MAE_P	0.0832	0.1768	0.2967	0.4762
	p_{I-S}	< 0,0001	< 0,0001	< 0,0001	< 0,0001
	p_{I-P}	< 0,0001	< 0,0001	< 0,0001	< 0,0001
	p_{S-P}	< 0,0001	< 0,0001	< 0,0001	< 0,0001
20 dB	MAE_I	0.1971	0.2943	0.4994	0.6685
	MAE_S	0.1135	0.2554	0.4349	0.6172
	MAE_P	0.1093	0.2437	0.3713	0.5640
	p_{I-S}	< 0,0001	< 0,0001	0,0001	< 0,0001
	p_{I-P}	< 0,0001	< 0,0001	< 0,0001	< 0,0001
	p_{S-P}	< 0,0001	< 0,0001	< 0,0001	< 0,0001
10 dB	MAE_I	0.3818	0.4396	0.5963	0.7445
	MAE_S	0.1663	0.2664	0.4805	0.6094
	MAE_P	0.1620	0.2444	0.4052	0.5652
	p_{I-S}	< 0,0001	< 0,0001	< 0,0001	< 0,0001
	p_{I-P}	< 0,0001	< 0,0001	< 0,0001	< 0,0001
	p_{S-P}	< 0,0001	< 0,0001	< 0,0001	< 0,0001

Tabla 4.2: Resultados de imputación de datos faltantes en señales sintéticas. Mediana de los valores de MAE para cada método inicial y el procedimiento HaLI para señales sintéticas en cada nivel de ruido junto con los p -valores para la comparación de medianas entre los diferentes métodos de imputación. I: Mejor imputación inicial. S: HaLI[MI](s). P: HaLI[MI](p), donde MI indica que se eligió el método de imputación inicial con mejor desempeño.

	$P_{ms} [\%]$			
	5	10	15	20
TLM	0.027	0.0245	0.0235	0.0225
LSE	0.0116	0.01	0.001	0.001
DMD	0.025	0.0235	0.0216	0.0211
EDMD	0.09	0.13	0.09	0.09
GPR	1.690	1.531	1.455	1.412
ARIMAF	5.06	4.180	4.519	4.467
ARIMAB	4.707	4.477	4.286	4.010
TBATS	11.192	10.917	10.866	11.038
DDTFA	241.28	230.21	210.59	210.76
LSW	6.112	9.786	19.021	37.277
DRAGO	16.389	15.484	14.922	13.909
HaLI[MI](s) (Etapas 2 y 3)	0.903	0.895	0.861	0.871
HaLI[MI](p) (Etapas 2 y 3)	0.901	0.893	0.859	0.870

Tabla 4.3: Tiempo medio de cálculo (en segundos) de los distintos algoritmos para el experimento con señales sintéticas para los diferentes porcentajes de omisión de datos. Los algoritmos fueron ejecutados en el entorno MATLAB[®] 2021b en una computadora con sistema operativo Ubuntu 22.04.4 LTS, procesador Intel[®] Core[™] i7-11700@2.5GHz x 16 núcleos y 64 GiB de memoria RAM.

la Fig. 4.5. Observamos que el algoritmo TLM supera sistemáticamente a los otros métodos para todos los porcentajes de omisión de datos. Entre ellos, los métodos de predicción de sistemas dinámicos (GPR, DMD, LSE y EDMD) siguen a TLM para $P_{ms} = 5\%$ y 10% . A partir de $P_{ms} = 15\%$, los métodos basados en predicción ARIMA muestran un rendimiento similar, superando sistemáticamente a otros métodos excepto TLM y GPR. Para $P_{ms} = 20\%$, TLM sigue siendo el método con mejores resultados, seguido de GPR y EDMD. Por último, el predominio del método TLM disminuye a medida que aumenta la longitud de los segmentos de datos faltantes, lo que podría estar relacionado con su limitación a la hora de encontrar una plantilla adecuada para imputar los datos para intervalos más largos. La Tabla 4.3 muestra los tiempos de cómputo promedio de los distintos algoritmos, incluido el tiempo combinado de los pasos 2 y 3 en HaLI para ambas alternativas de interpolación. Estos resultados ponen de manifiesto que TLM es preferible como método imputación inicial debido a su mayor precisión y menor costo computacional.

4.6. Experimentos en señales reales

En esta sección, presentamos los resultados de la aplicación del algoritmo propuesto para la imputación de datos faltantes en señales reales del ámbito biomédico. En este experimento, analizamos señales fisiológicas del mundo real procedentes de diversas fuentes, incluida la Base de Datos Integrada de Taiwán para el Sueño Inteligente (TIDIS)³, una base de datos de señales etiquetadas de acelerometría⁴, y señales de presión arterial de la base de datos CHARIS⁵. Nuestra selección de la base de datos TIDIS incluye registros de fotopletomografía (PPG), señal de flujo de aire (AF, del inglés *airflow*), señal de presión nasal (NP, del inglés *nasal pressure*) e impedanciometría torácica (THO). Para la acelerometría, nos centramos en la señal sensada a lo largo del eje x del sensor del tobillo derecho durante la marcha. Un ejemplo para cada una de estas señales se muestra en la Fig. 4.6. De esta gráfica, se observa que las señales de ABP de CHARIS muestran una tendencia significativa vinculada a la presión arterial media (MAP), vital para la evaluación clínica de las señales imputadas. En total, evaluamos 131 señales distintas, distribuidas en 25 registros de PPG, 17 de AF, 25 de NP, 19 de THO, 32 de ACC y 13 de ABP.

De cada señal, se seleccionaron segmentos libres de saturación o desconexiones del sensor y con un patrón oscilatorio claramente visible. Para CHARIS, se eligieron 3 segmentos para cada señal de la base de datos. De manera análoga al estudio de señales sintéticas, se generaron tres intervalos de datos faltantes para cada segmento, con tasas de omisión de datos entre 5% y 20% . En particular, consideramos a todas las señales conformadas por un único componente oscilatorio, excepto la señal de ABP, en la que se aprecia un componente oscilatorio no trivial de variación lenta ligado a la dinámica respiratoria. Por lo tanto, se fijó para el modelo (4.31) un valor $K = 1$ para todas las señales excepto ABP, donde se eligió $K = 2$. La prueba de estacionariedad propuesta por Borgnat y cols. [143] se aplicó a todos los segmentos de señales fisiológicas. Los resultados indican un alto grado de no estacionariedad en la mayoría de los segmentos, excepto en los registros de THO. En concreto, la prueba clasificó los siguientes porcentajes de señales de cada conjunto como no estacionarios: 100% para ACC, 92% para PPG, 87% para ABP,

³<https://tidis.org/en/>

⁴<https://physionet.org/content/accelerometry-walk-climb-drive/1.0.0/>

⁵<https://physionet.org/content/charisdb/1.0.0/>

4.6. EXPERIMENTOS EN SEÑALES REALES

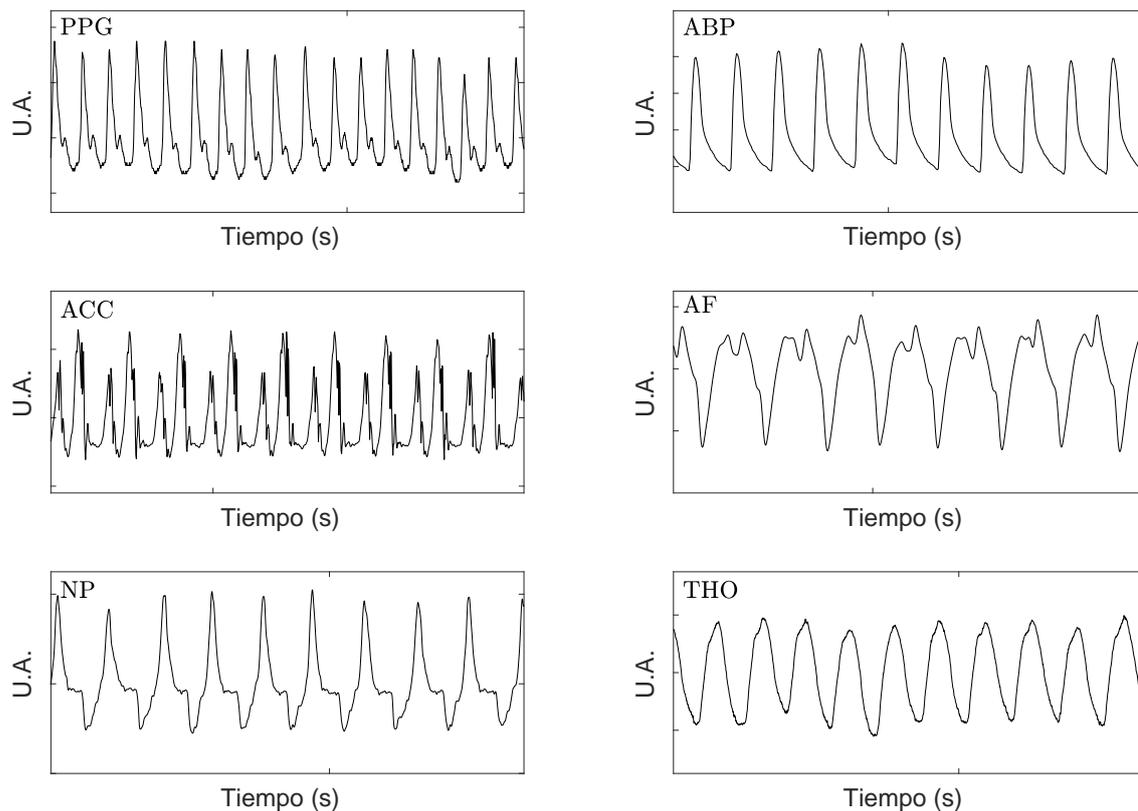


Figura 4.6: Ejemplos de las diferentes señales biomédicas consideradas para el experimento de imputación de datos faltantes en señales reales. Referencias: PPG - Fotoplethysmografía. ABP - Presión Sanguínea Arterial. ACC - Acelerometría. AF - Flujo de Aire. NP - Presión de Aire Nasal. THO: Impedanciometría Torácica.

60 % para NP, 53 % para AF y 21 % para THO.

La configuración más eficaz para HaLI se determina en base a la métrica del error, considerando el error medio absoluto normalizado (NMAE) debido a las diferentes escalas de amplitud de las señales estudiadas. Los resultados se resumen en los diagramas de cajas de la Fig. 4.7. En particular, la mediana del valor de NMAE para HaLI es sistemáticamente inferior en comparación con la imputación inicial en la mayoría de los casos (como se resume en la Tabla 4.4). La prueba de hipótesis de rango con signo de Wilcoxon confirma la significancia estadística de esta mejora, como indican los valores p correspondientes en la Tabla 4.4. En la mayoría de los casos, HaLI(p) mejora sistemáticamente la medida NMAE con respecto a la imputación inicial, excepto en casos como PPG y ABP con 15 % y 20 %, así como las señales AF con tasas de datos faltantes de 5 %. Esto se atribuye a las frecuencias fundamentales significativamente más altas de las señales PPG y ABP en comparación con las señales respiratorias (AF, NP y THO), lo que resulta en más ciclos dentro del intervalo o intervalos de valores faltantes. Esto pone de manifiesto el hecho de que los métodos de imputación inicial considerados en este experimento presentan limitaciones a la hora de tratar con un intervalo de datos faltantes que incluye un número elevado de ciclos de las señales.

No obstante, es evidente una mejora significativa en otros escenarios cuando se aplica el enfoque propuesto. La Fig. 4.8 ofrece un contraste visual entre la imputación inicial y

	PPG			ABP					
	5	10	15	20	5	10	15	20	
P_{ms} [%]									
NMAE _I	0,0572	0,0829	0,0958	0,1087	0,0440	0,0619	0,0659	0,0793	
NMAE _S	0,0580	0,0781	0,0910	0,1203	0,0378	0,0538	0,0749	0,0997	
NMAE _P	0,0561	0,0752	0,0832	0,1053	0,0347	0,0530	0,0736	0,0703	
p_{I-S}	0,016	0,0007	0,3533	0,0780	0,0019	0,0524	0,1546	< 0,0001	
p_{I-P}	0,011	0,0004	0,0615	0,1036	< 0,0001	0,0018	0,0048	0,0697	
p_{S-P}	0,925	0,3967	0,0653	0,0002	< 0,0001	0,0007	0,0003	< 0,0001	
	ACC			AF					
NMAE _I	0,0491	0,0559	0,0630	0,0702	0,1344	0,1361	0,1742	0,2218	
NMAE _S	0,0468	0,0519	0,0633	0,0766	0,1313	0,1197	0,1570	0,1806	
NMAE _P	0,0462	0,0504	0,0587	0,0664	0,1312	0,1211	0,1552	0,1788	
p_{I-S}	0,0225	0,0999	0,2699	0,8811	0,5228	0,0003	0,0495	0,0049	
p_{I-P}	0,0019	0,0042	0,0099	0,0009	0,4925	0,0003	0,0010	0,0003	
p_{S-P}	0,0029	< 0,0001	< 0,0001	< 0,0001	0,5228	0,5228	0,0016	0,0086	
	NP			THO					
P_{ms} [%]									
NMAE _I	0,1055	0,1233	0,1495	0,1568	0,1065	0,1444	0,1620	0,2230	
NMAE _S	0,0987	0,1050	0,1404	0,1425	0,0847	0,1211	0,1370	0,1859	
NMAE _P	0,0987	0,1024	0,1197	0,1320	0,0840	0,1219	0,1260	0,1747	
p_{I-S}	0,0054	0,0001	0,0021	0,0004	0,141	0,0001	0,0022	0,0002	
p_{I-P}	0,0049	0,0001	0,0001	0,0001	0,0141	0,0001	0,0001	0,0001	
p_{S-P}	0,1500	0,0027	0,0001	0,0004	0,0766	0,0043	0,0100	0,0011	

Tabla 4.4: Resultados de imputación de datos faltantes en señales reales. Mediana del error medio absoluto normalizado (NMAE) y valores p de las pruebas de Wilcoxon de rango con signo para el resultado de imputación inicial y el HaLI utilizando ambos esquemas de interpolación para todas las señales fisiológicas. I: Mejor imputación inicial. S: HaLI[MI](s). P: HaLI[MI](p). Los valores p estadísticamente significativos aparecen en negrita.

4.7. COMENTARIOS DE FIN DE CAPÍTULO

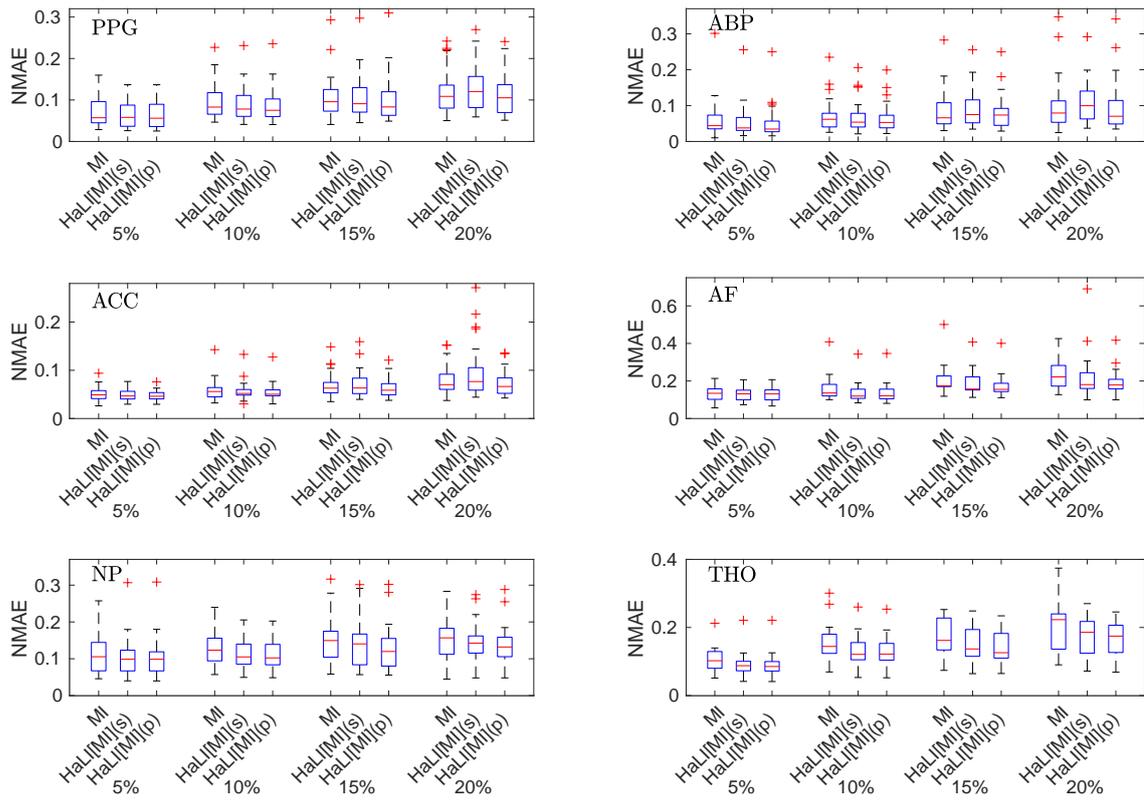


Figura 4.7: Resultados de imputación de datos faltantes en señales reales. Gráficos de caja de los errores absolutos medios normalizados (NMAE) para cada tipo señal fisiológica considerada para cada tasa de datos faltantes. MI: mejor imputación inicial; HaLI[MI](s): HaLI basado en la mejor imputación inicial e interpolación spline; HaLI[MI](p): HaLI basado en la mejor imputación inicial e interpolación cúbica que preserva la forma. Primera fila: resultados para PPG (izquierda) y ABP (derecha). Segunda fila: resultados para ACC (izquierda) y AF (derecha). Tercera fila: resultados para NP (izquierda) y THO (derecha).

HaLI para las diferentes señales consideradas. Vemos que HaLI mitiga las sobreamplitudes a través de la interpolación de amplitud y mejora la sincronización entre los ciclos originales e imputados a través de la interpolación de fase. Finalmente, la Fig. 4.9 presenta los histogramas para el método de imputación inicial de mejor rendimiento en cada señal, con TLM demostrando superioridad sobre el resto de los métodos en la totalidad de los casos.

4.7. Comentarios de fin de capítulo

En este capítulo presentamos una nueva estrategia para imputar datos faltantes en series temporales oscilatorias no estacionarias. Nuestra propuesta combina técnicas de imputación preestablecidas, con la descomposición armónica y la interpolación a nivel armónico de las amplitudes y fases variantes en el tiempo. Evaluamos diferentes algoritmos de imputación de datos faltantes y desarrollamos un código MATLAB[®] con el objetivo de que sea utilizado por la comunidad científica para analizar diferentes tipos de señales del ámbito biomédico. Nuestros experimentos, que incluyen el estudio de señales

4.7. COMENTARIOS DE FIN DE CAPÍTULO

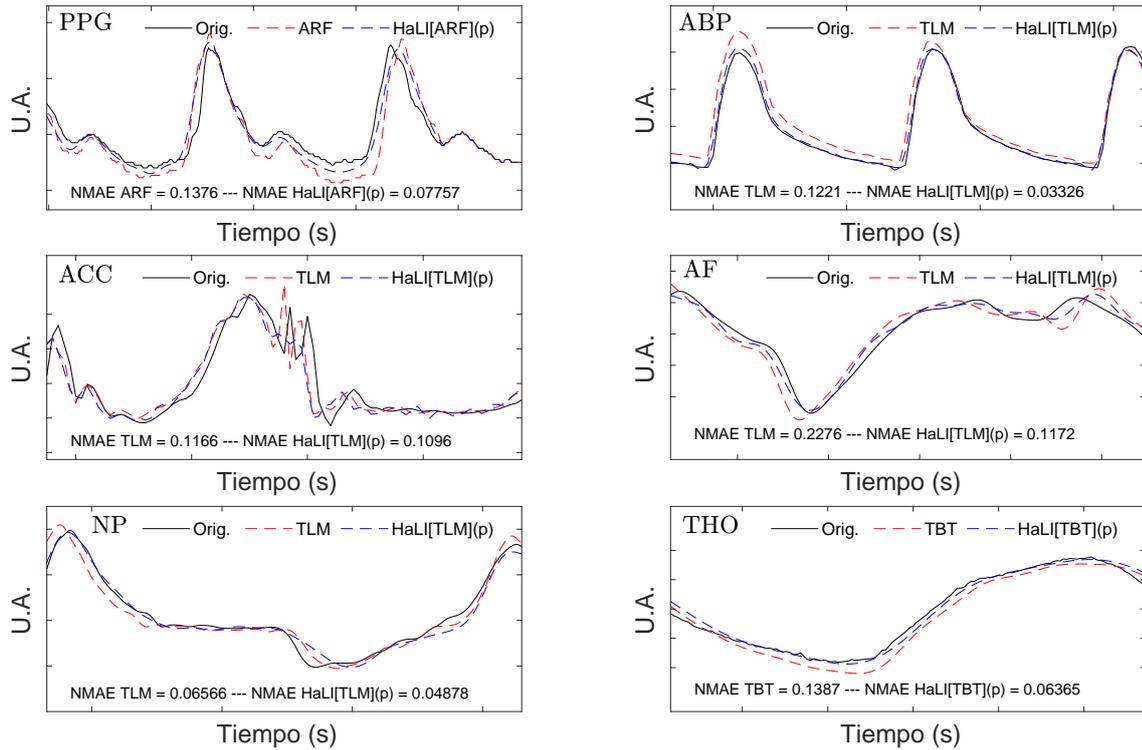


Figura 4.8: Comparación entre la imputación inicial de datos faltantes antes y después del posprocesamiento. Los registros originales se muestran en negro y el mejor resultado de imputación inicial y nuestra imputación mejorada utilizando HaLI se superponen en rojo y azul, respectivamente. El NMAE con y sin HaLI se muestra en la parte inferior de cada gráfico.

sintéticas y del mundo real, demuestran la eficacia del algoritmo HaLI, debido particularmente a la interpolación a nivel de los armónicos. Analizamos señales reales de diversas características, demostrando la robustez del método ante diferentes frecuencias de muestreo y frecuencias fundamentales.

En resumen, consideramos que el enfoque HaLI resulta muy prometedor, y buscaremos explorar posibles aplicaciones clínicas de estos algoritmos. Es importante destacar que la descomposición armónica y la interpolación a nivel armónico pueden aplicarse como pasos de posprocesamiento para mejorar cualquier método de imputación inicial, ofreciendo versatilidad para la integración de nuestra propuesta junto a técnicas existentes de imputación de datos faltantes.

Por otro lado, hay varias consideraciones técnicas que merecen ser discutidas. En primer lugar, los escenarios del mundo real a menudo involucran situaciones donde se produce la llamada mezcla de modos. Este fenómeno se da cuando las frecuencias instantáneas de diferentes componentes se solapan. Nuestro algoritmo actual tiene limitaciones en el tratamiento de este problema, y tenemos previsto explorar nuevos enfoques para la descomposición armónica. En particular, la transformada *chirplet* [144] resulta prometedora para abordar este reto.

Además, la suposición de que $B_{k,\ell}$ es una secuencia sumable indexada por ℓ exige un patrón oscilatorio “suave” es decir, sin picos. Sin embargo, muchas señales del mundo

4.7. COMENTARIOS DE FIN DE CAPÍTULO

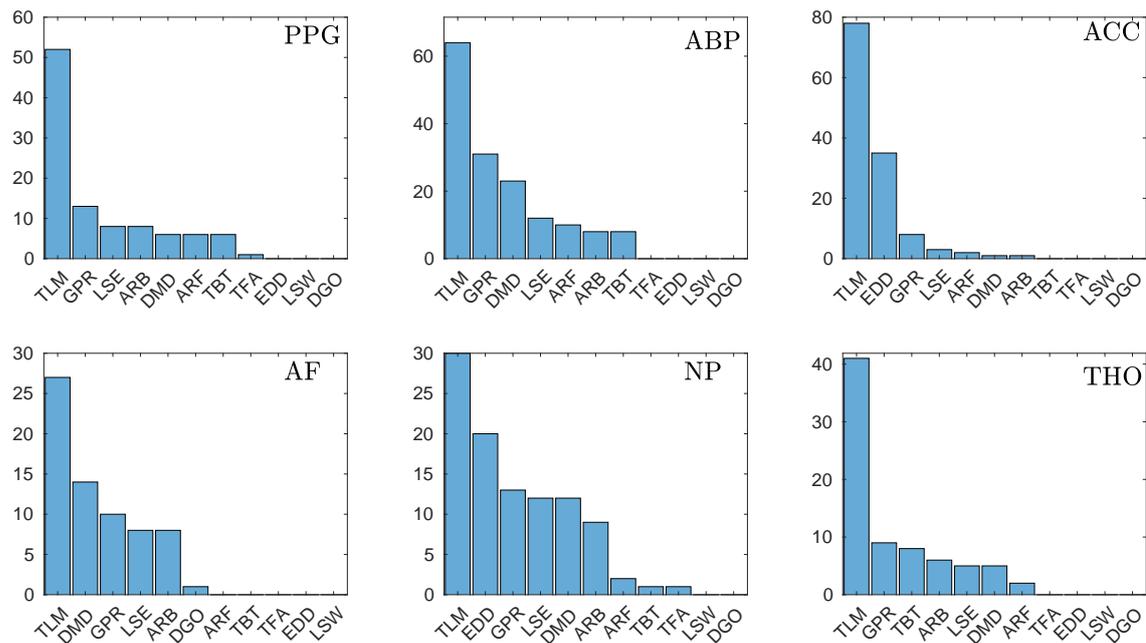


Figura 4.9: Desempeño de los métodos de imputación inicial para señales reales. Se muestran los histogramas de frecuencias del método de imputación inicial con mejores resultados en los distintos de señales fisiológicas consideradas. Referencias: TLM: Mapa de retardos de Takens. LSE: Estimación por mínimos cuadrados. DMD: Descomposición modal dinámica. EDD: Descomposición modal dinámica extendida. GPR: Regresión por procesos gaussianos. ARF: Predicción por Regresión ARIMA con paso hacia adelante. ARB: Predicción por Regresión ARIMA con paso hacia atrás. TBT: Predicción mediante modelo TBATS. TFA: Búsqueda de coincidencia en RTF ralas. LSW: Procesos ondita localmente estacionarios. DGO: Suavizado no lineal de datos con huecos aleatorios y valores atípicos

real, como los electrocardiogramas o los potenciales de acción, presentan patrones con picos. Es esencial adaptar el algoritmo de imputación propuesto para contemplar tales escenarios.

Por otro lado, nuestro modelo actual caracteriza las series temporales oscilatorias y sus tendencias describiendo su comportamiento oscilatorio, mediante un modelo fenomenológico. Se podría aprovechar conocimiento *a priori* de las señales para formular un modelo estadístico más complejo de los componentes armónicos y la tendencia. Sin embargo, este enfoque dependerá de cada aplicación específica.

Por último, esta propuesta se centra en series temporales unidimensionales. En la práctica, es común encontrar series temporales multidimensionales con valores faltantes. Aunque la imputación por separado es factible, resultaría interesante explorar alternativas de imputación conjunta de los valores faltantes de las distintas dimensiones de los datos. Se pueden encontrar ejemplos de imputación conjunta en [145]. Dentro de la comunidad estadística, la imputación conjunta o imputación múltiple es una tarea bien establecida y ampliamente estudiada. Cuando una serie temporal oscilatoria se ajusta a un modelo estadístico similar al descrito en [146], consideramos que HaLI podría adaptarse de una manera directa a este tipo de conjuntos de datos. Si consideramos una serie temporal como

una función definida sobre \mathbb{R} , es natural preguntarse si la idea de imputación propuesta podría generalizarse a conjuntos de datos espaciales, es decir funciones definidas sobre \mathbb{R}^2 [147], o incluso a espacios no lineales de mayor dimensión [148].

Cabe aclarar que existen varios paquetes de software fáciles de usar en aplicación de imputación de datos faltantes en series temporales. Algunos ejemplos son las librerías escritas en el lenguaje de programación R como “imputeTS”⁶ y los paquetes basados en Python como “impyute”⁷, entre otros. Estos paquetes ofrecen una serie de métodos de imputación, desde los más sencillos, como la imputación de la media y la última observación realizada (LOCF, del inglés *last observation carried forward*), hasta técnicas más avanzadas basadas en modelos no estacionarios, como los métodos ARMA y de series temporales estructurales. Además, suelen proporcionar herramientas para la visualización de datos y el cálculo de métricas de rendimiento. Sin embargo, es importante señalar que actualmente no se dispone de una toolbox con estas características para el entorno MATLAB[®] que sea sencilla de usar. Por lo tanto, otro aporte de esta tesis es la implementación en MATLAB[®] de una toolbox de acceso público, sirviendo como un módulo práctico para los usuarios que quieran aplicar técnicas de imputación de datos en series temporales oscilatorias.

El método de imputación de datos faltantes y los resultados presentados en este capítulo fueron incluidos en un artículo científico publicado en la revista *IEEE Transactions on Signal Processing* de la editorial del Instituto de Ingenieros Eléctricos y Electrónicos (IEEE) [4].

⁶<https://cran.r-project.org/web/packages/imputeTS/index.html>

⁷<https://impyute.readthedocs.io/en/master/>

Capítulo 5

Análisis de la señal de voz mediante medidas de variabilidad de la función de forma de onda

El estudio de la salud vocal se centra en evaluar el estado del aparato fonador, el cual es el conjunto de estructuras orgánicas involucradas en la generación de la voz humana. Estas estructuras incluyen diversos músculos, nervios, huesos y estructuras cartilaginosas que interactúan entre sí para dar como resultado la onda de presión que se conoce como señal de voz. El bienestar de la voz y, por consiguiente del aparato fonador, es de suma importancia para el normal desempeño de las personas en su vida cotidiana. La voz es la principal herramienta de comunicación de las personas, por lo cual las afecciones que dificultan el habla tienen consecuencias considerables en el bienestar físico, social y psíquico de las personas.

La señal de voz es la herramienta más empleada para el análisis de la salud vocal, debido a su facilidad de adquisición y a la gran cantidad de información que provee sobre el estado de salud del paciente. Una gran variedad de técnicas han sido desarrolladas para obtener información sobre diferentes porciones del aparato fonador. Entre ellas, podemos mencionar el filtrado inverso [149], análisis en espacio de estados [150], análisis tiempo-frecuencia [151, 152] e inteligencia computacional [153], entre muchas otras.

La evaluación clínica de la voz humana se basa tanto en la evaluación instrumental como en el análisis perceptivo. La evaluación instrumental implica el uso de diversos dispositivos de medición, como micrófonos, electroglotógrafos, acelerómetros y videolaringscopios, para explorar las causas subyacentes de los trastornos vocales. Estas herramientas permiten analizar los factores acústicos, aerodinámicos y anatómicos que contribuyen a la producción de la voz. En cambio, el análisis perceptivo implica valoraciones subjetivas por parte de profesionales de la salud vocal para evaluar la gravedad de los trastornos vocales [154]. Para ello, suelen emplearse escalas de valoración. A través de estas escalas se analizan distintos aspectos perceptuales de la voz, como la ronquera, la astenia, el murmullo y la tensión, entre otros.

La evaluación perceptiva de la voz es una práctica muy extendida en el ámbito clínico debido a su carácter no invasivo, su bajo costo y la información que proporciona sobre el estado del paciente. Sin embargo, una limitación importante de esta práctica radica en la variabilidad de los resultados entre profesionales [155]. Esta variabilidad puede dar lugar a errores de diagnóstico y complicar el tratamiento del paciente.

Para paliar esta situación, en los últimos años se han comenzado a utilizar sistemas de análisis computacional de señales que automatizan la tarea de interpretación de los datos y la toma de decisiones para aliviar la carga de los clínicos y mitigar el sesgo inherente, ofreciendo una precisión competitiva con la clasificación manual. Dada su naturaleza, estos sistemas obtienen o extraen medidas cuantitativas de los registros de las señales de voz y las utilizan para llevar adelante las tareas de análisis vocal. Para esto, diferentes tipos de sistemas basados en inteligencia computacional han sido utilizados para numerosas tareas de evaluación de la voz, como la detección de patologías [156], la identificación de patologías [157], la tipificación de la voz [158] y la emulación de la evaluación perceptiva de la voz [159]. Es esperable que estos sistemas otorguen resultados más objetivos, proporcionando un marco más coherente para la evaluación de la voz y, en última instancia, mejorar la precisión diagnóstica y la atención al paciente.

Recientemente, se han propuesto algoritmos de *aprendizaje profundo* (DL, por las siglas en inglés de *deep learning*) para algunas de estas tareas [160, 161, 162]. Estos algoritmos toman los registros como entrada y devuelven el estado de la voz como salida. En este enfoque, se hace hincapié en el diseño de la arquitectura de las redes que permitan obtener el mayor desempeño posible según las medidas de desempeño seleccionadas. No obstante, los sistemas basados en DL presentan una serie de inconvenientes, como la necesidad de grandes volúmenes de datos y largos ciclos de entrenamiento para alcanzar niveles de desempeño comparables a los de los enfoques de clasificación tradicionales. Además, los parámetros obtenidos a partir del entrenamiento de las redes no pueden interpretarse fácilmente en el contexto de la aplicación clínica del sistema. Este último aspecto motiva la necesidad de características que sean interpretables por los usuarios, permitiendo una comprensión más profunda de las capacidades del sistema. Esto resulta especialmente útil cuando los datos son escasos o cuando los largos procedimientos de entrenamiento resultan prohibitivos.

En este capítulo, detallaremos una propuesta de análisis automático de la condición vocal basado en el ANHM aplicado en señales de voz. Comenzaremos con una breve discusión sobre las herramientas de análisis y procesamiento de señales de voz en la Sec. 5.1. Posteriormente, en la Sec. 5.2 describiremos la aplicación del algoritmo detallado en el Cap. 3 para analizar señales de voz no estacionarias con función de forma de onda variable en el tiempo. A partir de los coeficientes óptimos del modelo, definiremos en la Sec. 5.3 un conjunto de características asociadas a las variaciones temporales de la forma de onda de la señal de voz. Esta variabilidad resultará ser un buen indicador del estado de la voz con una clara interpretación clínica asociada a la inestabilidad del proceso de fonación, dado que diferentes patologías se manifiestan como variaciones en el patrón oscilatorio de la señal de voz. Para validar esta propuesta, las medidas de variabilidad de la forma de onda se utilizan como características para entrenar y probar clasificadores en dos tareas del estudio de la voz: tipificación automática de voces (Sec. 5.4) y detección de enfermedad de Parkinson (Sec. 5.5). El desempeño global de estos clasificadores se evalúa utilizando medidas insensibles al desequilibrio de clases, un reto habitual en las aplicaciones de aprendizaje automático dentro del campo biomédico.

5.1. Análisis de la señal de voz

La voz humana es el resultado de la compleja interacción entre varios sistemas fisiológicos interconectados, incluidos los sistemas respiratorio, muscular, fonatorio y neu-

rológico [163]. Por consiguiente, la señal de voz contiene información importante sobre el estado de estos sistemas, lo que convierte el análisis de la señal de voz en una valiosa herramienta para evaluar la salud de los pacientes. Comúnmente, la generación de la voz se conceptualiza como el resultado de aplicar un filtro que modela el comportamiento del *tracto vocal* sobre una entrada obtenida a partir de una fuente de excitación. El tracto vocal consiste en una serie de articuladores que pueden ajustar su configuración para producir el sonido deseado. Dependiendo del tipo de excitación y la configuración del filtro, diferentes sonidos pueden ser generados mediante este sistema fuente-filtro. En particular, las emisiones de fonemas *sonoros* (por ejemplo, las vocales) se generan a partir de una fuente glótica excitatoria [164], la cual se manifiesta como una serie de pulsos cuasi-periódicos de presión pulmonar. La frecuencia de estos pulsos estará determinada por el movimiento de las cuerdas vocales.

Esta cuasi-periodicidad puede entenderse, en el contexto del análisis de señales, como una modulación en frecuencia. Además, es común encontrar pequeñas diferencias en la amplitud relativa de las oscilaciones de la voz, dando como resultado una modulación en amplitud. Debido a estas perturbaciones tanto en amplitud como en frecuencia, es necesario estudiar a la señal de voz como una señal no estacionaria. Sumado a esto, diversas afecciones pueden generar inestabilidad en las estructuras involucradas en el filtro de tracto vocal, provocando que la respuesta del filtro cambie a lo largo del tiempo, lo que provoca distorsiones en el patrón oscilatorio entre los diferentes ciclos de la señal de voz.

Esta interpretación de la señal de voz como una señal no estacionaria con un patrón oscilatorio variante en el tiempo motiva la aplicación de los algoritmos desarrollados en esta tesis para su análisis.

5.2. Modelado no armónico adaptativo aplicado a señales de voz

En el contexto del análisis de señales, el modelado de datos es útil como herramienta de preprocesamiento para caracterizar eficazmente las señales utilizando un número limitado de parámetros. Más específicamente, en el ámbito del análisis de la voz, se han utilizado modelos armónicos para representar señales asociadas al proceso fonatorio. Estos modelos han sido aplicados a tareas como la mejora del habla, la estimación de la frecuencia fundamental y la separación de fuentes [165, 166, 167, 168, 169]. Sin embargo, el modelado no armónico de la voz y el habla sigue siendo un área poco investigada. Comúnmente, los trastornos de la voz se deben a inestabilidades en el proceso fonatorio, que se manifiestan como cambios dependientes del tiempo en la amplitud y la frecuencia de los componentes de la señal. La principal ventaja del modelado no armónico es la capacidad de describir estas variaciones y extraer características cuantificables de estos registros.

En este marco, es posible estudiar las señales de voz a partir del modelo no armónico adaptativo. En este sentido, una señal monocomponente puede pensarse como una señal que registra una única fuente aislada de sonido, como por ejemplo un individuo. En cambio, las señales multicomponente se componen por varias fuentes, y el modelo no armónico adaptativo podría ser aplicado como una herramienta de separación ciega de fuentes. Sin embargo, por una cuestión de simplicidad, en esta tesis consideraremos el caso de fuentes aisladas, es decir, el caso ANHM monocomponente (1.31). Dada di-

cha condición, es posible extraer la información asociada a la variabilidad temporal de las características de la señal utilizando el modelo no armónico adaptativo con forma de onda variable en el tiempo. Como hemos visto, este modelo permite capturar la evolución temporal del patrón oscilatorio de la señal y es posible estimar el conjunto de parámetros óptimo para aproximar la señal de voz mediante el algoritmo de extracción de forma de onda variante en el tiempo desarrollado en el Cap. 3.

Para ilustrar la aplicabilidad de este algoritmo en señales de voz, analizamos una señal de la base de datos de Saarbruecken [170] correspondiente a un paciente que realiza un ejercicio de cambio de registro grave-agudo-grave (l-h-l, por las siglas en inglés de *low-high-low*). En la Fig. 5.1, mostramos la señal original coloreada según la función de amplitud del tercer armónico, denotada como $\alpha_3(t)$, la cual se obtiene al resolver el problema de regresión no lineal propuesto en (3.27). Se observa una transición brusca en la HAF al inicio y al final del segmento agudo del registro. Basándonos en esto, segmentamos la señal en tres segmentos -grave inicial, agudo y grave final- según el valor de $\alpha_3(t)$. La tercera fila de la Fig. 5.1 muestra la señal original con la forma de onda estimada superpuesta en color. Por último, mostramos los coeficientes de Fourier obtenidos a partir de la forma de onda del ciclo promedio para cada segmento.

Vemos que los segmentos primero y tercero tienen una forma de onda similar ya que están asociados al registro grave de la vocal /a/, mientras que la forma de onda es claramente diferente para el segmento agudo del registro de voz. Este sencillo ejemplo muestra cómo el algoritmo tvWSE puede codificar eficazmente las fuentes de variabilidad temporal de la forma de onda para una señal de voz no estacionaria. En la siguiente sección, veremos cómo obtener medidas cuantitativas asociadas a este modelo para caracterizar las señales de voz.

5.3. Medidas de variabilidad de la forma de onda de la VOZ

La salida del algoritmo tvWSE es un vector de coeficientes $\hat{\gamma}$ que describen la naturaleza variable en el tiempo del patrón oscilatorio de una señal dada. Para poder analizar un conjunto de señales, es necesario establecer de antemano ciertos parámetros del algoritmo para garantizar un marco de referencia coherente para la comparación entre las diferentes señales. En concreto, el orden r del modelo y el número de nodos I_ℓ utilizados para interpolar cada HAF. Para encontrar r , se utilizó el criterio de selección de modelos de regresión trigonométrica propuesto por Wang descrito en la Sec. 2.2.3. Se encontró el valor óptimo r_i para cada señal analizada y se estableció un valor fijo r_{fixed} para todas las señales a partir del percentil 95 de la distribución del orden óptimo para todas las señales. El número de nodos I_ℓ se obtuvo empleando el procedimiento de interpolación automática del número de nodos propuesto en la Sec. 3.5.2. De manera análoga, se obtuvo la distribución de I_ℓ sobre todas las HAFs de cada una de las señales y se eligió el valor fijo I_{fixed} como el percentil 95 de la distribución. Establecer estos valores para todas las señales permite analizar conjuntamente las características de la forma de onda en las distintas señales de voz consideradas.

Reiteramos que el objetivo es encontrar medidas que permitan cuantificar el grado de variabilidad de la forma de onda y utilizar estas características para entrenar modelos de aprendizaje automático para diferentes tareas de clasificación de voz. La Fig. 5.2 ofrece

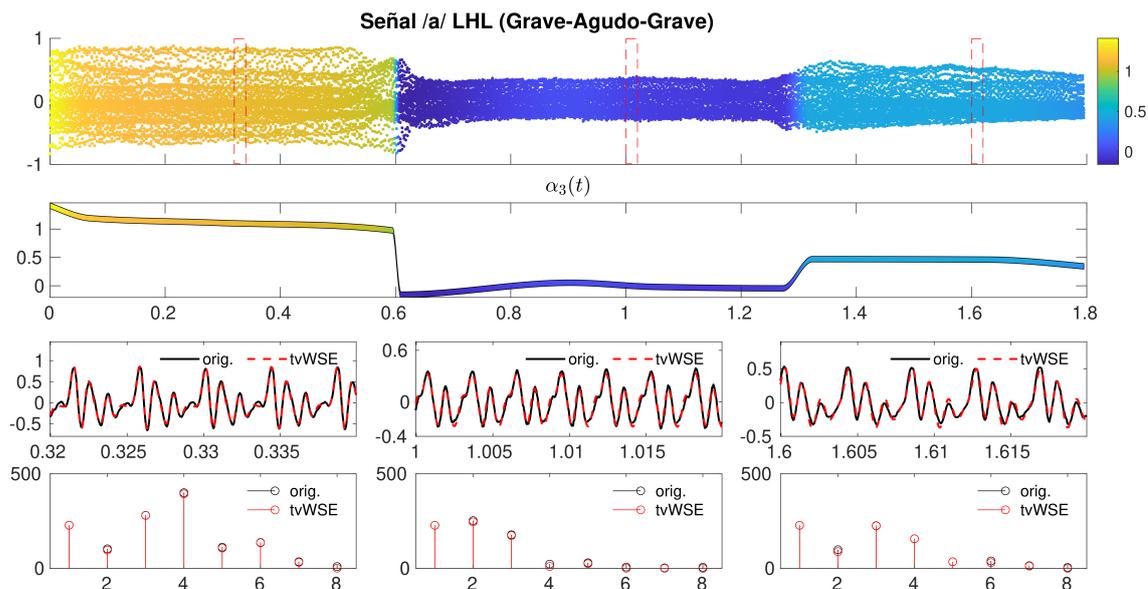


Figura 5.1: Ejemplo de análisis de una señal de voz utilizando el algoritmo tvWSE. Fila superior: Señal de vocal /a/ grave-agudo-grave. Segunda fila: Función de amplitud del tercer armónico $\alpha_3(t)$ extraída mediante el algoritmo tvWSE. Tercera fila: intervalos de 20 milisegundos en tres porciones diferentes de la señal, mostrando la forma de onda variable. Los intervalos corresponden a los recuadros rojos de la primera fila. Última fila: Coeficientes de Fourier de la forma de onda promedio de cada segmento de la señal.

una visión general de los pasos de extracción de características y clasificación de nuestro sistema propuesto. Las características empleadas para entrenar y validar los clasificadores incluyen el error del modelo, la variación total de las funciones (discretas) de amplitud armónica $\hat{\alpha}_\ell(n)$ y los coeficientes relativos de amplitud armónica \hat{s}_ℓ , que definiremos a continuación. El procedimiento de extracción de características se realiza del siguiente modo:

- El error de aproximación del modelo se calcula como

$$E = \frac{\|\mathbf{x}(n) - \hat{\mathbf{x}}(n)\|_2}{\|\mathbf{x}(n)\|_2}, \quad (5.1)$$

donde $\hat{\mathbf{x}}(n)$ es la señal aproximada sintetizada a partir del modelo (3.19) utilizando el vector de parámetros óptimo $\hat{\gamma}$. Aquí suponemos que se le ha restado la media a las señales originales.

- La variación total (TV, del inglés *total variation*) de las funciones de amplitud armónica se calcula a partir del resultado de interpolación de los nodos óptimos, denotado $\hat{\alpha}_\ell(n)$. Para cada HAF, la variación total se obtiene mediante la siguiente fórmula

$$\text{TV}(\boldsymbol{\alpha}_\ell) = \frac{1}{N} \sum_{n=2}^N |\hat{\alpha}_\ell(n) - \hat{\alpha}_\ell(n-1)|. \quad (5.2)$$

- La amplitud relativa para la ℓ -ésima HAF quedará codificada mediante una serie de coeficientes \hat{s}_ℓ que se obtienen a partir de la ecuación $\alpha_\ell(n) = \mathbf{Q}_\ell(n)a_\ell$. Observamos que, según esta fórmula $\sum_n \mathbf{Q}_\ell(n) = N$, por lo que podemos estimar el coeficiente de amplitud cosenoidal a_ℓ integrando la HAF en el tiempo discreto

$$\sum_{n=1}^N \hat{\alpha}_\ell(n) = N\hat{a}_\ell,$$

$$\hat{a}_\ell = \frac{1}{N} \sum_{n=1}^N \hat{\alpha}_\ell(n), \quad \ell = 2, \dots, r. \quad (5.3)$$

Seguidamente, los coeficientes sinusoidales b_ℓ pueden estimarse utilizando la relación $\hat{b}_\ell = \hat{a}_\ell \hat{c}_\ell$. Por último, calculamos el coeficiente de amplitud armónica relativa $\hat{s}_\ell = \sqrt{\hat{a}_\ell^2 + \hat{b}_\ell^2}$. Las variaciones totales y los coeficientes de amplitud de cada armónico se concatenan junto con el error relativo de estimación del modelo para generar el vector de características de cada señal.

Destacamos el hecho de que estas no son las únicas medidas de variabilidad de la forma de onda que pueden obtenerse a partir del ANHM. Otras medidas, como por ejemplo la variación total de la amplitud fundamental $\text{TV}(B_1(t))$ y la frecuencia fundamental $\text{TV}(\phi_1'(t))$, pueden añadirse al conjunto de características para incorporar más información acerca del fenómeno bajo estudio. Nótese que estas cantidades describen la variación *global* de la modulación de amplitud y frecuencia de la señal, sin tener en cuenta la forma de onda intrínseca.

En las siguientes secciones, veremos cómo estas características asociadas a la variabilidad de la forma de onda pueden ser aplicadas al análisis automático de la condición de la voz, en diferentes tareas comúnmente encontradas en el ámbito del estudio de la salud vocal. En particular, la Sec. 5.4 se abocará a la tarea de tipificación de voces según la escala propuesta por Ingo Titze [171]. La Sec. 5.5, en cambio, se enfocará en la tarea de detección de la enfermedad de Parkinson a partir de registros de vocales sostenidas.

5.4. Aplicación 1: tipificación automática de voces

El análisis de perturbaciones es un área del estudio de la voz que se encarga de analizar los pequeños cambios que se producen ciclo a ciclo en la señal de voz. Esto se logra a partir de la definición y la estimación de diversos parámetros asociados a los cambios en la frecuencia, amplitud y forma de onda de la señal. Entre estos parámetros, destacan el *jitter*, *shimmer*, y la variabilidad de la forma de onda [158]. El desvío de estos parámetros respecto a valores de referencia asociados a sujetos sanos determinan el grado y tipo de afección que presenta el paciente. El análisis de los parámetros de perturbación se encuentra muy difundido en la práctica fonoaudiológica, debido principalmente a la existencia de herramientas de software, como PRAAT o MDVP, que permiten calcularlos fácilmente a partir de los registros de voz. Sin embargo, el cálculo de estos parámetros se basa en la suposición de que las señales de voz son periódicas, lo cual no es generalmente cierto. En el caso de no cumplir con la condición de periodicidad, la aplicación de estas medidas de perturbación no refleja necesariamente la condición real de la salud del paciente. Es por

5.4. APLICACIÓN 1: TIPIFICACIÓN AUTOMÁTICA DE VOCES

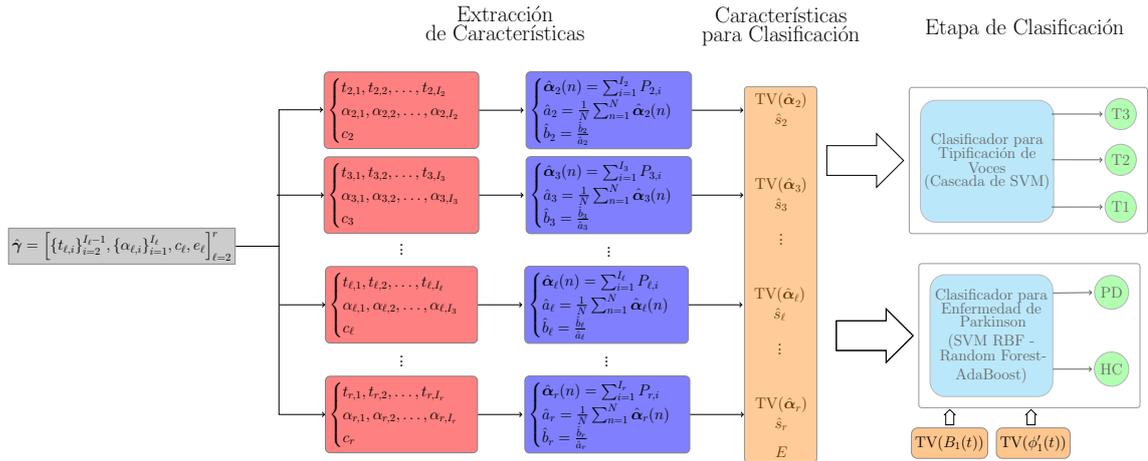


Figura 5.2: Diagrama de los pasos de extracción de características y clasificación de la tarea de tipificación de la voz.

esto que se suele realizar, como etapa previa al análisis de perturbaciones, un proceso de *tipificación* de voces.

Según Titze [171], las señales de voz pueden clasificarse en tres tipos en función de su periodicidad. Las señales de tipo 1 se definen como casi periódicas con una estructura armónica clara. Las señales de tipo 2 presentan una fuerte modulación y contienen frecuencias subarmónicas. Por último, las señales de tipo 3 carecen de estructura armónica perceptible. Bajo este esquema, las señales de tipo 1 puede estudiarse adecuadamente mediante el análisis de perturbaciones. Las señales de tipo 2 deben ser estudiadas preferentemente mediante espectrogramas de banda angosta o escalas perceptuales. Finalmente, las señales de tipo 3 deberán ser evaluadas únicamente mediante escalas perceptuales.

Una limitante del proceso de tipificación de señales de voz es la subjetividad inherente al proceso de clasificación realizado por el profesional de salud. Esto se debe a que las definiciones dadas por Titze no están ligadas a criterios cuantitativos, sino que requieren de la interpretación subjetiva por parte del individuo de las características de la señal. Es por esto que recientemente ha surgido interés en contar con medidas cuantitativas que permitan distinguir entre los diferentes tipos de voces.

A modo ilustrativo, la Fig. 5.3 muestra un ejemplo de cada tipo de señal, junto con su espectrograma correspondiente y las HAFs asociadas a cada una de las señales según el ANHM, considerando una WSF variable en el tiempo. Al examinar estas cantidades, se hace evidente que las señales de Tipo 3 presentan una mayor variabilidad ciclo a ciclo en su forma de onda que los otros dos tipos, y que la WSF de las señales de Tipo 2 es más variable que la de las señales de Tipo 1. Esto indica la posibilidad de realizar una tipificación automática de la voz basada en características asociadas a la variabilidad de la forma de onda, despreciando la modulación en amplitud y frecuencia.

Para validar esta propuesta, se entrenó un clasificador que combina distintos modelos de inteligencia computacional para realizar la tipificación de voces de manera automática, utilizando las medidas de variabilidad de la forma de onda como características para el entrenamiento y validación del clasificador. A continuación, se detallan los conjuntos de datos considerados, las características del clasificador propuesto y se discuten los resultados obtenidos.

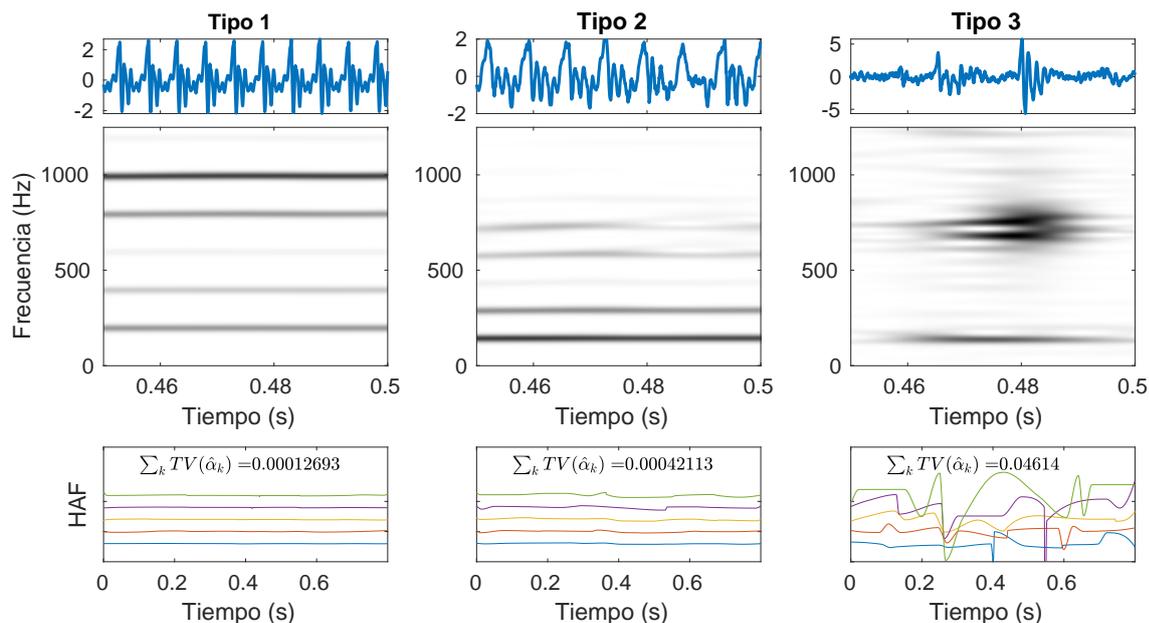


Figura 5.3: Señales de voces según la tipificación de Titze. Primera fila: Ejemplos de emisiones de /a/ sostenida para cada tipo de señal de voz. Segunda fila: espectrogramas correspondientes a cada vocal. Tercera fila: Funciones de amplitud armónica (desplazadas verticalmente por motivos de visibilidad) para cada señal de voz junto a la variación total acumulada para todas las HAFs.

5.4.1. Conjuntos de datos

Para el entrenamiento y validación de este clasificador, se utilizaron señales de voz de dos bases de datos distintas. La primera es la base de datos Massachusetts Eye and Ear Infirmary Voice Disorder (MEEI), distribuida por Kay Elemetrics [172]. Este corpus consta de unas 750 señales de voz que incluyen tanto voces normales como patológicas. Las grabaciones se realizaron a 25 kHz (condición patológica) y 50 kHz (condición normal), con 16 bits de resolución de amplitud. La segunda base de datos es la Saarbrücken voice database (SAAR), grabada por el Institut für Phonetik de la Universidad de Saarland y la Sección de Foniatría de la Caritas Clinic St. Theresia de Saarbrücken, Alemania [170]. Se trata de un corpus extenso, compuesto por más de 2000 señales de hablantes alemanes y que abarca una gama muy diversa de patologías. Las señales de este conjunto de datos se grabaron a una frecuencia de muestreo de 50 kHz.

Se llevó a cabo un proceso de selección de muestras en ambos conjuntos de datos. En primer lugar, se excluyeron las grabaciones con interrupciones y duración inferior a 800 ms. A continuación, se procesaron las señales seleccionadas de ambos conjuntos de datos para eliminar el inicio y el final de la emisión. Además, las señales muestreadas a 50 kHz se redujeron a 25 kHz. Las voces fueron etiquetadas como Tipo 1, 2 o 3 por profesionales expertos (autores tercero y cuarto de [158]) utilizando espectrogramas y medidas de perturbación como jitter y shimmer. Sólo se incluyeron en el estudio las señales para las cuales ambos expertos coincidieron en la clasificación. En total, se utilizaron 1262 voces en este estudio, cuya distribución por tipos y base de datos se muestra en la Tabla 5.1.

	Tipo 1	Tipo 2	Tipo 3	Total
SAAR	174	330	109	613
MEEI	185	335	129	649

Tabla 5.1: Distribución de los tipos de voz en ambos conjuntos de datos

5.4.2. Detalles del clasificador

Para esta tarea se consideraron varias configuraciones del clasificador, todas ellas con un desempeño similar. El mejor desempeño para la tipificación automática de voces se obtuvo mediante una serie de dos clasificadores basados en máquinas de soporte vectorial (SVM, por las siglas en inglés de *support-vector machine*) conectados en cascada. La primera etapa se implementa como un SVM con kernel lineal, entrenado para separar las señales de Tipo 3 (etiqueta T3) de las de Tipo 1 y Tipo 2 (etiqueta T1-2). A continuación, las características de las muestras clasificadas como T1-2 se introducen en la segunda etapa, que se implementa como un clasificador SVM con un kernel de funciones de base radial (RBF, por las siglas en inglés de *radial basis function*). Este clasificador se entrena para discriminar entre señales de Tipo 1 (T1) y Tipo 2 (T2).

Las características extraídas de las señales de voz se utilizaron para entrenar y validar la cascada de clasificadores SVM. Obtuvimos los parámetros mediante el algoritmo tvWSE para este experimento siguiendo el procedimiento descrito en el apartado anterior, obteniendo un valor de $r_{fixed} = 18$ componentes armónicos por señal e $I_{fixed} = 15$ nodos por armónico. Mediante el algoritmo propuesto, se obtuvo el vector de coeficientes óptimo $\hat{\gamma}$ y se calcularon el error de estimación, la variación total de las HAFs y los coeficientes de amplitud relativa. En total, se extrajeron $N_f = 36$ características para cada señal. La validación del modelo se realizó mediante un procedimiento de validación cruzada estratificada de 10 *folds*. En cada pliegue, se utilizó el 90 % de los datos para entrenar los clasificadores, y las medidas de desempeño se calcularon con el 10 % restante de los ejemplos (conjunto de validación). No hubo solapamiento entre los diferentes conjuntos de validación. Dado el desequilibrio de clases, se estratificaron los pliegues para que se respete la proporción de muestras para cada clase en el conjunto completo de datos en los diferentes pliegues durante el proceso de validación cruzada.

5.4.3. Medidas de desempeño y optimización de hiperparámetros

Para tener en cuenta el desequilibrio entre clases, se utilizaron la precisión balanceada y las matrices de confusión multiclase como medidas de desempeño de la clasificación. La precisión balanceada se obtiene como la media de la tasa de verdaderos positivos (TPR, del inglés *true positive rate*) por clase para todas las clases:

Conjunto de Datos	C_1	C_2	σ
MEEI	0.215	21.5	0.009
SAAR	0.1	215.4	0.0009
MEEI+SAAR	0.021	464	0.00028

Tabla 5.2: Hiperparámetros óptimos de los clasificadores para cada base de datos.

5.4. APLICACIÓN 1: TIPIFICACIÓN AUTOMÁTICA DE VOCES

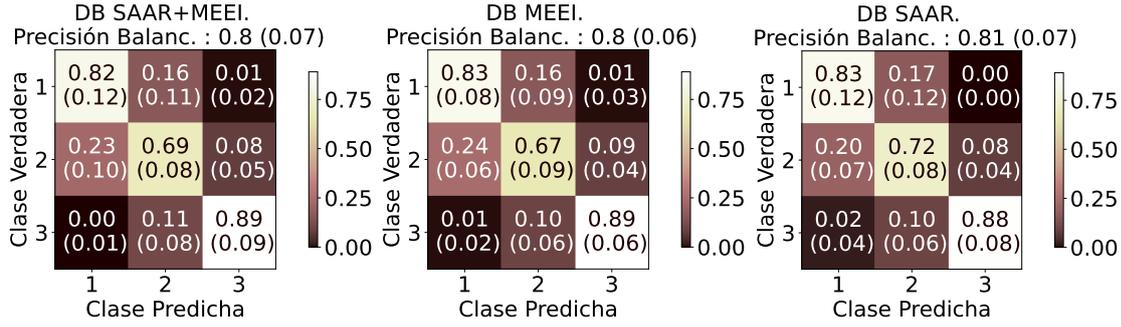


Figura 5.4: Resultados de la tarea de tipificación de voces. Izquierda : matriz de confusión promedio para la validación cruzada del clasificador del tipo de voz utilizando el conjunto de datos completo. Centro: Matriz de confusión promedio sólo para la base de datos MEEI. Derecha: Matriz de confusión promedio sólo para la base de datos SAAR.

$$\text{TPR}_k = \frac{\text{TP}_k}{\text{TP}_k + \text{FN}_k}, \quad (5.4)$$

$$\text{Balanced Acc} = \frac{1}{K} \sum_{k=1}^K \text{TPR}_k, \quad (5.5)$$

con $K = 3$ y los valores $k = 1, 2, 3$ indican las medidas de desempeño para los tipos de voz 1,2 y 3, respectivamente.

En cuanto a los hiperparámetros del clasificador, se determinaron las constantes de margen óptimas C_1 y C_2 de la primera y la segunda etapas, respectivamente, y el coeficiente óptimo σ para el kernel RBF. Los valores de estos hiperparámetros se optimizaron mediante una búsqueda en grilla para obtener la mayor precisión equilibrada media en todos los pliegues. La grilla se dividió en $L = 7$ segmentos para cada parámetro y los intervalos considerados, en escala logarítmica, para C_1 , C_2 y σ fueron, respectivamente, $[10^{-3}, 10]$, $[0,01, 1000]$ y $\left[\frac{0,01}{N_f}, \frac{10}{N_f}\right]$, siendo N_f el número total de características utilizadas para el entrenamiento. El procedimiento de optimización de hiperparámetros y validación del clasificador se realizaron para los conjuntos de datos MEEI y SAAR por separado, y también para el conjunto completo de datos. Los hiperparámetros óptimos para el clasificador en sus diferentes configuraciones se indican en la Tabla 5.2.

5.4.4. Resultados y discusión

En la Fig. 5.4 se presentan las matrices de confusión promediadas para todos los pliegues para los tres conjuntos de datos: SAAR, MEEI, y SAAR+MEEI. Además, se presenta la precisión balanceada promedio junto a su desviación estándar. Para el conjunto de datos completo, el desempeño global, medido por la precisión equilibrada promedio, fue de 80%. La TPR promedio para las señales de tipo 3 es de 89%, lo que indica un buen desempeño de la primera etapa de nuestro clasificador en cascada. La principal dificultad se encuentra en la segunda etapa encargada de discriminar entre las señales de Tipo 1 y Tipo 2. Esta condición se condice con la clasificación manual realizada por los profesionales de la salud vocal, quienes a menudo encuentran dificultad a la hora de diferenciar estos tipos de señales. La mayor desviación estándar de la TPR para los tipos

5.4. APLICACIÓN 1: TIPIFICACIÓN AUTOMÁTICA DE VOCES

Entrenamiento: SAAR+MEEI - Validación: SAAR+MEEI									
	Miramont20 [158]			Miramont22 [159]			Nuestra propuesta		
	1	2	3	1	2	3	1	2	3
1	0,86 (0,04)	0,14 (0,04)	0,00 (0,00)	0,76 (0,11)	0,24 (0,11)	0,00 (0,00)	0,82 (0,12)	0,16 (0,11)	0,01 (0,02)
2	0,15 (0,05)	0,78 (0,04)	0,07 (0,02)	0,09 (0,04)	0,86 (0,04)	0,05 (0,03)	0,23 (0,10)	0,69 (0,08)	0,08 (0,05)
3	0,00 (0,00)	0,08 (0,03)	0,92 (0,03)	0,00 (0,00)	0,21 (0,11)	0,79 (0,11)	0,00 (0,01)	0,11 (0,08)	0,89 (0,09)
	Precisión Balanceada: 0.85 (0.03)			Precisión Balanceada: 0.8 (0.06)			Precisión Balanceada: 0.8 (0.07)		

Tabla 5.3: Comparación de nuestros resultados con el estado del arte para la tarea de tipificación por voz.

1 y 2 indica un desempeño variable entre los pliegues, con algunos pliegues mostrando un buen desempeño mientras otros tienen un desempeño menos eficaz. Cabe destacar que la confusión entre las señales de tipo 1 y 3 es mínima en este conjunto de datos. Se trata de una condición favorable, ya que las señales de Tipo 3 no pueden analizarse con las medidas de perturbación que se utilizan normalmente para las señales de Tipo 1.

Los resultados de la tarea de clasificación sólo para la base de datos MEEI se muestran en el panel central de la Fig. 5.4, donde observamos que se alcanza un desempeño similar en comparación con el conjunto de datos completo, con una TPR más alta para las señales de Tipo 1 y una TPR ligeramente más baja para las señales de Tipo 2. Además, la varianza global entre los folds es menor para esta base de datos.

El panel de la derecha de la Fig. 5.4 muestra el desempeño de clasificación resultante en la base de datos SAAR. Para este conjunto de datos, se observa un desempeño global ligeramente superior, alcanzando una precisión balanceada promedio de 81%. Esta mejora se atribuye principalmente a una mejor discriminación entre las señales de Tipo 1 y Tipo 2, con una ligera disminución en el desempeño de la primera etapa en la tarea de discriminar las señales de Tipo 3 de las otras clases pero manteniendo una desviación estándar similar.

5.4.5. Comparación con el estado del arte

El análisis del estado del arte muestra que la tipificación automática de señales de voz sigue siendo un campo poco investigado. En el estudio de Miramont y cols.[158], los autores diseñaron un clasificador automático de señales de voz basado en la escala de periodicidad de Titze, utilizando SVMs. En este trabajo combinaron características obtenidas a partir del análisis de perturbación de periodicidad con medidas de la dinámica no lineal de las señales como dimensión de correlación y entropía de correlación. En un trabajo posterior [159], se propuso un nuevo conjunto de características, obtenidas a partir de la transformada *scattering* (ST) [173], y la clasificación se llevó a cabo de nuevo utilizando SVMs. Se utilizó el mismo conjunto de datos para ambos trabajos y para este artículo, lo que permitió una comparación directa.

Las matrices de confusión de la base de datos completa, publicadas por estos autores, se muestran en la Tabla 5.3 junto a nuestros resultados. Al comparar nuestro método con el clasificador basado en ST de [159], nuestro clasificador muestra una mayor TPR para las señales de Tipo 1 y Tipo 3. Sin embargo, nuestro clasificador obtiene peores resultados en la clasificación de las señales de Tipo 2, que se confunden principalmente con las de Tipo 1. En comparación con nuestro método y los resultados de [159], los resultados de [158] son comparables dada la varianza de la medida de precisión para los tres estudios,

aunque los valores promedios son ligeramente mayores en este último caso. No obstante, una ventaja notable de nuestra propuesta radica en que las características para el entrenamiento del clasificador se derivan de un modelo fenomenológico de la señal bajo estudio. Estas características están intrínsecamente relacionadas con la variabilidad de la forma de onda de la señal y la amplitud relativa de los armónicos en el dominio tiempo-frecuencia. La interpretabilidad de estas medidas de variabilidad se vincula estrechamente al análisis visual que los clínicos realizan al examinar los espectrogramas de las señales vocales. Nuestro modelo puede cuantificar las variaciones intrínsecas de las amplitudes armónicas de las señales y los resultados muestran que se pueden obtener resultados similares a los de otros conjuntos de características menos interpretables.

Es importante señalar que, aunque nuestras particiones de datos difieren de las utilizadas en los métodos con respecto a los cuales estamos comparando, esta comparación sigue siendo válida debido al uso de la validación cruzada, que reduce considerablemente el error debido al muestreo de datos.

5.5. Aplicación 2: detección de enfermedad de Parkinson

La enfermedad de Parkinson (PD, del inglés *Parkinson's Disease*) es un trastorno neurodegenerativo cada vez más prevalente en el mundo [174]. La principal afección asociada a esta enfermedad es la muerte progresiva de las neuronas encargadas de coordinar las funciones motoras del organismo. Un aspecto notable de la PD es la presencia de alteraciones vocales, que a menudo surgen en las etapas tempranas de la enfermedad. Estas alteraciones incluyen un volumen vocal reducido, un rango de frecuencia fundamental limitado, pronunciación imprecisa de las vocales, respiración entrecortada y pausas irregulares en el habla [175]. La presencia de estas alteraciones tiene un impacto considerable en la calidad de vida de las personas que sufren de PD, viendo fuertemente afectada su actividad social personal y profesional [176]. Es necesario entonces obtener marcadores clínicos de la presencia y severidad de la PD para poder tomar decisiones acordes respecto al tratamiento, con el objetivo de mitigar lo más posible los aspectos más perjudiciales de esta enfermedad. La identificación y el análisis de las anomalías del habla surge como una alternativa interesante para diferenciar a los pacientes con PD de los individuos sanos [177]. Además, la detección precoz de la PD mediante el análisis vocal podría facilitar la intervención temprana y mejorar los resultados de los procedimientos terapéuticos aplicados a los pacientes.

Debido a la naturaleza de las afecciones vocales asociadas a la PD, la inestabilidad y perturbación en la generación de la voz resulta ser un indicador útil para detectar la presencia de la enfermedad. Para ver esto, en la Fig. 5.5 se muestran emisiones de la vocal /a/ sostenida para un paciente con PD y un individuo de control, respectivamente. Al igual que con la tipificación de señales de voz, la variabilidad de la forma de onda interciclo es considerable en el caso patológico. Entonces, las medidas de variabilidad de forma de onda pueden resultar útiles para discriminar entre los dos grupos. Para verificar esta suposición, proponemos evaluar diferentes clasificadores entrenados para la tarea de detección automática de la presencia de PD a partir de registros de vocales sostenidas, basados en las medidas de variabilidad de la forma de onda propuestas en la Sec. 5.3. En las siguientes secciones detallaremos los aspectos de diseño de estos clasificadores, discutiremos el conjunto de datos utilizado para su entrenamiento y validación. Finalmente, discutiremos los resultados obtenidos y las perspectivas a futuro para esta propuesta.

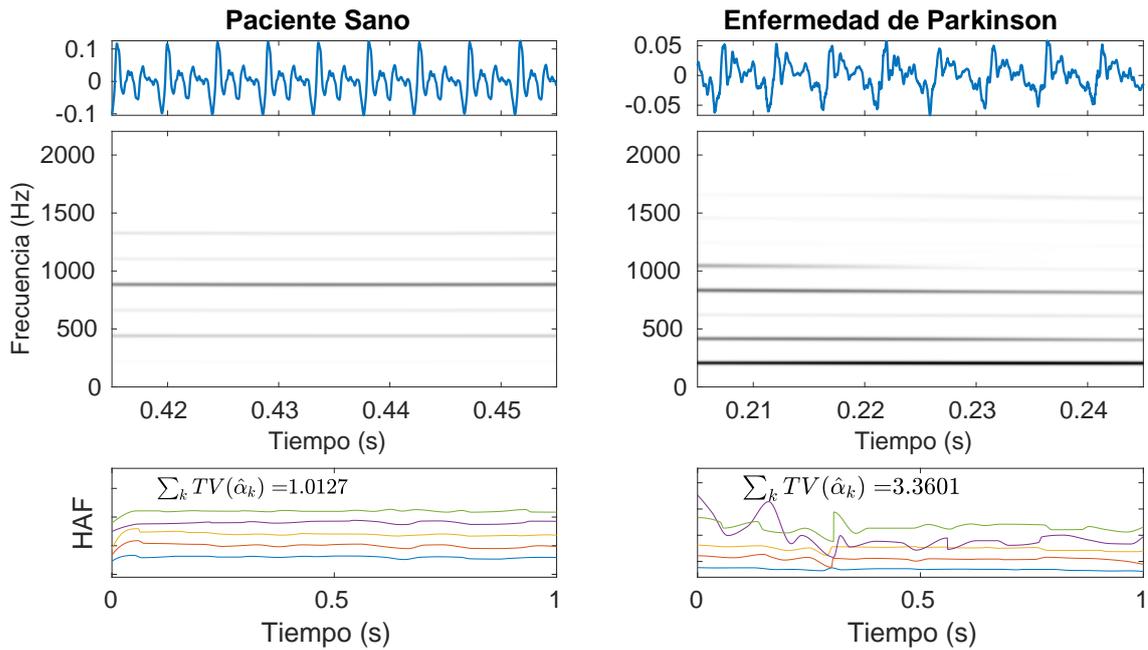


Figura 5.5: Señales de voz para detección de Enfermedad de Parkinson (PD). Primera fila: Ejemplos de señales de voces con forma de onda variable para el grupo de control y el de PD, respectivamente. Segunda fila: espectrograma correspondiente a cada señal de voz. Tercera fila: Funciones de amplitud armónica (desplazadas verticalmente por motivos de visibilidad) para cada señal de voz junto a la variación total acumulada para todas las HAFs.

5.5.1. Conjuntos de datos

Para desarrollar los clasificadores propuestos para la detección de la PD, utilizamos un conjunto de datos de habla española descrito originalmente en [178], bajo el nombre PC-GITA. Este conjunto de datos comprende muestras de habla recolectadas de una cohorte de 100 participantes (50 controles sanos y 50 pacientes con PD) que realizaron diversas tareas de habla. Todos los pacientes fueron diagnosticados por neurólogos expertos y categorizados según las escalas UPDRS (*Unified Parkinson's Disease Rating Scale*) y H&Y (Clasificación por estadios de Hoehn y Yahr). Los pacientes con PD fueron registrados en su estado “ON”, es decir, dentro de las 3 horas siguientes a la toma de su medicación matutina. Los participantes del grupo de control no mostraban ninguna sintomatología asociada a la PD ni a otros trastornos neurológicos en los cribados previos.

Los registros se obtuvieron a una frecuencia de muestreo de 44,1 kHz con una resolución de 16 bits. El proceso de recolección de datos se llevó a cabo en un entorno insonorizado bajo supervisión profesional. La duración media de los enunciados fue de 2,93 segundos. La edad de los participantes oscilaba entre los 31 y los 86 años. Para este estudio, nos centramos en los enunciados de vocales sostenidas. Se grabaron tres repeticiones de cada vocal por participante, lo que dio como resultado un total de 150 señales para cada tipo de vocal dentro del grupo de pacientes. En las grabaciones de más de 1 segundo de duración, sólo se tuvo en cuenta el primer segundo de la señal para su análisis.

5.5.2. Detalles del clasificador

Los clasificadores considerados para la tarea de detección de PD se entrenaron y validaron utilizando características extraídas de emisiones de vocales /a/ y /e/ sostenidas de la base de datos PC-GITA. Se eligieron estas vocales porque son las de uso más frecuente en la práctica clínica y además muestran una considerable variabilidad en la forma de onda en los casos patológicos. Sumado a esto, dada la naturaleza de las características obtenidas a partir del modelo de forma de onda variable, las medidas de variabilidad asociadas de diferentes vocales pueden considerarse conjuntamente para el entrenamiento y la validación de los clasificadores. Para el procedimiento de extracción de características, los parámetros del algoritmo tvWSE se fijaron en $r_{fixed} = 15$ armónicos por señal con $I_\ell = 20$ nodos por armónico según la metodología descrita en la Sec. 5.3. El número total de parámetros para cada señal fue $N_f = 32$. En cuanto a la implementación de los clasificadores, se evaluaron varios algoritmos de aprendizaje automático durante esta etapa. Además, se consideraron dos esquemas diferentes de validación cruzada: validación cruzada con k particiones (kFCV) y validación cruzada dejando un sujeto fuera (LOSOCV, del inglés *leave-one-subject-out cross-validation*). En el enfoque kFCV, el conjunto de datos se divide en k subconjuntos, con muestras de $k - 1$ subconjuntos utilizados para generar el conjunto de entrenamiento, y el desempeño del clasificador se evalúa en el subconjunto restante.

En cambio, en LOSOCV, cada sujeto se excluye sucesivamente del conjunto de entrenamiento y el clasificador se entrena con los datos de los sujetos restantes. El desempeño se evalúa sobre las muestras del sujeto excluido. Este enfoque maximiza el uso de la información durante el entrenamiento, mejorando la generalizabilidad del clasificador para los diferentes sujetos.

En cuanto a los clasificadores, se consideraron tres arquitecturas distintas: una máquina de soporte vectorial con kernel de función de base radial (SVM RBF), un clasificador de bosque aleatorio (RF, por las siglas en inglés de *Random Forest*) y un ensamble de clasificadores que emplea un conjunto de árboles truncados ensamblados mediante el algoritmo de *Boosting* adaptativo (AdaBoost). Para optimizar el desempeño de cada clasificador, se realizó una búsqueda en grilla de los hiperparámetros óptimos.

En el caso de SVM, los hiperparámetros considerados fueron la constante de margen C y el parámetro de kernel RBF σ . En el caso de RF, la búsqueda en grilla se realizó sobre el número de estimadores n_c y el número mínimo de muestras en cada rama terminal n_f . En el caso de AdaBoost, se debe optimizar el número de clasificadores del ensamble n_c .

5.5.3. Medidas de desempeño y optimización de hiperparámetros

En ambos enfoques de validación cruzada, se tuvo especial cuidado en evitar la inclusión de muestras del mismo paciente tanto en el conjunto de entrenamiento como en el de validación. Esta precaución se tomó para mitigar cualquier posible sesgo generado por la correlación entre las muestras de un mismo paciente.

En el caso del esquema kFCV, las medidas de desempeño elegidas son las matrices de precisión y confusión, promediadas sobre todos los *folds* junto con la desviación estándar correspondiente. Dado que el conjunto de datos utilizado está balanceado (mismo número de muestras para cada clase), no se realizó ninguna corrección de desbalance durante el entrenamiento ni al calcular las medidas de desempeño.

Para el esquema LOSOCV, el desempeño se evalúa mediante la precisión a nivel de

5.5. APLICACIÓN 2: DETECCIÓN DE ENFERMEDAD DE PARKINSON

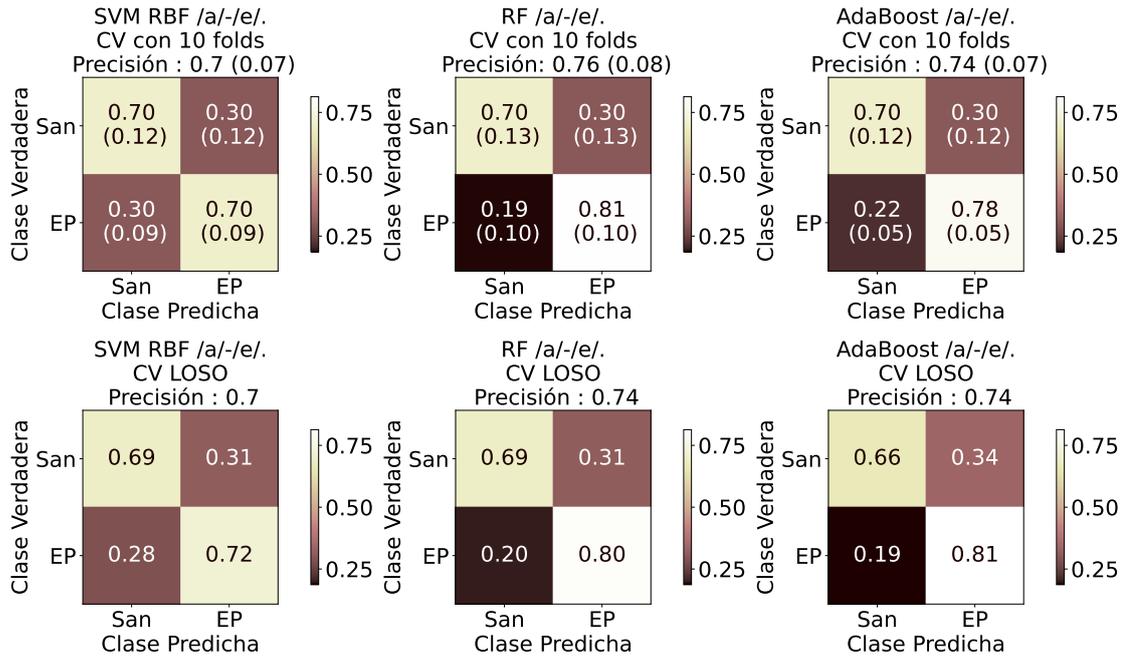


Figura 5.6: Resultados de detección de enfermedad de Parkinson. Fila superior: Matriz de confusión promedio para la validación cruzada de las tres arquitecturas para la detección de PD utilizando las emisiones vocálicas /a/ y /e/ con el esquema de kFCV. Fila inferior: Matriz de confusión promedio para la validación cruzada de las tres arquitecturas para la detección de PD utilizando las emisiones vocálicas /a/ y /e/ con el esquema LOSOCV.

los pacientes. Para cada paciente, todas las señales correspondientes se clasifican de forma independiente y al paciente se le asigna la etiqueta de clase más frecuente para ese subconjunto de señales. Por último, la precisión por paciente se calcula como el número de pacientes clasificados correctamente dividido por el número total de pacientes ($M = 50$).

La optimización de los hiperparámetros se realizó de forma análoga a la tarea de tipificación de voz, realizando una búsqueda en grilla en el espacio de hiperparámetros. Para el clasificador SVM, se realizaron $L = 13$ particiones para cada parámetro y los intervalos para C y σ fueron $[0, 1, 10^5]$ y $\left[\frac{10^{-4}}{N_f}, \frac{1000}{N_f}\right]$, respectivamente. Para el clasificador RF, el intervalo para el parámetro n_c es de $[2, 50]$, y para el parámetro n_f es de $[10, 500]$. Por último, para el clasificador AdaBoost, el intervalo del parámetro n_c es de $[10, 200]$. Los hiperparámetros óptimos obtenidos para cada clasificador se describen en la Tabla 5.4.

CV	SVM		RF		AdaBoost
	C	σ	n_c	n_f	n_c
10-fold CV	100	0.005	200	2	10
LOSOCV	1000	0.0009	100	10	10

Tabla 5.4: Hiperparámetros óptimos de los 3 clasificadores para la tarea de detección de la enfermedad de Parkinson. Se muestran los valores óptimos obtenidos para el conjunto de datos combinados de vocales /a/ y /e/ de la base de datos PC-GITA.

5.5. APLICACIÓN 2: DETECCIÓN DE ENFERMEDAD DE PARKINSON

Estudio	Señales	Características	Clasificador	Desempeño (Precisión)
Orozco-Arroyave (2014) [178]	Vocal /a/ Sostenida	Medidas de Estabilidad y Periodicidad	SVM	91 %
Rueda (2019) [180]	DDK y Vocal /a/ Sostenida	Medidas de Fonación	RF	73 %
Karan (2020) [181]	Vocales /a/ y /o/ Sostenidas	Medidas basadas en IMF	RF	100 %
Kadiri (2020) [182]	Vocales Sostenidas	Coefficiente Cepstral de Filtrado de Frecuencia Singular	SVM	73 %
Nuestro Estudio	Vocales /a/ y /e/ Sostenidas	Medidas de Variabilidad de Forma de Onda	RF	76 %

Tabla 5.5: Comparación de nuestra propuesta con el estado del arte en detección de PD.

5.5.4. Resultados y discusión

En la fila superior de la Fig. 5.6 mostramos las matrices de confusión promedio para los tres clasificadores utilizando el esquema kFCV con $k = 10$ *folds* al considerar la combinación de las emisiones vocálicas /a/ y /e/. Vemos que el mejor desempeño se obtiene con el clasificador RF, alcanzando una tasa de precisión promedio del 76 %. Los clasificadores SVM y AdaBoost alcanzan un desempeño inferior de 70 % y 74 %, respectivamente. Además, el clasificador RF puede identificar correctamente los casos patológicos en el 81 % de los casos.

En la fila inferior de la Fig. 5.6 se muestran los resultados para el esquema LOSOCV. En este esquema, la precisión promedio a nivel de los pacientes para cada clase se muestra en la diagonal de las matrices de confusión. Se observa un desempeño similar en comparación con el esquema kFCV, con la particularidad de que el clasificador AdaBoost alcanzó un desempeño similar al clasificador SVM. En ambos casos, RF es el algoritmo con mejor desempeño, alcanzando un 74 % de precisión a nivel de los pacientes en el esquema LOSOCV.

5.5.5. Comparación con el estado del arte

La detección de la PD basada en el habla es un campo muy estudiado y se han propuesto muchos métodos automáticos para llevar a cabo esta tarea. Muchos de estos algoritmos se centran en la elección de un conjunto relevante de características para la discriminación entre las señales de los grupos de PD y de pacientes sanos (control). Se han aplicado a esta tarea diversos modelos de aprendizaje maquina (ML, del inglés *machine learning*), como por ejemplo SVMs, árboles de decisión, RFs y modelos de mezcla de gaussianas, entre otros. En la presentación inicial de la base de datos PC-GITA [178], se entrenó un clasificador SVM utilizando características acústicas obtenidas de la vocal sostenida (/a/), logrando una precisión de aproximadamente 91 %. Un trabajo posterior [179], mejoró esta precisión hasta el 97 % entrenando un clasificador SVM similar sobre el comando diadococinético (DDK) /pa-ta-ka/, extrayendo una serie de características acústicas, espectrales y cepstrales.

Avances recientes en la detección de PD incluyen la aplicación de enfoques basados en el análisis tiempo-frecuencia. En Villa y cols. [183], los autores proponen utilizar características extraídas de la representación tiempo-frecuencia de señales de habla continua para entrenar un clasificador SVM, logrando una precisión del 72 %. En [180], se empleó EMD, como extractor de características. Un procedimiento similar fue seguido por los autores en [181], donde descompusieron las señales de voz en IMFs a través de EMD y

utilizaron características basadas en estas IMFs para entrenar una variedad de clasificadores incluyendo SVM, RF y árboles de decisión, reportando en algunos casos una precisión de 100 % para la clasificación de voces sanas y patológicas.

Más recientemente, se han desarrollado métodos basados en modelos de aprendizaje profundo para la tarea de detección de PD, algunos de estos modelos incluyen redes neuronales profundas (DNN, por las siglas en inglés de *deep neural networks*), redes neuronales convolucionales (CNN, por las siglas en inglés de *convolutive neural networks*) y redes neuronales recurrentes (RNN, por las siglas en inglés de *recurrent neural networks*). En Quan y cols. [184], un modelo de aprendizaje profundo de extremo a extremo entrenado sobre la base de datos PC-GITA utilizó una red CNN bidimensional para capturar la información espacial local del espectrograma y como segunda etapa se usó una CNN unidimensional para aprender las dependencias temporales. Otro estudio [185] combinó una CNN y una RNN para la detección de PD utilizando señales de habla continua, incorporando el procedimiento de descomposición variacional en modos (VMD, de las siglas en inglés de *variational model decomposition*) como herramienta de limpieza de ruido.

Dada la diversidad de enfoques adoptados para la tarea de clasificación de voces en pacientes de PD, diferentes dificultades y limitaciones han sido observadas en la práctica. Por un lado, no existe aún un consenso sobre las características más eficaces para la detección de la PD. Además, la falta de un marco coherente entre los estudios complica las comparaciones directas debido a las variaciones en los tipos de señales, los esquemas de validación cruzada y las arquitecturas de los clasificadores. En general, las señales más complejas, como las frases o el habla en movimiento, demuestran ser superiores a la hora de discriminar entre los grupos de PD y de control. Por ejemplo, las dificultades articulatorias asociadas a la PD se hacen más evidentes durante las transiciones de fonemas, como demostró Orozco-Arroyave [179]. Por otro lado, los enfoques de aprendizaje profundo, aunque ofrecen un desempeño competitivo, presentan desafíos como tiempos de entrenamiento extensos y la necesidad de muchas más muestras en comparación con los clasificadores tradicionales para obtener un desempeño adecuado.

Debido a lo anterior, y con el objetivo de llevar a cabo una comparación justa de nuestro método, evaluamos el desempeño de nuestra propuesta exclusivamente frente a clasificadores que utilizan vocales sostenidas como entrada, centrándonos en los modelos de ML clásicos. Esta comparación se resume en la Tabla 5.5.

A partir de estos resultados, vemos que nuestra propuesta supera a los clasificadores implementados en [182] y [180], que se basan en características fonatorias y coeficientes cepstrales de filtrado de frecuencia única, respectivamente. Vemos que aunque las propuestas de [178] y [181] pueden alcanzar desempeños superiores a nuestra propuesta, se basan en rasgos más generales que no pueden relacionarse directamente con un modelo fenomenológico de las señales. El primer trabajo se basa en características clásicas de periodicidad como el shimmer y el jitter, junto con características de estabilidad de fonación como el cociente de perturbación de amplitud y el cociente de perturbación de periodo de tono. El segundo se basa en la descomposición de las señales de voz en funciones modales intrínsecas, que se obtiene como resultado de aplicar el algoritmo de EMD, el cual es un procedimiento de descomposición guiado por los datos que carece de un marco matemático sólido. Cabe señalar que la modelización no estacionaria de las señales de voz, en particular para segmentos sonoros como las vocales sostenidas, es un enfoque bastante novedoso para la modelización de señales de voz. Nuestra propuesta es un primer paso en la aplicación de modelos no estacionarios como herramientas de análisis para señales

con características variables en el tiempo, y el uso de estas características como entradas de un sistema clasificador del estado de la voz requiere una mayor exploración, tanto en estas tareas como en otras similares.

5.6. Comentarios de fin de capítulo

En este capítulo discutimos la aplicación del modelo no armónico adaptativo para el análisis de señales de voz. Éstas pueden ser estudiadas eficazmente mediante este modelo al tratarse de señales no estacionarias con características variantes en el tiempo. Propusimos un conjunto de características obtenidas a partir de un procedimiento de estimación la función de forma de onda variable en el tiempo de la señal de voz. Estas características permiten codificar las variaciones temporales de la forma de onda de señales de voz, tales como las emisiones de vocales sostenidas. Esta variabilidad puede utilizarse como un indicador del estado del paciente a partir de cuantificar la variabilidad de la señal vocal. La utilidad de estas características fue validada para dos tareas de análisis del estado de la voz, a saber, la tipificación automática de la voz y la detección de la enfermedad de Parkinson. Ambas tareas se llevaron a cabo a partir de clasificadores entrenados en base al conjunto de características de forma de onda variante en el tiempo.

Los resultados para ambas tareas muestran que las características propuestas en esta tesis pueden diferenciar señales según la variabilidad de la forma de onda entre ciclos, permitiendo alcanzar un desempeño competitivo con el estado del arte para ambas tareas. Además, los resultados demuestran que es posible utilizar características fácilmente interpretables para discriminar las voces según su grado de armonicidad. Particularmente para la tarea de detección de PD, los resultados muestran una tendencia alentadora en cuanto a la capacidad de diferenciar las voces sanas de las emisiones patológicas. En lo que respecta a esta tarea, se plantean trabajos futuros analizando otras bases de datos en diferentes idiomas y aplicando diferentes técnicas para mejorar la generalización de los clasificadores.

Por último, cabe destacar que las medidas de variabilidad de la forma de onda podrían ser utilizadas para otras tareas de clasificación de voces en las que el análisis de las variaciones ciclo a ciclo de la forma de onda de la señal puede conducir a una mejor separación entre clases.

El estudio de las características propuestas y los resultados para las tareas de clasificación presentados en este capítulo fueron incluidos en un artículo científico [5] que se encuentra actualmente en revisión para publicarse en la revista *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, de publicación conjunta entre la editorial del IEEE y la *Association for Computing Machinery* (ACM).

Capítulo 6

Conclusiones y trabajo a futuro

Esta tesis se enfocó en el estudio de modelos fenomenológicos para señales oscilatorias no estacionarias. Dichos modelos permiten caracterizar eficazmente las diferentes fuentes de variación temporal en la amplitud, frecuencia y patrón oscilatorio en este tipo de señales. A continuación, detallaremos los diferentes aportes realizados en esta tesis para analizar este tipo de señales, discutiremos los resultados obtenidos y su implicancia en la disciplina del estudio de señales biomédicas. Finalmente, propondremos líneas de trabajo que aprovechen los desarrollos propuestos en esta tesis tanto desde el punto de vista metodológico como desde el punto de vista de las aplicaciones biomédicas.

6.1. Conclusiones

En primer lugar, se establecieron las bases teóricas y metodológicas para describir los aportes originales propuestos en esta tesis. Se discutió el concepto de sistema dinámico y las señales no estacionarias que generan este tipo de sistemas. Se hizo hincapié en la presencia de este tipo de sistemas y señales en el ámbito biomédico, el cual es de especial interés debido al gran número de señales diferentes que pueden estudiarse mediante este enfoque. Se detallaron las diferentes estrategias de modelado fenomenológico existentes para señales no estacionarias, detallando las ventajas y debilidades de cada uno de ellos. Se detallaron las condiciones del modelo no armónico adaptativo y se definió la función de forma de onda para capturar la forma de la oscilación de las señales. El ANHM junto a la WSF conformaron la base conceptual de todos los métodos desarrollados en esta tesis. Finalmente, se establecieron las líneas de trabajo que se realizaron en esta tesis, junto a los objetivos perseguidos y los métodos y aplicaciones propuestos para cumplirlos.

Como un primer aporte, se presentó un método automático para la estimación del número de armónicos necesarios para representar las WSFs de señales no estacionarias. Para esto, se extendió el uso de criterios de selección de modelos trigonométricos, considerados originalmente para series temporales estacionarias, al estudio de señales moduladas en amplitud y en frecuencia. Se desarrolló una propuesta algorítmica basada en estos criterios para el análisis guiado por los datos de las señales no estacionarias mediante el ANHM. La aplicabilidad de estos criterios se validó extensivamente en señales sintéticas, bajo diferentes condiciones de forma de onda y niveles de ruido. También se extendió el uso de estos criterios de selección al caso de señales multicomponentes, donde cada componente presenta un patrón oscilatorio característico. Posteriormente, se aplicó el algoritmo propuesto a diferentes tareas de procesamiento en señales biomédicas. En primer

lugar, se analizaron señales de onda de pulso contaminadas con diferentes niveles de ruido, y se demostró que el algoritmo propuesto es capaz de recuperar eficazmente la señal original a partir de los registros ruidosos. También se demostró, mediante un análisis estadístico, que la estimación automática del número de armónicos para la WSF es un factor esencial para obtener un rendimiento óptimo del algoritmo en esta tarea. Finalmente, se aplicó el algoritmo de estimación de WSFs a señales biomédicas reales de una base de datos que contiene señales cardíacas y respiratorias contaminadas con diferentes fuentes de ruido y corrimiento de la línea de base. Los resultados sobre estas señales aportaron más evidencia al hecho que las técnicas de análisis y procesamiento de señales no estacionarias basadas en el ANHM permiten realizar tareas de limpieza de ruido en señales considerando las diferencias existentes en el patrón oscilatorio de las señales de diferente naturaleza.

Una limitación del ANHM en su formulación original es que considera WSFs fijas en el tiempo. Para superar esta limitante, se ha extendido el ANHM para considerar WSFs variantes en el tiempo, caracterizadas mediante la variación temporal de las amplitudes y fases de los componentes armónicos para cada WSF. Sin embargo, la estimación guiada por los datos de estos parámetros de las WSFs variantes en el tiempo sigue siendo un área poco explorada en la disciplina. Por esto, se propuso en un método de estimación adaptativo de las WSFs variantes en el tiempo, el cuál es capaz de caracterizar la variabilidad temporal de la forma de onda mediante la proporcionalidad no entera de las fases armónicas y mediante funciones de variación relativa de la amplitud armónica. El método propuesto fue descrito detalladamente, haciendo hincapié en la estimación automática de todos los parámetros necesarios para poder aplicarlo adecuadamente a las señales no estacionarias. Este algoritmo fue validado mediante señales sintéticas para las tareas de limpieza de ruido, descomposición de señales y detección de los instantes de cambio de la forma de onda. Además, este algoritmo fue aplicado en diferentes señales biomédicas en estas mismas tareas, demostrando así la aplicabilidad del método en situaciones reales del ámbito biomédico.

Una tarea importante en el estudio de señales biomédicas es el tratamiento de datos faltantes. Este fenómeno se suele presentar regularmente en la adquisición de estas señales, ya que los entornos de registro suelen involucrar sujetos sometidos a movimientos o expuestos a fuentes de ruido e interferencia que degradan la calidad de los datos. Para tratar esta situación, propusimos en esta tesis un método de imputación de datos faltantes basado en el ANHM. Este método se basa en la descomposición de una imputación inicial de la señal en sus componentes armónicos y de tendencia para luego realizar una interpolación de la tendencia y de las amplitudes y fases armónicas con el objetivo de mejorar la aproximación obtenida mediante el resultado de imputación. En este aspecto, el algoritmo resulta altamente versátil, ya que admite diferentes procedimientos de imputación inicial y sistemáticamente otorga una mejora significativa sobre este primer resultado. El buen desempeño de nuestro algoritmo fue validado tanto en señales sintéticas que se rigen por el ANHM, como también señales reales del ámbito biomédico.

El último aporte original de esta tesis fue la aplicación del ANHM con WSF variante en el tiempo al análisis de señales de voz. Para esto, se caracterizó la variabilidad de la forma de onda de las señales de vocales sostenidas mediante medidas obtenidas a partir del resultado del algoritmo de estimación de WSFs variable en el tiempo. Dichas medidas fueron luego utilizadas para conformar un conjunto de características para el entrenamiento y validación de sistemas de análisis automático de la voz basados en algoritmos

de aprendizaje automático. En particular, se aplicaron estas medidas de variabilidad a la tarea de tipificación de voces y la detección de la enfermedad de Parkinson. En ambos casos, los resultados obtenidos por nuestro método fueron competitivos con el estado del arte. Además, las medidas obtenidas a partir del modelo fenomenológico con WSF en el tiempo tienen una interpretabilidad clara en el contexto del análisis de voces sujetas a perturbaciones por la inestabilidad del aparato fonador presente en situaciones patológicas. El desempeño obtenido resulta muy prometedor y da la pauta de que este modelo puede ser aplicado al análisis de voces en diferentes contextos.

6.2. Trabajos a futuro

La conclusión de esta tesis sienta las bases para continuar avanzando sobre el desarrollo de las técnicas basadas en modelos fenomenológicos de señales no estacionarias, así como también aplicar dichos métodos para diferentes tareas de procesamiento de señales en distintos ámbitos, incluyendo la ingeniería biomédica.

Respecto a la primera propuesta de la tesis, que implica el uso de criterios de selección de modelos para estimar el número óptimo de armónicos para representar la WSF de una señal no estacionaria, existen diversas cuestiones en las que se puede profundizar. En primer lugar, se debería llevar adelante una análisis más detallado de las bases matemáticas y estadísticas de los criterios de selección, buscando describir las condiciones de las señales y criterios que garanticen resultados adecuados. Respecto a esto, se pueden analizar las propiedades asintóticas de los criterios, es decir, a medida que el número de muestras tiende a infinito. Por otro lado, los criterios descritos en esta tesis no son los únicos que se han desarrollado para series trigonométricas, sino que existen muchos otros como el propuesto por Hannan en [79] y el criterio de Schwarz [68]. La posibilidad de estudiar el comportamiento de nuevos criterios de selección, o bien proponer criterios a medida de las señales no estacionarias consideradas en esta tesis, es una línea de trabajo interesante para desarrollar a futuro.

En contraposición, también es factible mantener los criterios de selección considerados hasta ahora y modificar la familia de modelos sobre la cual se realiza la selección. En este sentido, sería posible aplicar los mismos criterios ahora sobre modelos que consideren formas de onda variantes en el tiempo, como ser el modelo SAMD o el modelo propuesto en la Sec. 3.3.1 de esta tesis. Respecto a esta idea, una de las principales limitantes es el tamaño de este conjunto de búsqueda y el número de parámetros que deben ser considerados. Modelos más complejos, como tvWSE o SAMD, presentan un mayor número de parámetros como, por ejemplo, el número de nodos I_ℓ o bien el grado del polinomio para las funciones de fase en el modelo SAMD. La necesidad de ajustar un nuevo modelo para cada combinación posible de parámetros vuelve al procedimiento muy costoso computacionalmente y dificulta su aplicación en la práctica.

En parte por lo anterior es que surge una de las primeras líneas de desarrollo a futuro para el método tvWSE propuesto en el Cap. 3, que es optimizar el algoritmo de optimización para reducir los tiempos de convergencia. En este aspecto, el algoritmo de estimación de la forma de onda variante en el tiempo propuesto en esta tesis tiene la desventaja de ser considerablemente más lento en obtener un resultado óptimo respecto a los algoritmos basados en el problema de regresión lineal por mínimos cuadrados presentado en (1.35) o bien el problema de regresión no lineal para el método SAMD en (3.8). Entre las posibilidades para agilizar la convergencia del algoritmo, está la posibilidad de realizar una

optimización por subconjuntos de parámetros, así por ejemplo se podría alternar entre la optimización de la amplitud y ubicación de los nodos para estimar las HAFs, y los coeficientes de fase c_ℓ y e_ℓ para describir la variación temporal de la fase de los armónicos. En este caso, la calidad del ajuste podría verse afectada al considerar este esquema de optimización, por lo cual se deberá tener en cuenta la relación entre tiempo de convergencia y la precisión del resultado obtenido a la hora de evaluar estos procedimientos.

Por otro lado, también es factible considerar métodos de optimización más sofisticados, como el método de multiplicadores con dirección alternante (ADMM, del inglés *alternating direction method of multipliers*) [186] o bien métodos basados en redes neuronales [187]. La aplicación de estos métodos de optimización más potentes también haría factible resolver versiones del problema de regresión no lineal (3.27) que incorporen términos de regularización con el objetivo de promover comportamientos adecuados sobre las HAFs, como por ejemplo la rareza o propiedades estadísticas de interés.

Respecto a las aplicaciones consideradas en esta tesis, existen diversos aspectos que pueden perseguirse en desarrollos futuros. En lo que respecta a la imputación de datos en series temporales no estacionarias, una de las primeras cuestiones a considerar es la posibilidad de extender el método HaLI a series temporales *multidimensionales*. Para este caso, la tarea de imputación inicial debería realizarse mediante algoritmos diseñados para tratar con datos multivariados, como los propuestos en [188] y [189]. Las etapas posteriores de descomposición e interpolación pueden modificarse con el objetivo de aprovechar la información contenida en los diferentes canales de datos. El análisis de series temporales multivariadas es de especial interés en el ámbito biomédico, ya que es común realizar registros simultáneos de diferentes señales durante estudios clínicos como la polisomnografía o el registro simultáneo de señales como las derivaciones del electrocardiograma, la fotopletomografía y la medición de la presión sanguínea durante el estudio hemodinámico de pacientes en terapia intensiva.

El análisis de señales fonatorias no estacionarias resulta otra línea de trabajo interesante. Además de la señal de voz, otras señales como la vibración de piel de cuello (VPC) o el electroglotograma (EGG) aportan información relevante acerca del estado del aparato fonador, en particular respecto a la integridad de las cuerdas vocales y la glotis. En este sentido, aplicar los métodos de análisis basados en el modelo no armónico adaptativo a este tipo de señales resulta interesante para evaluar la factibilidad de describir diferentes situaciones fisiológicas y patológicas a partir de los parámetros del modelo y de las medidas de variabilidad de forma de onda que fueron descritas en esta tesis. Por otro lado, continuar con el análisis de la variabilidad de la forma de onda de la voz sigue siendo de especial interés, ya que es posible aplicar este análisis a situaciones patológicas que no fueron consideradas en esta tesis.

Finalmente, cabe aclarar que los métodos desarrollados en esta tesis son aplicables a las llamadas señales no estacionarias *oscilatorias*. Éstas se caracterizan por presentar un patrón oscilatorio “suave”, es decir, sin la presencia de picos abruptos. Sin embargo, diversas señales de interés para el ámbito clínico, como el electrocardiograma, son de naturaleza “puntiaguda” ó *spiky*, y los modelos para señales oscilatorias pueden no adecuarse de la mejor manera a señales con estas características, lo cual fue discutido brevemente al final de la Sec. 2.6. Resulta interesante, entonces, formular nuevos modelos fenomenológicos que consideren la condición de señales *spiky*, ya sea de manera aislada o en superposición con componentes oscilatorios cercanos a los estudiados a través del ANHM convencional.

Acrónimos

En el presente documento, utilizaremos los acrónimos en inglés, por ser la forma más difundida en la bibliografía.

AM-FM	Modulado en amplitud y modulado en frecuencia
FT	Transformada de Fourier
HT	Transformada de Hilbert
TFR	Representación tiempo-frecuencia
STFT	Transformada de Fourier de tiempo corto
CWT	Transformada ondita continua
AHM	Modelo armónico adaptativo
ANHM	Modelo no armónico adaptativo
WSF	Función de forma de onda
IMF	Función modales intrínsecas
AIC	Criterio de información de Akaike
BIC	Criterio de información bayesiano
MDL	Mínima longitud de descripción
MMD	Descomposición modal multirresolución
SAMD	Descomposición modal adaptable a la forma
tvWSE	Extracción de la forma de onda variante en el tiempo
HAF	Función de amplitud armónica
LR	Regresión lineal
NLR	Regresión no lineal
MSE	Error cuadrático medio
MAE	Error absoluto medio
NMAE	Error absoluto medio normalizado
HaLI	Interpolación a nivel de los armónicos
SVM	Máquina de soporte vectorial
RF	Bosque aleatorio
RBF	Función de base radial
EP	Enfermedad de Parkinson
kFCV	Validación cruzada de k pliegues
LOSO CV	Validación cruzada dejando un sujeto afuera
SNR	Relación señal a ruido
ECG	Electrocardiograma
EEG	Electroencefalograma
ABP	Presión arterial sanguínea
PPG	Fotopleletismografía
ACC	Señal de acelerometría

Notación

$x(t)$	Señal en tiempo continuo
$x = x[n]$	Serie de datos discreta
$A(t)$	Modulación en amplitud
$\phi(t)$	Modulación en fase
$\tilde{x}(t)$	Señal demodulada en amplitud en tiempo continuo
$s(\cdot)$	Función de forma de onda
$\alpha_\ell(t)$	Función de amplitud del ℓ -ésimo armónico
$\hat{x}(f) = \mathcal{F}\{x\}(f)$	Transformada de Fourier de $x(t)$
$\mathcal{H}\{x\}(t)$	Transformada de Hilbert de $x(t)$
$ z $	Módulo del número complejo z
$\angle z$	Argumento de un número complejo z
$\ v\ _2$	Norma-2 o norma euclidiana de v
$\ v\ _1$	Norma-1 de v
$F_x^g(t, f)$	Transformada de Fourier de tiempo corto de x con la ventana g
$W_x^\psi(t, s)$	Transformada ondita continua de x con la ondita ψ
σ	Varianza de la ventana gaussiana
κ	Umbral adaptativo para la STFT
r^*	Orden óptimo para la WSF
χ_I	Función característica del conjunto I
$\#\mathbb{A}$	Cardinalidad del conjunto \mathbb{A}
$\text{GP}(\mu, k(x, x'))$	Proceso gaussiano con media μ y función de covarianza $k(x, x')$
$TV(\cdot)$	Variación total

Lista de Publicaciones

- [1] **J. Ruiz**, G. Schlotthauer, M. A. Colominas, Denoising biomedical signals with wave-shape functions, in: Lopez, N.M., Tello, E. (eds) *Advances in Bioengineering and Clinical Engineering. SABI 2022. IFMBE Proceedings*, vol 105., Springer, Cham, 2024, pp. 414–421. doi:10.1007/978-3-031-51723-5_51.
- [2] **J. Ruiz**, M. A. Colominas, Wave-shape function model order estimation by trigonometric regression, *Signal Processing* 197 (2022) 108543. doi:10.1016/j.sigpro.2022.108543.
- [3] **J. Ruiz**, G. Schlotthauer, L. Vignolo, M. A. Colominas, Fully adaptive time-varying wave-shape model: Applications in biomedical signal processing, *Signal Processing* 214 (2024) 109258. doi:10.1016/j.sigpro.2023.109258.
- [4] **J. Ruiz**, H.-T. Wu, M. A. Colominas, Enhancing missing data imputation of non-stationary oscillatory signals with harmonic decomposition, *IEEE transactions on signal processing* (2024). doi:10.1109/TSP.2024.3508468.
- [5] **J. Ruiz**, G. Schlotthauer, L. Vignolo, M. A. Colominas, Adaptive oscillatory signal modeling with applications to automatic voice analysis, *IEEE transactions on audio, speech, and language processing*, *En etapa de revisión* (2024).

Bibliografía

- [6] P. X. Joris, C. E. Schreiner, A. Rees, Neural processing of amplitude-modulated sounds, *Physiological reviews* 84 (2) (2004) 541–577. doi:10.1152/physrev.00029.2003.
- [7] Z. Feng, M. J. Zuo, Vibration signal models for fault diagnosis of planetary gearboxes, *Journal of Sound and Vibration* 331 (22) (2012) 4919–4939. doi:https://doi.org/10.1016/j.jsv.2012.05.039.
- [8] H. Roder, Amplitude, phase, and frequency modulation, *Proceedings of the Institute of Radio Engineers* 19 (12) (1931) 2145–2176. doi:10.1109/JRPROC.1931.222283.
- [9] M. Malik, A. J. Camm, Heart rate variability, *Clinical cardiology* 13 (8) (1990) 570–576.
- [10] V. Bajaj, R. B. Pachori, Classification of seizure and nonseizure EEG signals using empirical mode decomposition, *IEEE Transactions on Information Technology in Biomedicine* 16 (6) (2011) 1135–1142. doi:10.1109/TITB.2011.2181403.
- [11] C. Yeh, A. Roebel, X. Rodet, Multiple fundamental frequency estimation and polyphony inference of polyphonic music signals, *IEEE transactions on audio, speech, and language processing* 18 (6) (2009) 1116–1126. doi:10.1109/TASL.2009.2030006.
- [12] M. Muller, D. P. Ellis, A. Klapuri, G. Richard, Signal processing for music analysis, *IEEE Journal of selected topics in signal processing* 5 (6) (2011) 1088–1110. doi:10.1109/JSTSP.2011.2112333.
- [13] L. Regnier, G. Peeters, Singing voice detection in music tracks using direct voice vibrato detection, in: *2009 IEEE international conference on acoustics, speech and signal processing, IEEE, 2009*, pp. 1685–1688. doi:10.1109/ICASSP.2009.4959926.
- [14] J. Han, M. van der Baan, Empirical mode decomposition for seismic time-frequency analysis, *Geophysics* 78 (2) (2013) O9–O19. doi:10.1190/geo2012-0199.1.
- [15] P. Wang, J. Gao, Z. Wang, Time-frequency analysis of seismic data using synchrosqueezing transform, *IEEE Geoscience and Remote Sensing Letters* 11 (12) (2014) 2042–2044. doi:10.1109/LGRS.2014.2317578.

- [16] B. P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., Observation of gravitational waves from a binary black hole merger, *Physical review letters* 116 (6) (2016) 061102. doi:10.1103/PhysRevLett.116.061102.
- [17] K. Jotania, S. Valluri, S. V. Dhurandhar, A study of the gravitational wave form from pulsars, *Astronomy and Astrophysics* 306 (1996) 317–325.
- [18] Y. Yang, X. Dong, Z. Peng, W. Zhang, G. Meng, Vibration signal analysis using parameterized time–frequency method for features extraction of varying-speed rotary machinery, *Journal of Sound and Vibration* 335 (2015) 350–366. doi:10.1016/j.jsv.2014.09.025.
- [19] Z. Peng, W. T. Peter, F. Chu, An improved Hilbert–Huang transform and its application in vibration signal analysis, *Journal of sound and vibration* 286 (1-2) (2005) 187–205. doi:10.1016/j.jsv.2004.10.005.
- [20] L. Aguiar-Conraria, N. Azevedo, M. J. Soares, Using wavelets to decompose the time–frequency effects of monetary policy, *Physica A: Statistical mechanics and its Applications* 387 (12) (2008) 2863–2878. doi:10.1016/j.physa.2008.01.063.
- [21] N. E. Huang, M.-L. Wu, W. Qu, S. R. Long, S. S. Shen, Applications of Hilbert–Huang transform to non-stationary financial time series analysis, *Applied stochastic models in business and industry* 19 (3) (2003) 245–268. doi:10.1002/asmb.501.
- [22] E. Bedrosian, Product theorem for Hilbert transforms, Rand Corporation, 1962. doi:10.1109/PROC.1963.2308.
- [23] A. Nuttall, E. Bedrosian, On the quadrature approximation to the Hilbert transform of modulated signals, *Proceedings of the IEEE* 54 (10) (1966) 1458–1459. doi:10.1109/PROC.1966.5138.
- [24] D. Iatsenko, P. V. McClintock, A. Stefanovska, Linear and synchrosqueezed time–frequency representations revisited: Overview, standards of use, resolution, reconstruction, concentration, and algorithms, *Digital Signal Processing* 42 (2015) 1–26. doi:10.1016/j.dsp.2015.03.004.
- [25] T. Oberlin, S. Meignen, V. Perrier, Second-order synchrosqueezing transform or invertible reassignment? Towards ideal time-frequency representations, *IEEE Transactions on Signal Processing* 63 (5) (2015) 1335–1344. doi:10.1109/TSP.2015.2391077.
- [26] S. Meignen, D.-H. Pham, M. A. Colominas, On the use of short-time Fourier transform and synchrosqueezing-based demodulation for the retrieval of the modes of multicomponent signals, *Signal Processing* 178 (2021) 107760. doi:10.1016/j.sigpro.2020.107760.
- [27] P. Flandrin, Time-frequency/time-scale analysis, Academic press, 1998.

- [28] F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, H.-T. Wu, Time-frequency reassignment and synchrosqueezing: An overview, *IEEE Signal Processing Magazine* 30 (6) (2013) 32–41. doi:10.1109/MSP.2013.2265316.
- [29] I. Daubechies, *Ten lectures on wavelets*, SIAM, 1992. doi:10.1137/1.9781611970104.
- [30] S. Mallat, *A wavelet tour of signal processing: The sparse way*, Elsevier, 2008. doi:10.1016/B978-0-12-374370-1.X0001-8.
- [31] A. Calderón, Intermediate spaces and interpolation, the complex method, *Studia Mathematica* 24 (2) (1964) 113–190.
- [32] S. Meignen, T. Oberlin, P. Depalle, P. Flandrin, S. McLaughlin, Adaptive multimode signal reconstruction from time–frequency representations, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374 (2065) (2016) 20150205. doi:10.1098/rsta.2015.0205.
- [33] R. A. Carmona, W. L. Hwang, B. Torrésani, Characterization of signals by the ridges of their wavelet transforms, *IEEE Transactions on Signal Processing* 45 (10) (1997) 2586–2590. doi:10.1109/78.640725.
- [34] D. Iatsenko, P. V. McClintock, A. Stefanovska, Extraction of instantaneous frequencies from ridges in time–frequency representations of signals, *Signal Processing* 125 (2016) 290–303. doi:10.1016/j.sigpro.2016.01.024.
- [35] G. Thakur, E. Brevdo, N. S. Fučkar, H.-T. Wu, The synchrosqueezing algorithm for time-varying spectral analysis: Robustness properties and new paleoclimate applications, *Signal processing* 93 (5) (2013) 1079–1094. doi:10.1016/j.sigpro.2012.11.029.
- [36] S. Meignen, T. Oberlin, S. McLaughlin, A new algorithm for multicomponent signals analysis based on synchrosqueezing: With an application to signal sampling and denoising, *IEEE Transactions on Signal Processing* 60 (11) (2012) 5787–5798. doi:10.1109/TSP.2012.2212891.
- [37] M. A. Colominas, S. Meignen, D.-H. Pham, Fully adaptive ridge detection based on STFT phase information, *IEEE Signal Processing Letters* 27 (2020) 620–624. doi:10.1109/LSP.2020.2987166.
- [38] K. Kodera, *Analyse numérique de signaux géophysiques non-stationnaires*, These de doctorat, Université de Paris VI, France (1976).
- [39] K. Kodera, C. De Villedary, R. Gendrin, A new method for the numerical analysis of non-stationary signals, *Physics of the Earth and Planetary Interiors* 12 (2-3) (1976) 142–150. doi:10.1016/0031-9201(76)90044-3.
- [40] F. Auger, P. Flandrin, Improving the readability of time-frequency and time-scale representations by the reassignment method, *IEEE Transactions on signal processing* 43 (5) (1995) 1068–1089. doi:10.1109/78.382394.

- [41] I. Daubechies, S. Maes, A nonlinear squeezing of the continuous wavelet transform based on auditory nerve models, *Wavelets in Medicine and Biology* (1996).
- [42] I. Daubechies, J. Lu, H.-T. Wu, Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool, *Applied and computational harmonic analysis* 30 (2) (2011) 243–261. doi:10.1016/j.acha.2010.08.002.
- [43] J. V. Candy, *Model-based signal processing*, John Wiley & Sons, 2005.
- [44] C. Li, M. Liang, Time–frequency signal analysis for gearbox fault diagnosis using a generalized synchrosqueezing transform, *Mechanical Systems and Signal Processing* 26 (2012) 205–217. doi:10.1016/j.ymsp.2011.07.001.
- [45] F. Li, R. Li, L. Tian, L. Chen, J. Liu, Data-driven time-frequency analysis method based on variational mode decomposition and its application to gear fault diagnosis in variable working conditions, *Mechanical Systems and Signal Processing* 116 (2019) 462–479. doi:10.1016/j.ymsp.2018.06.055.
- [46] Z. Feng, M. Liang, Y. Zhang, S. Hou, Fault diagnosis for wind turbine planetary gearboxes via demodulation analysis based on ensemble empirical mode decomposition and energy separation, *Renewable Energy* 47 (2012) 112–126. doi:10.1016/j.renene.2012.04.019.
- [47] N. A. Khan, S. Ali, A new feature for the classification of non-stationary signals based on the direction of signal energy in the time–frequency domain, *Computers in biology and medicine* 100 (2018) 10–16. doi:10.1016/j.combiomed.2018.06.018.
- [48] B. Boashash, G. Azemi, J. M. O’Toole, Time-frequency processing of nonstationary signals: Advanced TFD design to aid diagnosis with highlights from medical applications, *IEEE signal processing magazine* 30 (6) (2013) 108–119. doi:10.1109/MSP.2013.2265914.
- [49] N. Wang, E. Ambikairajah, B. G. Celler, N. H. Lovell, Feature extraction using an am-fm model for gait pattern classification, in: *2008 IEEE Biomedical Circuits and Systems Conference*, IEEE, 2008, pp. 25–28. doi:10.1109/BIOCAS.2008.4696865.
- [50] J. T. Astola, K. O. Egiazarian, G. I. Khlopov, S. I. Khomenko, I. V. Kurbatov, V. Y. Morozov, A. V. Totsky, Application of bispectrum estimation for time-frequency analysis of ground surveillance doppler radar echo signals, *IEEE Transactions on Instrumentation and Measurement* 57 (9) (2008) 1949–1957. doi:10.1109/TIM.2008.917192.
- [51] H.-T. Wu, Instantaneous frequency and wave shape functions (i), *Applied and Computational Harmonic Analysis* 35 (2) (2013) 181–199. doi:10.1016/j.acha.2012.08.008.
- [52] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, H. H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proceedings of the Royal Society*

- of London. Series A: mathematical, physical and engineering sciences 454 (1971) (1998) 903–995. doi:10.1098/rspa.1998.0193.
- [53] H.-T. Wu, H.-K. Wu, C.-L. Wang, Y.-L. Yang, W.-H. Wu, T.-H. Tsai, H.-H. Chang, Modeling the pulse signal by wave-shape function and analyzing by synchrosqueezing transform, *PloS one* 11 (6) (2016) e0157135. doi:10.1371/journal.pone.0157135.
- [54] Y.-T. Lin, J. Malik, H.-T. Wu, Wave-shape oscillatory model for nonstationary periodic time series analysis, *Foundations of Data Science* 3 (2) (2021) 99. doi:10.3934/fods.2021009.
- [55] L. Su, H.-T. Wu, Extract fetal ECG from single-lead abdominal ECG by de-shape short time Fourier transform and nonlocal median, *Frontiers in Applied Mathematics and Statistics* 3 (2017) 2. doi:10.3389/fams.2017.00002.
- [56] W. K. Huang, Y.-M. Chung, Y.-B. Wang, J. E. Mandel, H.-T. Wu, Airflow recovery from thoracic and abdominal movements using synchrosqueezing transform and locally stationary gaussian process regression, *Computational Statistics & Data Analysis* 174 (2022) 107384. doi:10.1016/j.csda.2021.107384.
- [57] Y. Lu, H.-T. Wu, J. Malik, Recycling cardiogenic artifacts in impedance pneumography, *Biomedical Signal Processing and Control* 51 (2019) 162–170. doi:10.1016/j.bspc.2019.02.027.
- [58] H.-T. Wu, R. Talmon, Y.-L. Lo, Assess sleep stage by modern signal processing techniques, *IEEE Transactions on Biomedical Engineering* 62 (4) (2014) 1159–1168. doi:10.1109/TBME.2014.2375292.
- [59] M. A. Colominas, H.-T. Wu, Decomposing non-stationary signals with time-varying wave-shape functions, *IEEE Transactions on Signal Processing* 69 (2021) 5094–5104.
- [60] H. Yang, Multiresolution mode decomposition for adaptive time series analysis, *Applied and Computational Harmonic Analysis* 52 (2021) 25–62. doi:10.1016/j.acha.2019.09.006.
- [61] H.-T. Wu, Current state of nonlinear-type time–frequency analysis and applications to high-frequency biomedical signals, *Current Opinion in Systems Biology* 23 (2020) 8–21. doi:10.1016/j.coisb.2020.07.013.
- [62] V. Sucic, N. Saulig, B. Boashash, Estimating the number of components of a multicomponent nonstationary signal using the short-term time-frequency Rényi entropy, *EURASIP Journal on Advances in Signal Processing* 2011 (1) (2011) 1–11. doi:10.1186/1687-6180-2011-125.
- [63] V. Bruni, M. Tartaglione, D. Vitulano, A signal complexity-based approach for AM–FM signal modes counting, *Mathematics* 8 (12) (2020) 2170. doi:10.3390/math8122170.
- [64] J. Ding, V. Tarokh, Y. Yang, Model selection techniques: An overview, *IEEE Signal Processing Magazine* 35 (6) (2018) 16–34. doi:10.1109/MSP.2018.2867638.

- [65] P. Stoica, Y. Selen, Model-order selection: A review of information criterion rules, *IEEE Signal Processing Magazine* 21 (4) (2004) 36–47. doi:10.1109/MSP.2004.1311138.
- [66] C. M. Bishop, *Pattern recognition and machine learning*, Springer google schola 2 (2006) 645–678.
- [67] H. Akaike, Information theory and an extension of the maximum likelihood principle, in: *Selected papers of Hirotugu Akaike*, Springer, 1998, pp. 199–213. doi:10.1007/978-1-4612-1694-0_15.
- [68] G. Schwarz, Estimating the dimension of a model, *The annals of statistics* (1978) 461–464 doi:https://www.jstor.org/stable/2958889.
- [69] J. Rissanen, Modeling by shortest data description, *Automatica* 14 (5) (1978) 465–471. doi:10.1016/0005-1098(78)90005-5.
- [70] M. Wax, T. Kailath, Detection of signals by information theoretic criteria, *IEEE Transactions on acoustics, speech, and signal processing* 33 (2) (1985) 387–392. doi:10.1109/TASSP.1985.1164557.
- [71] R. L. Eubank, *Nonparametric regression and spline smoothing*, CRC press, 1999. doi:10.1093/biomet/77.1.1.
- [72] D. L. Donoho, J. M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, *biometrika* 81 (3) (1994) 425–455. doi:10.1093/biomet/81.3.425.
- [73] D.-H. Pham, S. Meignen, A novel thresholding technique for the denoising of multicomponent signals, in: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2018, pp. 4004–4008. doi:10.1109/ICASSP.2018.8462216.
- [74] G. H. Golub, M. Heath, G. Wahba, Generalized cross-validation as a method for choosing a good ridge parameter, *Technometrics* 21 (2) (1979) 215–223. doi:10.1080/00401706.1979.10489751.
- [75] R. L. Eubank, P. Speckman, Curve fitting by polynomial-trigonometric regression, *Biometrika* 77 (1) (1990) 1–9. doi:10.1093/biomet/77.1.1.
- [76] X. Wang, An AIC type estimator for the number of sinusoids, *Journal of Time Series Analysis* 14 (4) (1993) 433–440. doi:10.1111/j.1467-9892.1993.tb00155.x.
- [77] B. G. Quinn, Estimating the number of terms in a sinusoidal regression, *Journal of time series analysis* 10 (1) (1989) 71–75. doi:10.1111/j.1467-9892.1989.tb00016.x.
- [78] L. Kavalieris, E. Hannan, Determining the number of terms in a trigonometric regression, *Journal of time series analysis* 15 (6) (1994) 613–625. doi:10.1111/j.1467-9892.1994.tb00216.x.

- [79] E. J. Hannan, L. Kavalieris, Multivariate linear time series models, *Advances in Applied Probability* (1984) 492–561 doi:10.2307/1427286.
- [80] C. L. Herry, M. Frasch, A. J. Seely, H.-T. Wu, Heart beat classification from single-lead ECG using the synchrosqueezing transform, *Physiological measurement* 38 (2) (2017) 171. doi:10.1088/1361-6579/aa5070.
- [81] P. H. Charlton, J. M. Harana, S. Vennin, Y. Li, P. Chowiencyk, J. Alastruey, Pulse wave database (PWDB): A database of arterial pulse waves representative of healthy adults, version 0.2.0, zenodo (2019). doi:10.5281/zenodo.3275625. URL <https://zenodo.org/record/3275625#.YSk9xt-vHIU>
- [82] P. H. Charlton, J. Mariscal Harana, S. Vennin, Y. Li, P. Chowiencyk, J. Alastruey, Modeling arterial pulse waves in healthy aging: A database for in silico evaluation of hemodynamics and pulse wave indexes, *American Journal of Physiology-Heart and Circulatory Physiology* 317 (5) (2019) H1062–H1085. doi:10.1152/ajpheart.00218.2019.
- [83] P. W. Kamen, H. Krum, A. M. Tonkin, Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans, *Clinical science* 91 (2) (1996) 201–208. doi:10.1042/cs0910201.
- [84] R. G. Baraniuk, P. Flandrin, A. J. Janssen, O. J. Michel, Measuring time-frequency information content using the Rényi entropies, *IEEE Transactions on Information theory* 47 (4) (2001) 1391–1409. doi:10.1109/18.923723.
- [85] S. Meignen, M. A. Colominas, D.-H. Pham, On the use of Rényi entropy for optimal window size computation in the short-time Fourier transform, in: *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2020, pp. 5830–5834. doi:10.1109/ICASSP40776.2020.9053392.
- [86] M. Friedman, The use of ranks to avoid the assumption of normality implicit in the analysis of variance, *Journal of the american statistical association* 32 (200) (1937) 675–701. doi:10.1080/01621459.1937.10503522.
- [87] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, H. E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals, *Circulation* 101 (23) (2000) e215–e220. doi:10.1161/01.CIR.101.23.e215.
- [88] N. Iyengar, C. Peng, R. Morin, A. L. Goldberger, L. A. Lipsitz, Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics, *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* 271 (4) (1996) R1078–R1084. doi:10.1152/ajpregu.1996.271.4.R1078.
- [89] C.-Y. Lin, L. Su, H.-T. Wu, Wave-shape function analysis: When cepstrum meets time–frequency analysis, *Journal of Fourier Analysis and Applications* 24 (2018) 451–505. doi:10.1007/s00041-017-9523-0.

- [90] R. Talmon, I. Cohen, S. Gannot, R. R. Coifman, Diffusion maps for signal processing: A deeper look at manifold-learning techniques based on kernels and graphs, *IEEE signal processing magazine* 30 (4) (2013) 75–86. doi:10.1109/MSP.2013.2250353.
- [91] F. N. Fritsch, R. E. Carlson, Monotone piecewise cubic interpolation, *SIAM Journal on Numerical Analysis* 17 (2) (1980) 238–246. doi:10.1137/0717021.
- [92] G. E. Box, G. M. Jenkins, G. C. Reinsel, G. M. Ljung, *Time series analysis: Forecasting and control*, John Wiley & Sons, 2015.
- [93] E. A. Yildirim, S. J. Wright, Warm-start strategies in interior-point methods for linear programming, *SIAM Journal on Optimization* 12 (3) (2002) 782–810. doi:10.1137/S1052623400369235.
- [94] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, *Journal of the Society for Industrial and Applied Mathematics* 11 (2) (1963) 431–441. doi:10.1137/0111030.
- [95] R. Killick, P. Fearnhead, I. A. Eckley, Optimal detection of changepoints with a linear computational cost, *Journal of the American Statistical Association* 107 (500) (2012) 1590–1598. doi:10.1080/01621459.2012.737745.
- [96] N. J. Stevenson, K. Tapani, L. Lauronen, S. Vanhatalo, A dataset of neonatal EEG recordings with seizure annotations, *Scientific data* 6 (1) (2019) 1–8. doi:10.1038/sdata.2019.39.
- [97] G. B. Bucklar, V. Kaplan, K. E. Bloch, Signal processing technique for non-invasive real-time estimation of cardiac output by inductance cardiography (thoracocardiography), *Medical and Biological Engineering and Computing* 41 (2003) 302–309. doi:10.1007/BF02348435.
- [98] F. Nolle, F. Badura, J. Catlett, R. Bowser, M. Sketch, CREI-GARD, a new concept in computerized arrhythmia monitoring systems, *Computers in Cardiology* 13 (1) (1986) 515–518.
- [99] D. B. Rubin, Inference and missing data, *Biometrika* 63 (3) (1976) 581–592. doi:10.1093/biomet/63.3.581.
- [100] J. G. Ibrahim, H. Chu, M.-H. Chen, Missing data in clinical studies: Issues and methods, *Journal of clinical oncology* 30 (26) (2012) 3297. doi:10.1200/JCO.2011.38.7589.
- [101] R. J. Little, D. B. Rubin, *Statistical analysis with missing data*, Vol. 793, John Wiley & Sons, 2019. doi:10.1038/npre.2008.2298.1.
- [102] E. Damsleth, Interpolating missing values in a time series, *Scandinavian Journal of Statistics* (1980) 33–39doi:https://www.jstor.org/stable/4615768.
- [103] M. Pourahmadi, Estimation and interpolation of missing values of a stationary time series, *Journal of Time Series Analysis* 10 (2) (1989) 149–169. doi:10.1111/j.1467-9892.1989.tb00021.x.

- [104] Y. Walter, J. Kihoro, K. Athiany, H. Kibunja, Imputation of incomplete non-stationary seasonal time series data, *Math. Theory Model* 3 (2013) 142–154.
- [105] S. Oh, Multiple imputation on missing values in time series data, Master’s thesis, Duke University (2015).
- [106] F. Lobato, C. Sales, I. Araujo, V. Tadaiesky, L. Dias, L. Ramos, A. Santana, Multi-objective genetic algorithm for missing data imputation, *Pattern Recognition Letters* 68 (2015) 126–131. doi:10.1016/j.patrec.2015.08.023.
- [107] A. S. Dhevi, Imputing missing values using inverse distance weighted interpolation for time series data, in: 2014 Sixth international conference on advanced computing (ICoAC), IEEE, 2014, pp. 255–259. doi:10.1109/ICoAC.2014.7229721.
- [108] M. Amiri, R. Jensen, Missing data imputation using fuzzy-rough methods, *Neurocomputing* 205 (2016) 152–164. doi:10.1016/j.neucom.2016.04.015.
- [109] P. G. Papaioannou, R. Talmon, I. G. Kevrekidis, C. Siettos, Time-series forecasting using manifold learning, radial basis function interpolation, and geometric harmonics, *Chaos: An Interdisciplinary Journal of Nonlinear Science* 32 (8) (2022) 083113. doi:10.1063/5.0094887.
- [110] M. West, J. Harrison, Bayesian forecasting and dynamic models, Springer Science & Business Media, 2006.
- [111] S. Sridevi, S. Rajaram, C. Parthiban, S. SibiArasan, C. Swadhikar, Imputation for the analysis of missing values and prediction of time series data, in: 2011 international conference on recent trends in information Technology (ICRTIT), IEEE, 2011, pp. 1158–1163. doi:10.1109/ICRTIT.2011.5972466.
- [112] A. Moghtaderi, P. Borgnat, P. Flandrin, Gap-filling by the empirical mode decomposition, in: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2012, pp. 3821–3824. doi:10.1109/ICASSP.2012.6288750.
- [113] T. Sidekerskiene, R. Damasevicius, Reconstruction of missing data in synthetic time series using EMD, in: CEUR Workshop Proceedings, Vol. 1712, 2016, pp. 7–12.
- [114] W. Cao, D. Wang, J. Li, H. Zhou, L. Li, Y. Li, BRITS: Bidirectional recurrent imputation for time series, *Advances in neural information processing systems* 31 (2018).
- [115] V. Fortuin, D. Baranchuk, G. Rätsch, S. Mandt, GP-VAE: Deep probabilistic time series imputation, in: International conference on artificial intelligence and statistics, PMLR, 2020, pp. 1651–1661.
- [116] R. Qin, Y. Wang, ImputeGAN: Generative adversarial network for multivariate time series imputation, *Entropy* 25 (1) (2023) 137. doi:10.3390/e25010137.
- [117] B. Xu, Y. Sheng, P. Li, Q. Cheng, J. Wu, Causes and classification of EMD mode mixing, *Vibroengineering Procedia* 22 (2019) 158–164. doi:10.21595/vp.2018.20250.

- [118] F. Takens, Detecting strange attractors in turbulence, in: *Dynamical systems and turbulence*, Warwick 1980, Springer, 1981, pp. 366–381. doi:10.1007/BFb0091924.
- [119] A. Meynard, H.-T. Wu, An efficient forecasting approach to reduce boundary effects in real-time time-frequency analysis, *IEEE Transactions on Signal Processing* 69 (2021) 1653–1663. doi:10.1109/TSP.2021.3062181.
- [120] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, *Journal of fluid mechanics* 656 (2010) 5–28. doi:10.1017/S0022112010001217.
- [121] M. O. Williams, I. G. Kevrekidis, C. W. Rowley, A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition, *Journal of Nonlinear Science* 25 (2015) 1307–1346. doi:10.1007/s00332-015-9258-5.
- [122] C. E. Rasmussen, Gaussian processes in machine learning, in: *Summer school on machine learning*, Springer, 2003, pp. 63–71. doi:10.1007/978-3-540-28650-9_4.
- [123] A. M. De Livera, R. J. Hyndman, R. D. Snyder, Forecasting time series with complex seasonal patterns using exponential smoothing, *Journal of the American statistical association* 106 (496) (2011) 1513–1527. doi:10.1198/jasa.2011.tm09771.
- [124] T. Y. Hou, Z. Shi, Data-driven time–frequency analysis, *Applied and Computational Harmonic Analysis* 35 (2) (2013) 284–308. doi:10.1016/j.acha.2012.10.001.
- [125] R. Killick, M. I. Knight, G. P. Nason, M. A. Nunes, I. A. Eckley, Automatic locally stationary time series forecasting with application to predicting UK gross value added time series, *Journal of the Royal Statistical Society Series C: Applied Statistics* (2024) qlae043doi:10.1093/jrsssc/qlae043.
- [126] A. Parekh, I. Selesnick, A. Baroni, O. Bubu, A. Varga, D. Rapoport, I. Ayappa, E. Blessing, R. Osorio, Nonlinear smoothing of core body temperature data with random gaps and outliers (DRAGO), *Biomedical Signal Processing: Innovation and Applications* (2021) 63–84doi:10.1007/978-3-030-67494-6_3.
- [127] O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein, R. B. Altman, Missing value estimation methods for DNA microarrays, *Bioinformatics* 17 (6) (2001) 520–525. doi:10.1093/bioinformatics/17.6.520.
- [128] J.-C. Hua, F. Noorian, D. Moss, P. H. Leong, G. H. Gunaratne, High-dimensional time series prediction using kernel-based Koopman mode regression, *Nonlinear Dynamics* 90 (2017) 1785–1806. doi:10.1007/s11071-017-3764-y.
- [129] C. Williams, C. Rasmussen, Gaussian processes for regression, *Advances in neural information processing systems* 8 (1995).

- [130] P. J. Brockwell, R. A. Davis, Time series: Theory and methods, Springer science & business media, 2009.
- [131] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, D. Gorinevsky, An interior-point method for large-scale ℓ_1 -regularized least squares, *IEEE journal of selected topics in signal processing* 1 (4) (2007) 606–617. doi:10.1109/JSTSP.2007.910971.
- [132] G. P. Nason, R. Von Sachs, G. Kroisandt, Wavelet processes and adaptive estimation of the evolutionary wavelet spectrum, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 62 (2) (2000) 271–292. doi:10.1111/1467-9868.00231.
- [133] P. Fryzlewicz, S. Van Bellegem, R. Von Sachs, Forecasting non-stationary time series by wavelet process modelling, *Annals of the Institute of Statistical Mathematics* 55 (4) (2003) 737–764. doi:10.1007/BF02523391.
- [134] M. L. Stein, Interpolation of spatial data, Springer Series in Statistics (1999).
- [135] R. Killick, M. I. Knight, G. P. Nason, I. A. Eckley, The local partial autocorrelation function and some applications, *Electronic Journal of Statistics* (2020). doi:10.1214/20-EJS1748.
- [136] Y. Xie, J. Yu, B. Ranney, Forecasting using locally stationary wavelet processes, *Journal of Statistical Computation and Simulation* 79 (9) (2009) 1067–1082. doi:10.1080/00949650802087003.
- [137] G. Nason, A test for second-order stationarity and approximate confidence intervals for localized autocovariances for locally stationary time series, *Journal of the Royal Statistical Society Series B: Statistical Methodology* (2013) 879–904 doi:10.1111/rssb.12015.
- [138] I. Obeid, I. Selesnick, J. Picone, Biomedical signal processing: Innovation and applications, Springer, 2021. doi:10.1007/978-3-030-67494-6.
- [139] D. J. Stekhoven, P. Bühlmann, MissForest—non-parametric missing value imputation for mixed-type data, *Bioinformatics* 28 (1) (2012) 112–118. doi:10.1093/bioinformatics/btr597.
- [140] S. Van Buuren, K. Groothuis-Oudshoorn, mice: Multivariate imputation by chained equations in R, *Journal of statistical software* 45 (2011) 1–67. doi:10.18637/jss.v045.i03.
- [141] M. A. Colominas, S. Meignen, D.-H. Pham, Fully adaptive ridge detection based on STFT phase information, *IEEE Signal Processing Letters* 27 (2020) 620–624. doi:10.1109/LSP.2020.2987166.
- [142] N. Laurent, S. Meignen, A novel ridge detector for nonstationary multicomponent signals: Development and application to robust mode retrieval, *IEEE Transactions on Signal Processing* 69 (2021) 3325–3336. doi:10.1109/TSP.2021.3085113.

- [143] P. Borgnat, P. Flandrin, P. Honeine, C. Richard, J. Xiao, Testing stationarity with surrogates: A time-frequency approach, *IEEE Transactions on Signal Processing* 58 (7) (2010) 3459–3470. doi:10.1109/TSP.2010.2043971.
- [144] Z. Chen, H.-T. Wu, Disentangling modes with crossover instantaneous frequencies by synchrosqueezed chirplet transforms, from theory to application, *Applied and Computational Harmonic Analysis* 62 (2023) 84–122. doi:10.1016/j.acha.2022.08.004.
- [145] M. Weber, M. Denk, Imputation of cross-country time series: Techniques and evaluation, in: *European Conference on Quality in Official Statistics*, 2010, pp. 1–10.
- [146] D. B. Rubin, Multiple imputation, in: *Flexible Imputation of Missing Data*, Second Edition, Chapman and Hall/CRC, 2018, pp. 29–62.
- [147] T. Paiva, A. Chakraborty, J. Reiter, A. Gelfand, Imputation of confidential data sets with spatial locations using disease mapping models, *Statistics in medicine* 33 (11) (2014) 1928–1945. doi:10.1002/sim.6078.
- [148] A. C. Gilbert, R. Sonthalia, Unsupervised metric learning in presence of missing data, in: *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2018, pp. 313–321. doi:10.1109/ALLERTON.2018.8635955.
- [149] H. Wakita, Direct estimation of the vocal tract shape by inverse filtering of acoustic speech waveforms, *IEEE Transactions on Audio and Electroacoustics* 21 (5) (1973) 417–427. doi:10.1109/TAU.1973.1162506.
- [150] S. Malik, G. Enzner, State-space frequency-domain adaptive filtering for nonlinear acoustic echo cancellation, *IEEE Transactions on audio, speech, and language processing* 20 (7) (2012) 2065–2079. doi:10.1109/TASL.2012.2196512.
- [151] J. W. Pitton, K. Wang, B.-H. Juang, Time-frequency analysis and auditory modeling for automatic recognition of speech, *Proceedings of the IEEE* 84 (9) (1996) 1199–1215. doi:10.1109/5.535241.
- [152] M. Kepesi, L. Weruaga, Adaptive chirp-based time–frequency analysis of speech signals, *Speech communication* 48 (5) (2006) 474–492. doi:10.1016/j.specom.2005.08.004.
- [153] L. Verde, G. De Pietro, G. Sannino, Voice disorder identification by using machine learning techniques, *IEEE access* 6 (2018) 16246–16255. doi:10.1109/ACCESS.2018.2816338.
- [154] S. Fex, Perceptual evaluation, *Journal of Voice* 6 (2) (1992) 155–158. doi:10.1016/S0892-1997(05)80130-4.
- [155] I. V. Bele, Reliability in perceptual analysis of voice quality, *Journal of Voice* 19 (4) (2005) 555–573. doi:10.1016/j.jvoice.2004.08.008.

- [156] R. Behroozmand, F. Almasganj, Optimal selection of wavelet-packet-based features using genetic algorithm in pathological assessment of patients' speech signal with unilateral vocal fold paralysis, *Computers in Biology and Medicine* 37 (4) (2007) 474–485. doi:10.1016/j.compbiomed.2006.08.016.
- [157] A. Benba, A. Jilbab, A. Hammouch, Discriminating between patients with Parkinson's and neurological diseases using cepstral analysis, *IEEE transactions on neural systems and rehabilitation engineering* 24 (10) (2016) 1100–1108. doi:10.1109/TNSRE.2016.2533582.
- [158] J. M. Miramont, J. F. Restrepo, J. Codino, C. Jackson-Menaldi, G. Schlotthauer, Voice signal typing using a pattern recognition approach, *Journal of Voice* (2020). doi:10.1016/j.jvoice.2020.03.006.
- [159] J. M. Miramont, M. A. Colominas, G. Schlotthauer, Emulating perceptual evaluation of voice using scattering transform based features, *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2022). doi:10.1109/TASLP.2022.3178239.
- [160] S.-H. Fang, Y. Tsao, M.-J. Hsiao, J.-Y. Chen, Y.-H. Lai, F.-C. Lin, C.-T. Wang, Detection of pathological voice using cepstrum vectors: A deep learning approach, *Journal of voice* 33 (5) (2019) 634–641. doi:10.1016/j.jvoice.2018.02.003.
- [161] B. Sisman, J. Yamagishi, S. King, H. Li, An overview of voice conversion and its challenges: From statistical modeling to deep learning, *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 29 (2020) 132–157. doi:10.1109/TASLP.2020.3038524.
- [162] M. Alhussein, G. Muhammad, Voice pathology detection using deep learning on mobile healthcare framework, *IEEE Access* 6 (2018) 41034–41041. doi:10.1109/ACCESS.2018.2856238.
- [163] I. R. Titze, D. W. Martin, *Principles of voice production* (1998).
- [164] G. Fant, Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations, no. 2 in *Description and Analysis of Contemporary Standard Russian*, Walter de Gruyter, 1971.
- [165] M. G. Christensen, A. Jakobsson, S. H. Jensen, Joint high-resolution fundamental frequency and order estimation, *IEEE Transactions on audio, speech, and language processing* 15 (5) (2007) 1635–1644. doi:10.1109/TASL.2007.899267.
- [166] P. Mowlaee, M. G. Christensen, S. H. Jensen, New results on single-channel speech separation using sinusoidal modeling, *IEEE Transactions on audio, speech, and language processing* 19 (5) (2010) 1265–1277. doi:10.1109/TASL.2010.2089520.
- [167] T. W. Parsons, Separation of speech from interfering speech by means of harmonic selection, *The Journal of the Acoustical Society of America* 60 (4) (1976) 911–918. doi:10.1121/1.381172.

- [168] S. M. Nørholm, J. R. Jensen, M. G. Christensen, Enhancement and noise statistics estimation for non-stationary voiced speech, *IEEE Transactions on audio, speech, and language processing* 24 (4) (2016) 645–658. doi:10.1109/TASLP.2016.2514492.
- [169] S. M. Nørholm, J. R. Jensen, M. G. Christensen, Instantaneous fundamental frequency estimation with optimal segmentation for nonstationary voiced speech, *IEEE Transactions on audio, speech, and language processing* 24 (12) (2016) 2354–2367. doi:10.1109/TASLP.2016.2608948.
- [170] B. Woldert-Jokisz, Saarbruecken voice database, Tech. rep., Institut für Phonetik, Universität des Saarlandes (2007).
- [171] I. Titze, Workshop on acoustic voice analysis: summary statement. national center for voice and speech., Salt Lake City, UT, USA (1995).
- [172] M. Eye, E. Infirmity, Voice disorders database, version. 1.03 (cd-rom), Lincoln Park, NJ: Kay Elemetrics Corporation (1994).
- [173] S. Mallat, Group invariant scattering, *Communications on Pure and Applied Mathematics* 65 (10) (2012) 1331–1398. doi:10.1002/cpa.21413.
- [174] A. Lee, R. M. Gilbert, Epidemiology of Parkinson disease, *Neurologic clinics* 34 (4) (2016) 955–965. doi:10.1016/j.ncl.2016.06.012.
- [175] C. Stewart, L. Winfield, A. Hunt, S. B. Bressman, S. Fahn, A. Blitzer, M. F. Brin, Speech dysfunction in early Parkinson’s disease, *Movement disorders: official journal of the Movement Disorder Society* 10 (5) (1995) 562–565. doi:10.1002/mds.870100506.
- [176] K. Whetten-Goldstein, F. Sloan, E. Kulas, T. Cutson, M. Schenkman, The burden of Parkinson’s disease on society, family, and the individual, *Journal of the American Geriatrics Society* 45 (7) (1997) 844–849. doi:10.1111/j.1532-5415.1997.tb01512.x.
- [177] M. Little, P. McSharry, E. Hunter, J. Spielman, L. Ramig, Suitability of dysphonia measurements for telemonitoring of Parkinson’s disease, *Nature Precedings* (2008) 1–1.
- [178] J. R. Orozco-Arroyave, J. D. Arias-Londoño, J. F. Vargas-Bonilla, M. C. Gonzalez-Rátiva, E. Nöth, New Spanish speech corpus database for the analysis of people suffering from Parkinson’s disease., in: *LREC*, 2014, pp. 342–347.
- [179] J. R. Orozco-Arroyave, F. Hönig, J. D. Arias-Londoño, J. F. Vargas-Bonilla, S. Skodda, J. Ruz, E. Nöth, Automatic detection of Parkinson’s disease from words uttered in three different languages, in: *Fifteenth annual conference of the international speech communication association*, 2014, pp. 1573–1577.
- [180] A. Rueda, J. C. Vásquez-Correa, C. D. Rios-Urrego, J. R. Orozco-Arroyave, S. Krishnan, E. Nöth, Feature representation of pathophysiology of Parkinsonian dysarthria., in: *INTERSPEECH*, 2019, pp. 3048–3052.

- [181] B. Karan, S. S. Sahu, K. Mahto, Parkinson disease prediction using intrinsic mode function based features from speech signal, *Biocybernetics and Biomedical Engineering* 40 (1) (2020) 249–264. doi:10.1016/j.bbe.2019.05.005.
- [182] S. Kadiri, R. Kethireddy, P. Alku, Parkinson’s disease detection from speech using single frequency filtering cepstral coefficients, in: *Interspeech, International Speech Communication Association (ISCA)*, 2020, pp. 4971–4975. doi:10.21437/Interspeech.2020-3197.
- [183] T. Villa-Cañas, J. Arias-Londoño, J. Vargas-Bonilla, J. Orozco-Arroyave, Time-frequency approach in continuous speech for detection of Parkinson’s disease, in: *2015 20th Symposium on Signal Processing, Images and Computer Vision (STSIVA), IEEE*, 2015, pp. 1–6. doi:10.1109/STSIVA.2015.7330421.
- [184] C. Quan, K. Ren, Z. Luo, Z. Chen, Y. Ling, End-to-end deep learning approach for Parkinson’s disease detection from speech signals, *Biocybernetics and Biomedical Engineering* 42 (2) (2022) 556–574. doi:10.1016/j.bbe.2022.04.002.
- [185] M. B. Er, E. Isik, I. Isik, Parkinson’s detection based on combined CNN and LSTM using enhanced speech signals with variational mode decomposition, *Biomedical Signal Processing and Control* 70 (2021) 103006. doi:10.1016/j.bspc.2021.103006.
- [186] V. Singhal, A. Majumdar, R. K. Ward, Semi-supervised deep blind compressed sensing for analysis and reconstruction of biomedical signals from compressive measurements, *IEEE Access* 6 (2017) 545–553. doi:10.1109/ACCESS.2017.2771536.
- [187] A. Landi, P. Piaggi, M. Laurino, D. Menicucci, Artificial neural networks for nonlinear regression and classification, in: *2010 10th International Conference on Intelligent Systems Design and Applications, IEEE*, 2010, pp. 115–120. doi:10.1109/ISDA.2010.5687280.
- [188] R. E. Wilson, I. A. Eckley, M. A. Nunes, T. Park, A wavelet-based approach for imputation in nonstationary multivariate time series, *Statistics and Computing* 31 (2) (2021) 18. doi:10.1007/s11222-021-09998-2.
- [189] Y. Zhang, B. Zhou, X. Cai, W. Guo, X. Ding, X. Yuan, Missing value imputation in multivariate time series with end-to-end generative adversarial networks, *Information Sciences* 551 (2021) 67–82. doi:10.1016/j.ins.2020.11.035.

Doctorado en Ingeniería
mención inteligencia computacional, señales y sistemas

Título de la obra:

Procesamiento de señales
basado en funciones de forma de onda:
contribuciones algorítmicas y
aplicaciones biomédicas

Autor: Joaquin Victorio Ruiz
Director: Marcelo Alejandro Colominas
Codirector: Leandro Vignolo

Palabras Claves:

señales biomédicas,
señales no estacionarias,
análisis tiempo-frecuencia,
funciones de forma de onda,
modelado fenomenológico de señales,
métodos guiados por los datos,
descomposición de señales,
imputación de datos, análisis de la voz